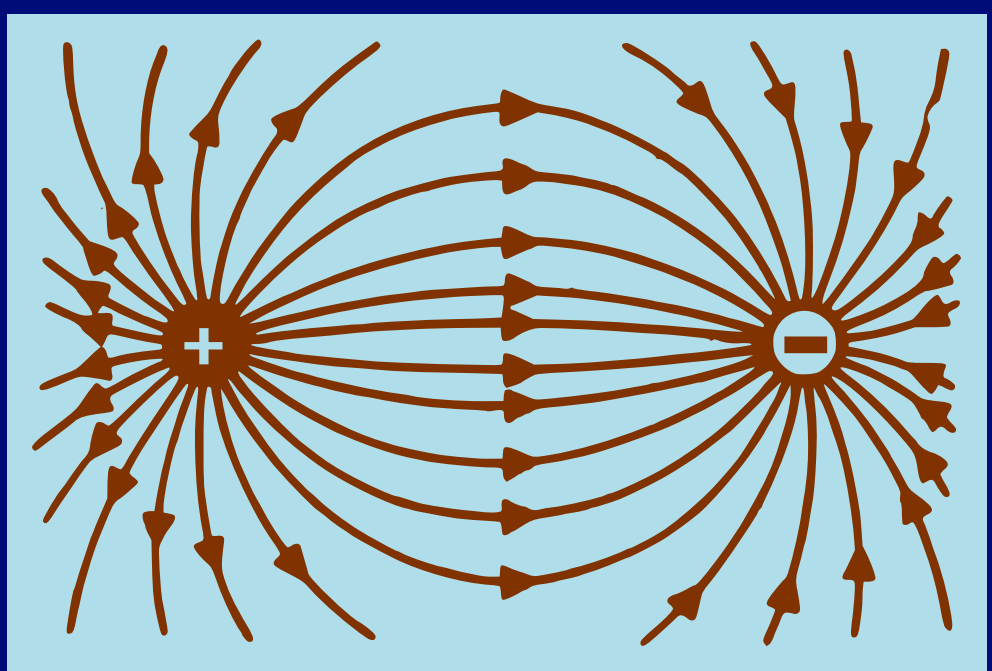
Electricity and Magnetism

A. N. Matveev



Mir Publishers Moscow

www.EngineeringBooksPDF.com

A. N. Matveev

Electricity and Magnetism

Translated from the Russian
by
Ram Wadhwa and Natalia Deineko



Mir Publishers Moscow

www.EngineeringBooksPDF.com

First published 1986
Revised from the 1983 Russian edition

© Издательство «Высшая школа», 1983
© English translation, Mir Publishers, 1986

www.EngineeringBooksPDF.com

Preface

This course reflects the present level of advancement in science and takes into account the changes in the general physics curriculum.

Since the basic concepts of the theory of relativity are known from the course on mechanics, we can base the description of electric and magnetic phenomena on the relativistic nature of a magnetic field and present the mutual correspondence and unity of electric and magnetic fields. Hence we start this book not with electrostatics but with an analysis of basic concepts associated with charge, force, and electromagnetic field. With such an approach, the information about the laws of electromagnetism, accumulated by students from school-level physics, is transformed into modern scientific knowledge, and the theory is substantiated in the light of the current state of experimental foundations of electromagnetism, taking into account the limits of applicability of the concepts involved. Sometimes, this necessitates a transgression beyond the theory of electromagnetism in the strict sense of this word. For example, the experimental substantiation of Coulomb's law for large distances is impossible without mentioning its connection with the zero rest mass of photons. Although this question is discussed fully and rigorously in quantum electrodynamics, it is expedient to describe its main features in the classical theory of electromagnetism. This helps the student to acquire a general idea of the problem and of the connection of the material of this book with that of the future courses. The latter circumstance is quite significant from the methodological point of view.

This course mainly aims at the description of the experimental substantiation of the theory of electromagnetism and the formulation of the theory in the local form, i.e. in the form of relations between physical quantities at the same point in space and time. In most cases, these relations are expressed in the form of differential equations. However, it is not the differential form but the local nature which is important. Consequently, the end product of the course are Maxwell's equations obtained as a result of generalization and mathematical formulation of experimentally established regularities. Consequently, the analysis is mainly based on induction. This, however, does not exclude the application of the deductive method but rather presumes the combination of the two methods of analysis in accordance with the principles of scientific perception of physical laws. Hence, Maxwell's equations appear in this book not only as a result of

www.EngineeringBooksPDF.com

mathematical formulation of experimentally established regularities but also as an instrument for investigating these laws.

The choice of experimental facts which can be used to substantiate the theory is not unique. Thus, the theory of electromagnetism is substantiated here with and without taking the theory of relativity into account. The former approach is preferable, since in this case the theory of relativity appears as a general space-time theory on which all physical theories must be based. Such a substantiation has become possible only within the framework of the new general physics curriculum.

An essential part of the theory is the determination of the limits of its applicability and the ranges of concepts and models employed in it. These questions, which are described in this book, are of vital importance. In particular, the analysis of the force of interaction between charges in the framework of the classical theory (i.e. without employing any quantum concepts) shows that the classical theory of electricity and magnetism cannot be applied for analyzing the interaction between isolated charged particles.

The author is grateful to his colleagues at Moscow State University as well as other universities and institutes for a fruitful discussion of the topics covered in this book. He is also indebted to Acad. A. I. Akhiezer of the Academy of Sciences of the Ukrainian SSR, Prof. N. I. Kaliteevskii and the staff of the Department of General Physics at the Leningrad State University who carefully reviewed the manuscript and made valuable comments.

A. Matveev

www.EngineeringBooksPDF.com

Contents

Preface 5
Introduction 13

Chapter 1. Charge. Field. Force

Sec. 1. Microscopic Charge Carriers 15
Classification. Electron. Proton. Neutron. What does the continuous distribution of an elementary electric charge mean? Spin and magnetic moment.

Sec. 2. Charged Bodies. Electrostatic Charging 19
Thermionic work function. Energy spectrum of electrons. Fermi energy. Contact potential difference. Electrostatic charging.

Sec. 3. Elementary Charge and Its Invariance 26
Millikan oil-drop experiment. Resonance method for measurement of charge. Nonexistence of fractional charges. Equality of positive and negative elementary charges. Invariance of charge.

Sec. 4. Electric Current 30
Motion of charges. Continuous distribution of charges. Volume charge density. Charge concentration. Surface charge density. Current density. Current through a surface.

Sec. 5. Law of Charge Conservation 35
Two aspects of the concept of charge conservation. Integral form of charge conservation law. Divergence. Gauss' theorem. Differential form of charge conservation law.

Sec. 6. Coulomb's Law 42
Experimental verification of Coulomb's law. The Cavendish method. Verification of Coulomb's law for large distances. Verification of Coulomb's law for small distances. Field form of Coulomb's law. Electric field. On the limits of applicability of the classical concept of field.

Sec. 7. Superposition Principle 50
Superposition principle for interaction of point charges. Field form of the superposition principle. Test charges. Limits of applicability of the principle of superposition.

Sec. 8. Magnetic Field 53
Inevitability of magnetic field generation due to motion of charges. Interaction between a point charge and an infinitely long charged filament. Relativistic nature of magnetic field. Forces of interaction between parallel current-carrying conductors. Unit of current. Magnetic field.

Sec. 9. Lorentz Force. Ampère Force 58
Transformation of forces. Lorentz force. Magnetic induction. Ampère force. Transformation from steady volume currents to linear currents. Magnetic field of a rectilinear current.

Sec. 10. Biot-Savart Law 63
Interaction between current elements. On experimental verification of the law of inter-

www.EngineeringBooksPDF.com

action. Field form of interaction. Biot-Savart law. Force of interaction between rectilinear currents.

Sec. 11. Field Transformation 69
Invariance of the expression for force in an electromagnetic field. Transformation of fields. Application of formulas (11.15). Field of a point charge moving uniformly in a straight line. Problems.

Chapter 2. Constant Electric Field

Sec. 12. Constant Electric Field 76
Fixed charge. The essence of the model. Limits of applicability of the model.

Sec. 13. Differential Form of Coulomb's Law 77
Gauss' theorem. Measurement of charge. Physical foundation of the validity of Gauss' theorem. Differential form of Coulomb's law. Maxwell's equation for div E. Lines of force. Sources and sinks of field E. Charge invariance.

Sec. 14. Potential Nature of Electrostatic Field 82
Work in an electric field. Potential nature of a Coulomb field. Curl of a vector. Stokes' integral theorem. Differential form of the potential nature of the field. Gradient. Scalar potential. Ambiguity of scalar potential. Normalization. Expression of work in terms of potential. Field potential of a point charge. Field potential of a system of point charges. Field potential of continuously distributed charges. Field potential of surface charges. Infinite value of the field potential of a point charge. Finite value of the potential for a continuous charge distribution with a finite density. Continuity of potential. Earnshaw's theorem.

Sec. 15. Electrostatic Field in Vacuum 94
Formulation of the problem. Direct application of Coulomb's law. Calculation of potential. Application of Gauss' theorem. Laplace's equation and Poisson's equation. A very long uniformly charged circular cylinder.

Sec. 16. Electrostatic Field in the Presence of Conductors 100
Differential form of Ohm's law. Classification of materials according to conductivity. Absence of electric field inside a conductor. Absence of volume charges inside a conductor. Electrostatic induction. The field near the surface of a conductor. Mechanism of creation of the field near the surface of a conductor. Dependence of the surface charge density on the curvature of the surface. Charge leakage from a tip. Electroscopes and electrometers. Metallic screen. Potential of a conductor. Capacitance of an isolated conductor. A system of conductors. Capacitors. A conducting sphere in a uniform field. The field of a dipole. Method of image charges.

Sec. 17. Electrostatic Field in the Presence of a Dielectric 129
Dipole moment of a continuous charge distribution. Polarization of dielectrics. Molecular pattern of polarization. Dependence of polarization on the electric field strength. The effect of polarization on electric field. Volume and surface density of bound charges. Electric displacement. Gauss' electrostatic theorem in the presence of dielectrics. Boundary conditions. Boundary conditions for the normal component of vector D. Boundary conditions for the tangential component of vector E. Refraction of field lines at the interface between dielectrics. Signs of bound charges at the interface between dielectrics. Method of images. Dielectric sphere in a uniform field.

Sec. 18. Energy of Electrostatic Field 148
Energy of interaction between discrete charges. Energy of interaction for a continuous distribution of charges. Self-energy. Energy density of a field. Energy of the field of surface charges. Energy of charged conductors. Energy of a dipole in an external field. Energy of a dielectric in an external field.

Sec. 19. Forces in an Electric Field 156
Nature of forces. Force acting on a point charge. Force acting on a continuously distributed charge. Force acting on a dipole. Moment of force acting on a dipole. Volume forces acting on a dielectric. Forces acting on a conductor. Surface forces acting on a dielectric. Volume forces acting on a compressible dielectric. Calculation of forces from the expression for energy. Problems.

www.EngineeringBooksPDF.com

Chapter 3. Dielectrics

Sec. 20. Local Field 172
The difference between a local field and an external field. Calculation of local field strength.

Sec. 21. Nonpolar Dielectrics 175
Molecular dielectric susceptibility. Rarefied gases. Dense gases.

Sec. 22. Polar dielectrics 177
Temperature dependence of polarization. Saturation field. Rarefied gases. Quantum interpretation of polarization of polar gaseous dielectrics. Dense gases. Polar liquids. Ionic crystals.

Sec. 23. Ferroelectrics 183
Definition. Hysteresis loop. Curie point. Molecular mechanism of spontaneous polarization. Dielectric domains. Antiferroelectrics.

Sec. 24. Piezoelectrics 187
Properties of piezoelectrics. Longitudinal and transverse piezoelectric effects. Mechanism of piezoelectric effect. Inverse piezoelectric effect. Pyroelectrics. Problems.

Chapter 4. Direct Current

Sec. 25. Electric Field in the Case of Direct Currents 191
The field in a conductor. The sources of a field. Field outside a conductor. Surface charges. Volume charges. Mechanism of generating direct currents. Change in potential along a current-carrying conductor.

Sec. 26. Extraneous Electromotive Forces 195
The origin of extraneous e.m.f. s. Mechanical extraneous e.m.f. Galvanic cells. Voltaic cell. Range of action of extraneous e.m.f. s. Law of conservation of energy. Polarization of a cell. Methods of depolarization. Accumulator.

Sec. 27. Differential Form of Joule's Law. Work Done During the Passage of Current and Power Developed 202
Work performed during passage of current. Power. Differential form of Joule's law. The source of energy for the work done by current. Derivation of Ohm's law from the electron pattern of electrical conductivity. Derivation of Joule's law from the electron theory of electrical conductivity. Drawbacks of the classical theory of electrical conductivity. Main features of quantum-mechanical interpretation of electrical conductivity.

Sec. 28. Linear Circuits. Kirchhoff's Laws 206
An isolated closed loop. Branched circuits. Kirchhoff's laws.

Sec. 29. Currents in a Continuous Medium 209
Formulation of the problem. Derivation of the formula. Conditions of applicability of Eq. (29.6). Coaxial electrodes. Nonhomogeneous medium.

Sec. 30. Earthing of Transmission Lines 213
Formulation of the problem. Calculation of resistance. Experimental verification. Step voltage. Problems.

Chapter 5. Electrical Conductivity

Sec. 31. Electrical Conductivity of Metals 218
The proof of the absence of mass transport by electric current in metals. The Tolman and Stewart experiments. On the band theory. Temperature dependence of resistance. Hall effect. Magnetoresistance. Mobility of electrons. Superconductivity. Critical temperature. Critical field. Meissner effect. Surface current. Soft and hard superconductors. The theory of superconductivity.

Sec. 32. Electrical Conductivity of Liquids 225
Dissociation. Calculation of electrical conductivity. Dependence of electrical conductivity on concentration. Temperature dependence of electrical conductivity. Electrolytes.

www.EngineeringBooksPDF.com

Sec. 33. Electrical Conductivity of Gases 228
Self-sustained and non-self-sustained currents. Non-self-sustained current. Saturation current density. The characteristic of current. Self-sustained current. The effect of volume charge. Mobility of charges. Comparison of results with experiment.

Sec. 34. Electric Current in Vacuum 232
Thermionic emission. The characteristics of an electron cloud. Saturation current density. Three-halves power law. Problems.

Chapter 6. Stationary Magnetic Field

Sec. 35. Ampère's Circuital Law 240
Formulation of the problem. The integral form of Ampère's circuital law. Differential form of Ampère's circuital law. Experimental verification of Ampère's circuital law. The derivation of the differential form of Ampère's law by direct differentiation of the Biot-Savart law.

Sec. 36. Maxwell's Equations for a Stationary Magnetic Field 245
Equation for div B. Maxwell's equations. The types of problems involved.

Sec. 37. Vector Potential 247
The possibility of introducing a vector potential. Ambiguity of vector potential. Potential gauging. Equation for vector potential. Biot-Savart law. The field of an elementary current.

Sec. 38. Magnetic Field in the Presence of Magnetics 254
Definition. Mechanisms of magnetization. Magnetization. Vector potential in the presence of magnetics. Volume density of molecular currents. Surface molecular currents. Uniformly magnetized cylinder. Magnetic field strength. Equation for the magnetic field strength. Relation between magnetization and magnetic field strength. Field in a magnetic. Permanent magnets. Boundary conditions for the field vectors. The boundary condition for the normal component of vector **B**. The boundary condition for the tangential component of vector **H**. Refraction of magnetic field lines. The measurement of magnetic induction. The fields of a very long solenoid and a uniformly magnetized very long cylinder. The measurement of permeability, magnetic induction and the field strength inside magnetics. A magnetic sphere in a uniform field. Magnetic shielding.

Sec. 39. Forces in a Magnetic Field 270
Forces acting on a current. Lorentz's force. The force and the torque acting on a magnetic dipole. Body forces acting on incompressible magnetics. Problems.

Chapter 7. Magnetics

Sec. 40. Diamagnetics 277
Larmor precession. Diamagnetism. Diamagnetic susceptibility. Temperature independence of diamagnetic susceptibility.

Sec. 41. Paramagnetics 282
Mechanism of magnetization. Temperature dependence of paramagnetic susceptibility. Magnetic moments of free atoms. Magnetic moments of molecules. Magnetism due to free electrons. Paramagnetic resonance.

Sec. 42. Ferromagnetics 287
Definition. Magnetization curve and hysteresis loop. Permeability curve. Classification of ferromagnetic materials. Interaction of electrons. Basic theory of ferromagnetism. Curie-Weiss law. Magnetization anisotropy. Domains. Domain boundaries. Magnetic reversal. Antiferromagnetism. Ferrimagnetism. Ferromagnetic resonance.

Sec. 43. Gyromagnetic Effects 295
Relation between angular momentum and magnetic moment. Einstein-de Haas experiment. Barnett effect. Problems.

Chapter 8. Electromagnetic Induction and Quasistationary Alternating Currents

Sec. 44. Currents Induced in Moving Conductors 300

www.EngineeringBooksPDF.com

Emergence of an e.m.f. in a moving conductor. Generalization to an arbitrary case. A.c. generators. The law of conservation of energy.

Sec. 45. Faraday's Law of Electromagnetic Induction 304
Definition. Physical essence of the phenomenon. A conductor moving in a varying magnetic field. Application of the law of electromagnetic induction to a.c. generators.

Sec. 46. Differential Form of the Law of Electromagnetic Induction 306
Differential form of Faraday's law. Nonpotential nature of induced electric field. Vector and scalar potentials in a varying electromagnetic field. Ambiguity of potentials and gauge transformation.

Sec. 47. Magnetic Field Energy 309
Magnetic field energy for an isolated current loop. Magnetic field energy for several current loops. Magnetic field energy in the presence of magnetics. Magnetic energy density. Inductance. The field of a solenoid. Energy of a magnetic in an external field. Calculation of forces from the expression for energy. Body forces acting on compressible magnetics. Energy of a magnetic dipole in an external magnetic field.

Sec. 48. Quasistationary A.C. Circuits 323
Definition. Self-inductance. Connection and disconnection of an *RL*-circuit containing a constant e.m.f. Generation of rectangular current pulses. *RC*-circuit. Connection and disconnection of an *RC*-circuit containing a constant e.m.f. *LCR*-circui containing a source of extraneous e.m.f.s. Alternating current. Vector diagrams. Kirchhoff's laws. Parallel and series connections of a.c. circuit elements. Mesh-current method.

Sec. 49. Work and Power of Alternating Current 334
Instantaneous power. Mean value of power. Effective (r.m.s.) values of current and voltage. Power factor. Electric motors. Synchronous motors. Asynchronous motors. Generation of rotating magnetic field. Load matching with a generator. Foucaul (eddy) currents.

Sec. 50. Resonances in A.C. Circuits 343
Voltage resonance. Current resonance. Oscillatory circuit.

Sec. 51. Mutual Inductance Circuit 347
Mutual Inductance. Equation for a system of conductors taking into account the self-inductance and mutual inductance. The case of two loops. Transformer. Vector diagram of a transformer at no-load. Vector diagram of a transformer under load. Autotransformer. Transformer as a circuit element. Real transformer.

Sec. 52. Three-Phase Current 353
Definition. Generation of three-phase current. Star connection of generator windings. Delta connection of generator windings. Load connection. Generation of a rotating magnetic field.

Sec. 53. Skin Effect 357
Essence of the phenomenon. Physical pattern of the emergence of skin effect. Basic theory. Skin depth. Frequency dependence of ohmic resistance of a conductor. Frequency dependence of inductance of a conductor.

Sec. 54. Four-Terminal Networks 361
Definition. Equations. Reciprocity theorem. Impedance of a four-terminal network. Simple four-terminal networks. Input and output impedances. Gain factor.

Sec. 55. Filters 366
Definition. Low-pass filter. High-pass filter. Iterative filter. Band filter.

Sec. 56. Betatron 369
Function. Operating principle. The betatron condition. Radial stability. Vertical stability. Betatron oscillations. Energy limit attainable in betatron. Problems.

Chapter 9. Electromagnetic Waves

Sec. 57. Displacement Current 376
The nature of displacement current. Why do we call the rate of variation of displacement the displacement current density? Maxwell's equations including displacement current. Relativistic nature of displacement current.

www.EngineeringBooksPDF.com

Sec. 58. Maxwell's Equations 381
Maxwell's equations. Physical meaning of Maxwell's equations. Conditions of applicability of Maxwell's equations. Completeness and compatibility of Maxwell's equations.

Sec. 59. The Law of Conservation of Electromagnetic Energy. Energy Flux 383
Formulation of the law of conservation of energy. Energy flux.

Sec. 60. Transmission of Electromagnetic Energy along Transmission Lines 385
Compensation of energy losses due to liberation of Joule's heat. Energy transmission along a cable. Transmission line for an alternating current. Equations for current and voltage. Characteristic impedance and propagation constant. Characteristic resistance. Velocity of propagation. Reflection.

Sec. 61. Electromagnetic Wave Radiation 392
Equation for vector potential. The choice of gauge function χ. Equation for scalar potential. Solution of the wave equation. Retarded and advanced potentials. Hertzian oscillator. The scalar potential of a dipole with a time-dependent moment. Vector potential. Electric and magnetic fields. The field of an oscillator in the wave zone. Power radiated by an oscillator. Radiation of current-carrying loop. Radiation of an electron moving with an acceleration. Decelerative force due to radiation.

Sec. 62. Propagation of Electromagnetic Waves in Dielectrics 405
Plane waves. Equations for the field vectors of a wave. Field vectors of a wave. Phase velocity. Wavelength. Properties of waves. Energy flux density.

Sec. 63. Propagation of Electromagnetic Waves in Conducting Media. 409
Complex permittivity. Penetration depth. Physical grounds for absorption. Interpretation of the skin effect. Phase velocity and the wavelength in a conducting medium. Relation between the phases of field vectors. Relationship between the amplitudes of field vectors.

Sec. 64. Invariance of a Plane Wave 413
Field transformation. Invariants of electromagnetic field transformations. Analysis of field invariants.

Sec. 65. Pressure of Electromagnetic Waves. Photon Momentum 415
Emergence of pressure. Pressure. The momentum of an electromagnetic wave train. Electromagnetic wave momentum per unit volume. Momentum of a photon.

Sec. 66. Waveguides and Resonators 418
Subcircuit. A section of a conductor. Induction coil. Capacitor. Radiation. Waveguides. Rectangular waveguide. Cut-off frequency. Phase velocity. Wavelength in a waveguide. Application of the method of images to the analysis of waveguides. Discrete nature of directions of propagation of plane waves from system oscillators. Boundary wavelength. Wavelength and phase velocity in a waveguide. Group velocity. Relation between the group velocity and phase velocity. Magnetic field. Classification of waves in waveguides. Resonators. Problems.

Chapter 10. Fluctuations and Noises

Sec. 67. Fluctuations in a Current-Carrying Loop. Resistance Noise 429
Theorem on equipartition of energy. Application of the equipartition theorem to a freely suspended mirror. Fluctuations in an oscillatory circuit. Frequency distribution of fluctuations. Resistance noise. Equivalent noise generator. Power of noise generator. Maximum sensitivity. Equivalent noise temperature of a receiver. Noise factor of a receiver. Signal-to-noise ratio.

Sec. 68. Schottky Noise and Current Noise 436
Source of Schottky noise. Frequency distribution of noise. Current noise. Methods of reducing noise.

Appendices 440
Subject Index 440

www.EngineeringBooksPDF.com

Introduction

At present, four types of interactions between material bodies are known to exist, viz. gravitational, strong, weak, and electromagnetic interactions. They are manifested on different three-dimensional scales and are characterized by different intensities.

Gravitational interaction is noticeable only for bodies on astronomical scale. Strong interactions can be observed only between certain particles when they approach each other to quite small distances (10^{-15} m). Weak interactions are exhibited during mutual conversion of certain kinds of particles and become insignificant as the particles are separated by large distances. Only electromagnetic interactions are manifested in our everyday life. Practically, all "forces" which are involved in physical phenomena around us, except for gravitational forces, are ultimately electromagnetic forces. Naturally, all diverse relations and phenomena due to electromagnetic interactions cannot be described by the laws of electrodynamics since on each level of a phenomenon there exist specific features and regularities that cannot be reduced to regularities on another level. However, electromagnetic interactions on all levels are to a certain extent an elementary link with the help of which the entire chain of relations is formed. This makes electromagnetic phenomena important from a practical point of view.

The theory of electromagnetic phenomena plays an extremely important role. This theory is the first relativistically invariant theory, which played a decisive role in the creation and substantiation of the theory of relativity and served as the "training ground" on which many new ideas have been verified. Quantum electrodynamics is the most elaborate branch of quantum theory, whose predictions are in astonishingly good agreement with experiment, although at present it is still not complete and free of internal contradictions. The philosophical aspect of electromagnetism is also very important. For example, specific features of the field form of existence of matter are clearly manifested within the framework of electromagnetic phenomena. The mutual conversion of different forms of matter and energy is also clearly reflected in these phenomena.

The substantiation of the theory is presented in the book in two ways. When the theory is substantiated without taking into account relativistic effects, the experimental basis of the theory of electricity and magnetism is formed by the invariance of an elementary charge, Coulomb's law, the principle of superposi-

www.EngineeringBooksPDF.com

tion for electric fields, the Biot-Savart law, the superposition principle for magnetic fields, Lorentz force, Faraday's law of electromagnetic induction, Maxwell's displacement currents, and the laws of conservation of charge and energy. When relativistic effects are taken into account for substantiating the theory, the Biot-Savart law, the principle of superposition for magnetic fields, and the Lorentz force no longer serve as independent experimental facts in the formulation of the theory. The second way of substantiating the theory of electricity and magnetism is presented not as the main line but as a side track chosen so as to simplify the mathematical aspect of the problem. It includes the following stages.

The relativistic nature of the magnetic field is demonstrated in Sec. 8, where the formula for interaction of currents flowing in infinitely long parallel conductors is derived and Lorentz force is obtained from electric interaction of charges. The field interpretation of these results allows us to find the magnetic induction of current passing through an infinitely long conductor. The principle of superposition for a magnetic field now becomes a corollary of the principle of superposition for an electric field. The transition to magnetic induction for arbitrary currents and the derivation of the corresponding equations are given in Sec. 35, where the independence of local relations from the values of physical quantities at other points is effectively used. After this, the Biot-Savart law is theoretically derived in Sec. 37, thus concluding the analysis of the connection existing in the relativistic concept of space and time between the invariance of an elementary electric charge, Coulomb's law, the principle of superposition for an electric field and the Biot-Savart law, as well as between the Lorentz force and the principle of superposition for a magnetic field.

www.EngineeringBooksPDF.com

CHAPTER 1

Charge. Field. Force

Charge is the source and the object of action of an electromagnetic field.
Field is the material carrier of electromagnetic interactions between charges, and is a form of the existence of matter.
Force is a quantitative measure of the intensity of interaction between charges.
Charges, fields, and forces are inseparably linked with space, time, and motion of matter.
Their interrelation cannot be understood without taking into account the connection with space, time and motion.

Sec. 1. Microscopic Charge Carriers

The properties of basic microscopic charge carriers are described. The distribution of electric charge in a proton and a neutron is discussed, and the physical meaning of electric charge is analyzed.

Classification. By microscopic charge carriers we mean charged particles and ions which can carry both positive and negative charge. The numerical value of a charge can only be an integral multiple of the elementary charge

$$| e | = 1.6021892(46) \cdot 10^{-19} \text{ C}. \tag{1.1}$$

In spite of persistent experimental attempts, it has not been possible so far to detect microscopic carriers with a fractional charge (see Sec. 3).

About 200 particles and an enormous number of ions, atoms, and molecules are known at present. A large number of particles exist only for a short time after their creation and then disintegrate into other particles. In other words, *particles have a finite lifetime.* In most cases, this lifetime is extremely small and is of the order of a very small fraction of a second. *Only a small number of charged particles have an infinite lifetime. These are the electron, the proton, and their antiparticles.* Atomic nuclei contain protons, while the electron shells of atoms contain electrons. It is these particles that are responsible for almost all phenomena analyzed in a course on electricity and magnetism. In addition to protons, nuclei also contain neutrons. These are electrically neutral and have an infinite lifetime in nuclei. However, their average lifetime outside nuclei is

www.EngineeringBooksPDF.com

about 17 min, after which they disintegrate into protons, electrons, and antineutrinos.

The charge of ions is due to the fact that the electron shell of the corresponding atom or molecule lacks one or several electrons (positive ions) or, on the contrary, has extra electrons (negative ions). Consequently, the treatment of ions as microscopic charge carriers boils down to an investigation of electron and proton charges.

Electron. An electron is the material carrier of an elementary negative charge. *It is usually assumed that an electron is a structureless point particle*, i.e. the entire charge of an electron is concentrated at a point. *Such a representation is intrinsically contradictory* since the energy of the electric field created by a point charge is infinite, and hence the inertial mass of the point charge must also be infinite. This is in contradiction with the experiment since the electron rest mass is $m_e = 9.1 \times 10^{-31}$ kg. *However, we must reconcile ourselves with this contradiction in the absence of a more satisfactory and less contradictory view on the structure (or absence of a structure) of electron.* The difficulties associated with an infinite rest mass can be successfully overcome in calculation of various effects with the help of **mass renormalization** which essentially consists in the following. Suppose that it is required to calculate a certain effect and an infinite rest mass appears in the calculations. The quantity obtained as a result of calculations is infinite and is consequently devoid of any physical meaning. In order to obtain a physically reasonable result, another calculation is carried out, in which all factors, except those associated with the phenomenon under consideration, are present. This calculation also includes an infinite rest mass and leads to an infinite result. Subtraction of the second infinite result from the first leads to the cancellation of infinite quantities associated with the rest mass. The remaining quantity is finite and characterizes the phenomenon being considered. Thus, we can get rid of the infinite rest mass and obtain physically reasonable results which are confirmed by experiment. Such a method is used, for example, to calculate the energy of an electric field (see Sec. 18).

Proton. A proton is the carrier of a positive elementary charge. Unlike an electron, a proton is not considered as a point particle. The distribution of the electric charge in a proton has been thoroughly investigated in experiments. The method of investigation is similar to that used at the beginning of this century by Rutherford in investigations of the atomic structure, which led to the discovery of the nucleus. The collisions between electrons and protons are analyzed. If we assume the proton to be a *spherically symmetric* distribution of charge in a finite volume, *the electron trajectory which does not pass through this volume is independent of the law of charge distribution*, and is the same as if the entire charge of the proton were concentrated at its centre. *The trajectories of electrons passing through the volume of the proton depend on the specific form of charge distribution in it.* These trajectories can be calculated. Hence, by carrying out quite a large number of observations of the results of collisions between electrons and protons, we can draw conclusion about the charge distribution inside the proton. Since very small volumes in space are involved, electrons with very high energies are required for experiments. This necessity is

www.EngineeringBooksPDF.com

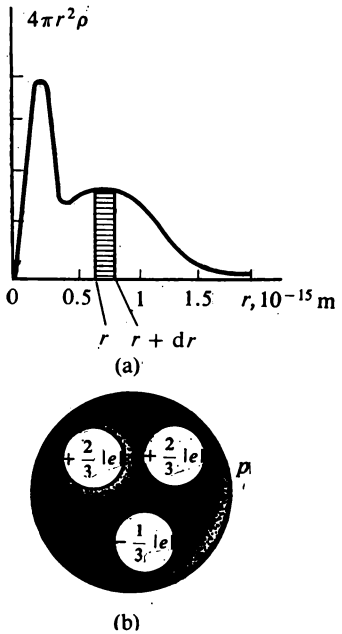


Fig. 1. Electromagnetic structure of a proton. Almost the entire charge is concentrated in a sphere of radius r_0.

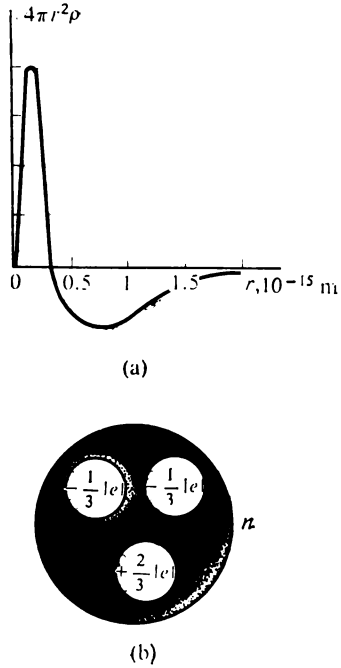


Fig. 2. Electromagnetic structure of a neutron. The positive charge is located near the centre, while the negative charge is at the periphery. The positive and negative charges compensate for each other, and hence the neutron is electrically neutral as a whole.

dictated by the quantum theory. According to the de Broglie relations material particles have wave properties, and the wavelength of a particle is inversely proportional to its momentum. In order to perceive a certain part in space, it is obviously necessary to use particles whose wavelengths are less than the corresponding spatial dimensions of this part. This involves quite high momenta of particles. Therefore the investigation of the electromagnetic structure of a proton became possible only after the creation of electron accelerators with an energy of several billion electron-volts. The result of these experiments is shown in Fig. 1*a*. Here, the ordinate represents not the charge density ρ at a distance r from the centre of the proton but the quantity $4\pi r^2\rho$ which is the density of the overall charge in all directions at a distance r from the centre. This is so because $4\pi r^2\rho\,(r)\,dr$ is the total charge in a spherical layer of thickness dr. It can be seen from the figure that *the entire proton charge is practically concentrated in a sphere of radius of about* 10^{-15} m. *After the first maximum,* $4\pi r^2\rho\,(r)$ *does not decrease monotonically, but another maximum exists.*

www.EngineeringBooksPDF.com

Neutron. Similar experiments, carried out on scattering of electrons by neutrons, showed that the neutron has an electromagnetic structure and is not an electrically neutral point particle. The distribution of electric charge in a neutron is shown in Fig. 2*a*.

Obviously, *a positive charge is located near the centre of the neutron, while a negative charge exists at its periphery.* The areas bounded by the curves and the abscissa axis are equal. Consequently, the positive charge is equal to the negative charge, and the neutron is electrically neutral as a whole. *The sizes of the regions in which electric charges are concentrated are approximately the same in a proton and a neutron.*

What does the continuous distribution of an elementary electric charge mean? The area bounded by the curve and the abscissa axis (see Fig. 1*a*) is numerically equal to the proton charge, while the shaded area is equal to the charge inside a proton in a spherical layer of thickness dr at a distance r from the centre of the proton. Clearly, this charge is just a small part of the total proton charge, i.e. a small part of an elementary charge. However, *it has not been possible to discover in nature physical objects whose charge is a fractional part of the elementary charge.* What, then, is the meaning of the statement that a small part of an elementary charge is located in the volume $4\pi r^2 dr$?

At present, it is assumed that a proton consists of two point quarks with a charge $+2|e|/3$ and one point quark with a charge $-|e|/3$ (see Fig. 1*b*). The quarks move inside a proton. Their relative duration of stay at different distances from the proton centre can be effectively represented as a spreading of the charge over the proton volume, as shown in Fig. 1*a*. A neutron consists of two quarks with a charge $-|e|/3$ and one quark with a charge $+2|e|/3$ (Fig. 2*b*). The charge distribution in a proton can be explained similarly.

Quarks have not been observed in free state in spite of considerable experimental efforts. At present it is assumed that it is practically impossible to detect quarks in free state since it requires an infinite energy. They do, however, exist inside a proton. Such an assumption provides an explanation for a large number of phenomena and is therefore accepted by physicists as a possible hypothesis.

There is no direct evidence of the presence of quarks inside a proton.

Spin and magnetic moment. In addition to charge, particles may possess angular momentum, or spin. *Spin is not due to the rotation of a particle* since for such an explanation under reasonable assumptions concerning the particle size, linear velocities exceeding the velocity of light would have to be admitted. This, however, is impossible. Consequently, *spin is considered as an intrinsic property of a particle.*

Spin is associated with the magnetic moment of a charged particle. This also cannot be explained by the motion of the charge and is considered as a fundamental property.

In classical electrodynamics, magnetic moment appears only due to motion of charges along closed trajectories. Consequently, the spin magnetic moment of particles cannot be described in the classical theory of electricity and magnetism. However, the magnetic field created by the spin magnetic moment can be described phenomenologically if necessary. As a rule, the strength of this field is very

www.EngineeringBooksPDF.com

small and attains large values only in the case of permanent magnets. The classical theory is unable to explain the mechanism of creation of this field, although the field itself outside permanent magnets can be completely described by the classical theory (see Sec. 38).

Electron is considered as a point particle, although it leads to difficulties. It has not been possible to experimentally determine the internal electromagnetic structure of an electron. Continuous distribution of an elementary electric charge is not connected with its division into parts and only means that the law of motion of this charge in space is taken into account.

There is no charge smaller than an elementary charge. What, then, is the idea behind the distribution of charge in a proton if the total charge in it is equal to an elementary charge? What is the main difficulty associated with the representation of an electron as a point particle? What artificial method is used for overcoming this difficulty?

Sec. 2. Charged Bodies. Electrostatic Charging

The physical nature of processes resulting in electrostatic charging of bodies in contact is elucidated. Some information on the energy spectrum of electrons in solids is given.

Thermionic work function. The forces that keep neutral atoms in a molecule and neutral molecules in a solid are considered in the course on molecular physics. The very *fact of the existence of solids indicates that there are forces confining electrons inside a solid.* In order to extract an electron from the solid, a certain work against the forces retaining the electrons inside the solid must be performed. Suppose that a solid body together with the surrounding medium is enclosed in an adiabatic shell and is kept at a constant temperature T. *Owing to thermal motion and the velocity distribution of electrons inside the body, there will be electrons whose kinetic energy is sufficiently high to allow them to overcome the forces keeping them inside the body and thus leave the body.* As a result, an electron "gas" is formed near the surface of the body. In the course of their motion, the electrons of this gas approach the surface of the body and are captured by it. *If the number of electrons leaving the volume of the solid is, on the average, equal to the number of electrons entering the volume of the solid from the layer of the electron "gas" adjoining its surface, thermodynamic equilibrium is attained.* In this case, the electron concentration near the surface of the solid has a definite value n_0. This electron gas is non-degenerate, and its density can be represented in the form of the Boltzmann distribution:

$$n_0 = A \exp [-\Phi/(kT)], \tag{2.1}$$

www.EngineeringBooksPDF.com

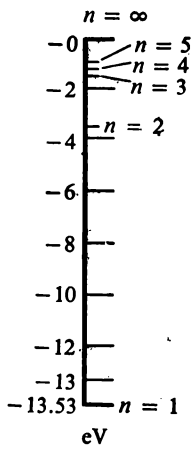


Fig. 3. Energy spectrum of a hydrogen atom.

where A depends only on the temperature T, and Φ is the thermionic work function.

According to the content of the Boltzmann distribution, the work function is the difference in energies of an electron outside a solid and inside it. However, electrons in a solid have different energies, and only an analysis of their energy spectrum clarifies what energy is meant while determining Φ.

Energy spectrum of electrons. The laws of motion of microparticles are given in quantum mechanics, which allows us to calculate the energy spectrum of electrons if we know the law of variation of their potential energy. These calculations are complicated by the fact that we must also take into account the mutual interaction of electrons. The exact solution of such problems cannot be obtained even with the help of modern computers and will hardly be possible in future. But there is no need for them, since it is possible to work out approximate methods which meet practical requirements sufficiently well. It is important to establish that this spectrum exists and is discrete for electrons contained in a finite region of space. It determines various properties of a body. Experimental investigation of these properties allows us to reveal the peculiarities of the energy spectrum. Consequently, the energy spectrum can be studied both theoretically and experimentally.

The energy spectrum of electrons in solids is investigated in detail. Its basic features consist in the following. Energy levels in an isolated atom form a discrete set of energies.

Figure 3 represents idealized energy levels of a hydrogen-like atom. In analytical form, the electron energy on the nth level is given by

$$W_n = -A/n^2,$$

where A is a positive quantity expressed in terms of an elementary charge, mass of the nucleus and electron, and Planck's constant. Electrons on the level $n = 1$ have the lowest energy. The separation between the levels amounts to several electron-volts, these distances decreasing with increasing n.

Since electrons obey the Fermi-Dirac statistics, only one electron can exist in each quantum-mechanical state. The quantum state is characterized not only by energy. In a hydrogen-like atom, it is also characterized by the angular momentum of the electron in its orbital motion in an atom, its orientation in space, as well as by the orientation of the electron spin. These two latter characteristics are also quantized, i.e. have a discrete set of numerical values. As a result, it turns out that *each energy level contains not one but several electrons*. Calculations show that the level $n = 1$ may contain two electrons which differ in the spin orientation (only two spin orientations are possible). The angular momentum on this

www.EngineeringBooksPDF.com

level may only be equal to zero. On the next level $n = 2$, the angular momentum of the electron may have, in addition to the zero value, a value differing from zero. For the zero value of the angular momentum, there is no sense in determining its spatial orientation, in contrast to the case when the angular momentum has a nonzero value. For $n = 2$, the angular momentum has three possible orientations. Consequently, there are four quantum-mechanical states on the level $n = 2$ corresponding to different magnitudes and spatial orientations of the angular momentum. In each of these states, the electron spin may have two orientations, and hence in total there are eight quantum-mechanical states on the energy level $n = 2$. This means that this level may contain eight electrons altogether. It turns out that the next levels may contain 18, 32, 50, etc. electrons. Since the stable state of an atom (ground state) corresponds to the state with the lowest energy, the energy levels must be filled starting from the level $n = 1$, and the filling of the next level starts only when the previous level is completely filled by the electrons. A complex of electrons with a certain value of n is called an atom shell. Atom shells are usually denoted by the letters K, L, M, N, etc. according to the following array:

n	1	2	3	4	5
Electron shell	K	L	M	N	O

For example, instead of saying "an electron on the level $n = 2$" we say "an electron of the L-shell".

The situation changes for the atoms constituting the crystal lattice of a solid. The very existence of crystal lattices indicates that there is an interaction between atoms, which substantiates the lattice formation. Consequently, the atoms in a lattice cannot be considered isolated. We must, therefore, consider *the entire crystal lattice as a single system and speak about the energy levels of this system.* It turns out that the energy spectrum of the crystal lattice is connected with the energy spectrum of isolated atoms through a simple relation, viz. *as a result of interaction between the atoms, each of the energy levels* $n = 1, 2, \ldots$ *splits into a large number of closely spaced sublevels on which all the electrons, which were initially on the corresponding level of isolated atoms, can be arranged.* For example, the K-shell of an isolated atom is occupied by two electrons. If atoms constitute a crystal lattice consisting of N_0 atoms, the level $n = 1$ splits into N_0 sublevels each of which contains two electrons with different spin orientations, i.e. $2N_0$ different quantum-mechanical states occupied by $2N_0$ electrons which formerly belonged to the K-shells are formed in total in the crystal lattice.

A set of closely lying energy levels formed as a result of splitting of a certain energy level of an isolated atom is called the energy band, or simply the band. We can speak of the K-,L-, etc. bands corresponding to the K-, L-, . . . shells of isolated atoms. The schematic diagram of band formation is shown in Fig. 4. As was mentioned above, the separation between different levels inside the bands is extremely small. On the other band, the distance between different bands remains considerable and is equal, in order of magnitude, to the distance between the energy levels of isolated atoms. The spacings between the energy bands occupied by electrons

www.EngineeringBooksPDF.com

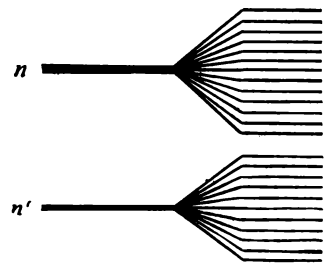


Fig. 4. Formation of energy bands.

are also called bands. These bands are termed *forbidden bands* since electrons cannot exist in these bands.

Thus, the energy spectrum of electrons in a solid consists of allowed and forbidden bands. *The distance between the energy levels within each allowed band is extremely small in comparison with the forbidden bandwidth.* The energy level diagram considered above for an isolated atom is idealized. If we take into account the interaction between the electrons in greater detail, it turns out that the energy of electrons in a shell is not the same but depends, for example, on the angular momentum. The energy of the electron with a higher value of n may be not higher but lower than the energy of electrons on the preceding level. As a result, the sequence of filling shells with electrons may change. The structure of the energy bands in a crystal and their filling with electrons will change accordingly. However, the general nature of the spectrum of a solid remains unchanged.

Fermi energy. The ground state of a solid is the state with the lowest possible energy. Consequently, *at the temperature* 0 K *all quantum states of electrons must be filled successively, without gaps, starting from the level with the lowest energy.* Since the number of electrons is finite, there is a finite filled level corresponding to the highest energy, while the upper-lying levels are vacant. Thus, *at* 0 K *there exists a distinct boundary between the filled and unoccupied levels.*

At a temperature other than 0 K, *this boundary is blurred*, since as a result of thermal motion, the energy of some electrons turns out to be higher than the boundary energy corresponding to $T = 0$ K, while the energy of some other electrons is lower than the boundary energy. Thus, some energy levels which were free at 0 K will be occupied, while some of the previously occupied levels will become empty. *The width of the transition region from almost completely filled levels to almost completely unoccupied energy levels is of the order of* kT. In this case, the energy distribution of electrons is described by the Fermi-Dirac function

$$f(E, T) = \{1 + \exp[(E - \mu)/(kT)]\}^{-1}, \qquad (2.2)$$

where E is the electron energy and μ the Fermi energy which depends on temperature. The Fermi energy is defined as the energy for which the Fermi-Dirac function is equal to 1/2.

The concept of the Fermi energy for metals is quite obvious. *In this case, the Fermi energy is the energy of electrons on the level which is filled at* $T = 0$ *and above which the levels are empty.* This definition is exact for $T = 0$ K and fairly accurate for all temperatures at which "blurring" of the Fermi distribution is slight (for most metals, this statement is valid up to the melting point and even higher).

The Fermi energy for dielectrics corresponds to the middle of the forbidden band (for $T = 0$ K) lying above the uppermost completely filled band. Since no elec-

www.EngineeringBooksPDF.com

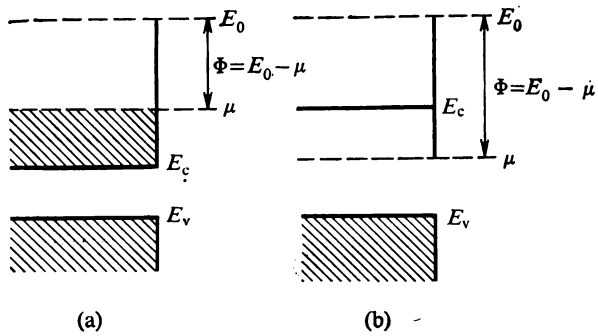


Fig. 5. Potential well for an electron in a metal (a) and in a dielectric (b). The thermionic work function Φ is the difference between the energy E_0 of an electron at rest in vacuum and the energy μ of the Fermi level.

tron can occupy this level, *the Fermi energy does not correspond to the energy of any real electron in a dielectric.* This, of course, does not make it less important for the description of statistical properties of electrons in dielectrics by using formula (2.2).

The theory shows that the **thermionic work function Φ appearing in formula** (2.1) is connected with the energy μ of the Fermi level through the relation

$$\Phi = E_0 - \mu, \tag{2.3}$$

where E_0 is the energy of an electron at rest outside the conductor in a vacuum. Thus Φ *is the work performed in shifting an electron from the Fermi level to beyond the limits of the solid.* For metals, this statement has a literal meaning, while for dielectrics it is conditional to a certain extent since there are no real electrons on the Fermi level. *In both cases, however, it is the work done to extract an electron from a solid against the forces confining electrons in it.* The existence of work function is manifested, for instance, in photoelectric effect, when the energy of a photon absorbed by a metal is completely transferred to an electron. The work function can be directly determined from the photoelectric threshold. Hence we can say that *electrons inside a solid are in a potential well of depth* Φ. The form of potential wells for metals (a) and for dielectrics (b) is shown in Fig. 5 (the energy levels occupied by electrons are hatched). The gap between the levels E_c and E_v is the forbidden band. It should be noted that the *work function for dielectrics strongly depends on the degree of purity of the material.* Even slight impurities can considerably change the work function. Besides, the *work function depends on even very low contamination of the surface.* The work function for pure metals is of the order of several electron-volts. For example, it is equal to 4.53 eV for tungsten, 4.43 for molybdenum, 4.39 for copper, etc.

Contact potential difference. The forces confining electrons to a solid are electric in origin. They are due to the potential difference between the points outside a body and its inner points. In other words, the *electron gas near the surface is*

www.EngineeringBooksPDF.com

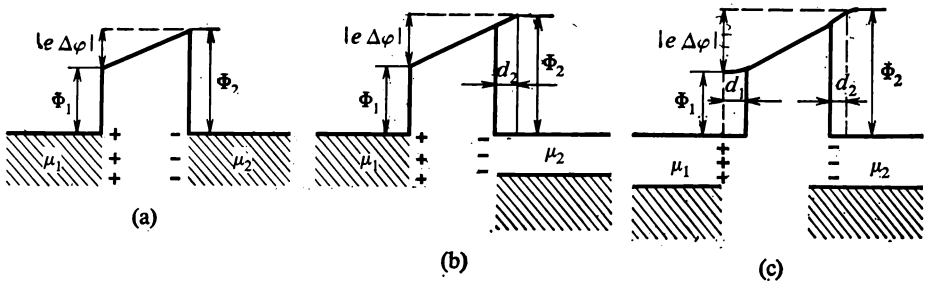


Fig. 6. Formation of contact potential difference between metal-metal (a) metal-dielectric (b), and dielectric-dielectric (c) surfaces.

subjected to the action of electric forces that tend to pull electrons into the body. These forces are the stronger, the larger the work function Φ. They act in a very thin layer of molecular dimensions ($d \simeq 10^{-10}$ m). Consequently, the effective intensity of the electric field due to which these forces appear is very high:

$$E_{\text{eff}} \sim \Phi/(|e|d) \sim 10^{10} \text{ V/m}, \tag{2.4}$$

where we took into account the fact that in order of magnitude, the work function is equal to several electron-volts.

Let us bring together the surfaces of two solids so that the layers of the electron gas near these surfaces overlap in the gap between them. Consequently, the bodies will exchange electrons. *Since the forces pulling electrons into a body are stronger in the body with a larger work function, the electrons of the body with a smaller work function will go over to the body with a larger work function when two bodies approach each other. As a result, the former body will acquire a positive charge, and the latter a negative charge.* Hence an electric field arises between the surfaces and prevents the motion of electrons which generated this field. The intensity of this field attains a certain value, after which a further transition of electrons from one body to the other is terminated, and equilibrium state sets in. The surfaces turn out to have charges opposite in sign but equal in magnitude. *A potential difference, called the contact potential difference, appears between these surfaces, as between the plates of a capacitor.*

The contact potential difference can be found from the following considerations. Since electron equilibrium sets in between the bodies, the Fermi energies of these bodies must be equal, and hence the upper points of the potential wells are displaced relative to one another. Consequently, a potential difference and electric field intensity appear between the surfaces of the bodies.

Figure 6 illustrates the contact potential difference across a gap between two metals (Fig. 6*a*), between a metal and a dielectric (Fig. 6*b*) and between two dielectrics (Fig. 6*c*). The difference in the formation of contact potential difference between a metal and a dielectric consists in that the electric field does not penetrate the metal but penetrates the dielectric to a small depth (in Fig. 6*b*, *c*, the penetration depths are denoted d_1 and d_2). Consequently, *a potential drop in dielectrics occurs not only between the surfaces but partially in a thin layer near*

www.EngineeringBooksPDF.com

its surface as well. The thickness of this layer, however, is usually small in comparison with the distance between the surfaces, so that this circumstance can be ignored to a high degree of accuracy.

Figure 6 shows that the difference in energies corresponding to upper points of the potential wells is equal to $\Phi_2 - \Phi_1$. Hence, the contact potential difference between the surfaces of bodies in electron equilibrium is given by

$$| \Delta\varphi | = | \Phi_2 - \Phi_1 | / | e |. \tag{2.5}$$

It should be noted that the potential decreases in the direction from positively charged bodies to negatively charged ones. Therefore, the change in potential is opposite to the change in the potential energy of an electron, i.e. the potential decreases in the direction from the first body to the second.

Electrostatic charging. If flat surfaces of two bodies between which a contact potential difference exists are moved apart and kept strictly parallel, the charges on the surfaces will remain, and the bodies will carry unlike charges. However, it is practically impossible to move the surfaces apart in such a way that they remain strictly parallel, since different regions move apart with different velocities. *The results of moving conductors and dielectrics apart are different in principle.*

When flat surfaces of conductors are drawn apart, the charges on them can move over the surface. If some regions of the surfaces are drawn apart before the others, the charge density on them will decrease at the same potential difference as in the case of a capacitor. As a result, the bodies will exchange charges in order to restore the electron equilibrium. This exchange occurs through the electron cloud at a given region of the surface and as a result of motion of charges over the surface in other regions. The regions of the conductor surface, which were drawn apart to a sufficiently large distance and thus lost contact through the electron cloud near the surface, turn out to be practically uncharged. The charge is retained only on those regions of the surface which are still in electron contact. Finally, a moment comes when the electron contact is observed only on an infinitely small surface area containing very small charge. For this reason, *no electric charge remains on the conductors when they are drawn apart completely.*

The situation is different when dielectrics are drawn apart. The charges on the dielectrics cannot move along the surfaces, and the potential itself may be different in different regions of the surface. When these regions of the surface are drawn apart, the potential difference then increases in the same way as the potential difference between the capacitor plates when the charge on the plate is constant and the distance between the plates increases. The charge density on the surfaces does not change significantly. After the electron contact through the electron cloud near the surface has disappeared, electric charges remain on the regions of the surface. *As a result of complete separation of the dielectric surfaces, they become carriers of equal but opposite charges.* This process is called **electrostatic charging**.

In order to ensure a closer approach of dielectric surfaces and the formation of a contact potential difference, the bodies are usually rubbed against each other. This process is called triboelectrification. However, friction in this case

www.EngineeringBooksPDF.com

has nothing in common with electrostatic charging. It would be more correct to call this effect charging by contact. The terminology was established before the physical nature of the phenomenon was clarified.

The work function of dielectrics depends on the purity of composition and the state of the surface.
When two dielectrics are brought in contact, electrons are transferred from the body with a smaller work function to the body with a greater work function.

The separation between energy levels inside each allowed energy band is extremely small in comparison with the width of the forbidden band. The Fermi energy in a dielectric does not correspond to the energy of any real electron in the dielectric.
The thermionic work function is the work required to transfer an electron from the Fermi level outside the solid.

What is the relation between the energy levels of an isolated atom and the energy bands of a solid? Which factors are responsible for the formation of energy bands?
How can the Fermi energy in metals be visually interpreted?
Why is this interpretation inapplicable to dielectrics?
How can the signs of charges of bodies in contact be determined? Why cannot metals be electrically charged by contact?

Sec. 3. Elementary Charge and Its Invariance

The experiments proving the existence of an electric charge and the absence of charges that are fractions of an elementary charge are described. The experimental evidence on the identity of positive and negative charges as well as the invariance of these charges are discussed.

Millikan oil-drop experiment. Although the idea of discrete nature of electric charge was put forth in a clear form by Franklin as early as in 1752, it was rather speculative. The discreteness of electric charge was established as a fundamental experimental result following the discovery of the laws of electrolysis by Faraday (1791-1867) in 1834. However, such a conclusion was drawn only in 1881 by Helmholtz (1821-1894) and Stoney (1826-1911). Soon afterwards, Lorentz (1853-1928) developed in 1895 the theory of electromagnetism which is based on the existence of elementary charges (electrons). The numerical value of an elementary charge was theoretically calculated from the laws of electrolysis, since the value of the Avogadro constant was known. A direct experimental measurement of the elementary charge was made in 1909 by Millikan (1868-1953).

The experimental set-up used by Millikan is shown in Fig. 7. Minute spherical particles move in a viscous liquid in a uniform electric field E. The particle

www.EngineeringBooksPDF.com

is subjected to a lifting force acting against the force of gravity (the density of the particle is higher than that of the liquid), and to the force f_{fr} of viscous friction acting against the velocity.

According to Stokes' formula, the force of viscous friction is proportional to velocity. At a constant particle velocity, the sum of forces acting on it is equal to zero.

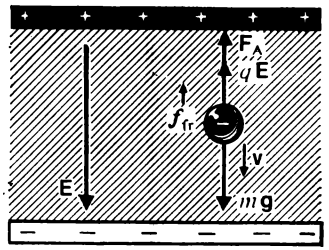


Fig. 7. Schematic diagram of Millikan's experiment.

All forces, with the exception of the force acting on the particle due to the electric field, can be measured experimentally as the particle moves in a medium in the absence of an electric field. Having studied the motion of the particle in the electric field we calculate the force $q\mathbf{E}$. This allows us to calculate the particle charge q, since the field strength E is known.

We can also change the field strength and ensure that the particle is in a state of rest. In this case, there is no friction and the other forces are known. Consequently, we can calculate the value of q if we know the value of E.

The charge of a particle changes with time. This is reflected in the motion of the particle. Having determined the charges q_1 and q_2 of the particle at different instants of time, we can determine the variation of charge

$$\Delta q = q_2 - q_1. \tag{3.1}$$

After carrying out a large number of measurements of charges, Millikan found that Δq is always an integral multiple of the same quantity $|e|$:

$$\Delta q = n\,|e|,\ n = \pm 1,\ \pm 2,\ \ldots, \tag{3.2}$$

$$|e| = 1.6 \cdot 10^{-19}\ \text{C}. \tag{3.2a}$$

Resonance method for measurement of charge. The methods used for directly measuring an electric charge were later perfected. *At present, measurements can be made with such precision that it is possible to detect decimal fractions of an elementary charge.* The most effective method used for such measurements is the resonance method shown in Fig. 8. A sphere of a very small mass m is fixed to a very thin elastic rod. Under the action of elastic forces resulting from bending of the rod, the sphere oscillates about its equilibrium position with a natural frequency ω_0 which can be measured experimentally. If the sphere carries a certain charge q, it undergoes forced vibrations under the action of an alternating electric field. The amplitude of these vibrations depends on the ratio of the frequencies ω and ω_0. The maximum amplitude of vibrations is attained at resonance ($\omega \simeq \omega_0$). In the resonance state, the amplitude of vibrations of the sphere is equal to

$$A_{res} = qE_0Q/(m\omega_0^2), \tag{3.3}$$

where Q is the quality of the system and E_0 is the amplitude of the electric field. Let us estimate the potentialities of this method. Suppose that $m = 1$ mg =

www.EngineeringBooksPDF.com

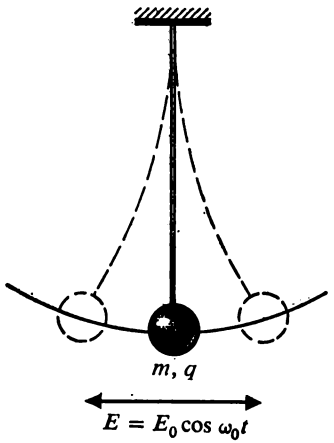


Fig. 8. Resonance method for measuring electric charge.

10^{-6} kg; $E_0 \simeq 10^5$ V/m; $q = 1.6 \times 10^{-19}$ C; $\omega_0 = 10^{-1}$ s^{-1}; and $Q \simeq 100$. In this case,

$$A_{res} \approx \frac{1.6 \cdot 10^{-19} \cdot 10^5 \cdot 10^2}{10^{-6} \cdot 10^{-2}} \text{ m} \simeq 1.6 \cdot 10^{-4} \text{ m} = 160 \text{ μm}. \quad (3.4)$$

The value of A_{res} (160 μm) is quite large and we can easily measure a small part of this quantity. Consequently, this method can be used for measuring electric charges that are much smaller than 1.6×10^{-19} C. This method has attained such a degree of perfection that it can be used for measuring, in principle, a fraction of an elementary charge, if only it existed.

As the charge on the sphere changes by Δq, the amplitude of resonance vibrations changes abruptly:

$$\Delta A_{res} = \Delta q E_0 Q/(m\omega_0^2). \quad (3.5)$$

As a result of measurements, it was established with a very high degree of precision that the charge on the sphere always changes by an integral multiple of an elementary charge, and that there are no charges that are fractions of an elementary charge.

Nonexistence of fractional charges. Following the predictions concerning the existence of quarks, many serious attempts were made to detect fractional charges It is assumed that quarks are particles which constitute most of the heavy elementary particles (protons, etc.). It was predicted that quarks must carry an electric charge of 1/3 and 2/3 of an elementary charge (with appropriate sign). Various methods, including the resonance method, were employed by many scientists to detect quarks. All these attempts proved unsuccessful. Thus, at present it *has been experimentally established to a sufficiently high degree of accuracy that fractional charges do not exist in free state.*

The words "in free state" are quite significant, since the experiments were directed at quests for free quarks. *This, however, does not mean that quarks do not exist in bound states inside elementary particles.* A direct experimental verification of this statement, however, has not been made.

Equality of positive and negative elementary charges. Negative, as well as positive, elementary charge was measured in the experiments described above. The results of these experiments proved their equality to the same degree of accuracy as the precision with which the values of a charge were measured. This is not a very high degree of accuracy. For example, it can be stated that a positive and a negative elementary charge differ in their absolute value by not more than one-tenth of their magnitude, i.e.

$$\frac{||e_+| - |e_-||}{|e_\pm|} \leqslant \frac{1}{10}. \quad (3.6)$$

www.EngineeringBooksPDF.com

This accuracy is quite unsatisfactory since the theory presumes an exact equality of absolute values of negative and positive elementary charges.

The accuracy of estimation can be immensely improved if we do not directly measure the value of an elementary charge. It is well known that equal numbers of protons and electrons are present in atoms. Bodies also contain the same number of electrons and protons, and hence the equality of a proton and an electron charge can be estimated by measuring the neutrality of bodies. This can be accomplished with an extremely high degree of accuracy since even a slight violation of neutrality results in enormous forces of electric interaction between bodies, which can be easily detected. Suppose, for example, that two iron balls of mass 1 g each are separated by 1 m and are not neutral because the charge of a proton differs from that of an electron by a millionth part of the absolute magnitude of the charge. Let us estimate the repulsive force between the spheres. Each gram of $^{26}_{56}$Fe contains $6 \times 10^{23} \times 26/56$ charges of each type. Consequently, a departure from neutrality just by 10^{-6} results in a charge

$$q = [1.6 \cdot 10^{-19} \cdot 10^{-6} \cdot 6 \cdot 10^{23} \cdot 26/56]\ \mathrm{C} = 4.46 \cdot 10^{-2}\ \mathrm{C} \qquad (3.7)$$

on each sphere. The repulsive force between these spheres is equal to

$$F = \frac{1}{4\pi\varepsilon_0}\frac{q^2}{r^2} = (4.46 \cdot 10^{-2})^2 \cdot 9 \cdot 10^9\ \mathrm{N} = 1.8 \cdot 10^7\ \mathrm{N} = 18\ \mathrm{MN}. \qquad (3.8)$$

This means that the repulsive force between the spheres is equal to the force exerted on the railway track by a goods train weighing 2000 tonnes. And this is the force resulting in just 2 grams of iron if the proton and electron charges differ by a millionth part! It is obvious that forces between two iron spheres which are extremely small in comparison with (3.8) can be easily measured. If such forces are not detected experimentally, this indicates a corresponding increase in the accuracy with which the absolute magnitude of the charge of an electron is equal to that of a proton. *At present, it has been experimentally established that the magnitude of the negative elementary charge of an electron is equal to the positive charge of a proton with a relative accuracy of* 10^{-21}, i.e.

$$\boxed{\frac{\|e_+| - |e_-\|}{|e_\pm|} \leqslant 10^{-21}.} \qquad (3.9)$$

The proof described above for the equality of absolute values of positive and negative elementary charges may appear to be not quite rigorous. We can imagine a body consisting of atoms or molecules, in which the elementary charges are not equal in magnitude, although their number in each atom or molecule is the same. In this case, atoms or molecules must have a charge. But the body as a whole may remain neutral if, in addition to these atoms or molecules, it contains the required number of electrons or positive ions (depending on the sign of the charge on atoms or molecules). Such an assumption, however, leads to complications which are difficult to reconcile with. For example, we have to discard the notion that bodies have a homogeneous structure, and accept that

www.EngineeringBooksPDF.com

their structure depends on their size, etc. Nevertheless, it is desirable to have a more straightforward and direct proof for the equality of absolute values of positive and negative elementary charges in atoms. Such a proof has actually been obtained.

The neutrality of individual atoms was verified by direct experiments. The deviation of a beam of neutral atoms in electrostatic fields was investigated. From the magnitude of deviation, we can determine the charge of the atom and draw conclusions about the equality of electron and proton charge in an atom. Investigations carried out on cesium ($Z = 55$) and potassium ($Z = 19$) beams have proved that the absolute values of the charge of an electron and a proton are equal with a relative accuracy of 3.5×10^{-19}.

Invariance of charge. The independence of the numerical value of an elementary charge of velocity is also proved by the fact that an atom is neutral. The difference in masses of an electron and a proton suggests that electrons move much faster in an atom than protons. *If the charge were dependent on velocity, the neutrality of atoms would be violated.* For example, electrons in a helium atom move about twice as fast as in a hydrogen molecule, while the neutrality of a helium atom and a hydrogen molecule has been proved to a very high accuracy. It can be concluded that with the same accuracy the charge is independent of velocity right up to the velocities of electrons in a helium atom, which is approximately equal to $0.02c$. In heavier atoms, whose neutrality has also been proved, electrons in inner shells move with velocities that are about half the velocity of light. Thus, it has been proved that the elementary charge is invariant up to $0.5c$. There are no reasons to believe that this is not so at higher velocities. Hence, the *invariance of electric charge is taken as an experimental foundation of the theory of electricity.*

The quest for quarks proved with high accuracy that fractional charges do not exist in nature. The absence of quarks in free state does not prove their nonexistence in bound state inside elementary particles.

What is the principle underlying the resonance method of measuring an elementary charge? What is the precision of this method at present? Give quantitative estimates.

Sec. 4. Electric Current

Basic concepts and values characterizing the distribution and motion of electric charges are discussed.

Motion of charges. The motion of electrons and protons involves the motion of their charges. Therefore, we can simply speak about the motion of charges without stipulating their carriers each time. This is not only convenient but also makes the consideration more general, since many phenomena depend only

www.EngineeringBooksPDF.com

on the charges and their motion and do not depend on the properties of charge carriers, say, their mass. When the properties of a charge carrier (for example, the mass of the carrier) are also important, besides the charge itself, we must take into account not only the charge but other characteristics of the charge carrier as well.

In the theory of electricity an elementary charge, including the charge of a proton, is assumed to be a point charge. The position of a charge, its velocity and acceleration have the same meaning as in the case of point particles.

Continuous distribution of charges. An elementary charge is very small. For this reason, *most macroscopic phenomena in electricity involve a huge number of electric charges, and their discrete nature is not manifested in any way.* For instance, each plate of a parallel-plate capacitor with a 10 μF capacitance contains about 7×10^{15} elementary charges at a potential difference of 100 V. About 6×10^{18} elementary charges pass each second through the cross-sectional area of a conductor carrying a current of 1 A. Hence *in most cases we can assume the charge to be continuously distributed in space and disregard its discreteness.*

Volume charge density. The volume density of a continuous distribution of charges is defined as the ratio of the charge to the volume occupied by it:

$$\rho = \frac{1}{\Delta V_{\text{ph}}} \sum_{\Delta V_{\text{ph}}} e_i = \frac{\Delta Q}{\Delta V_{\text{ph}}}, \tag{4.1}$$

where e_i are the elementary charges in the volume ΔV_{ph} (taking into account their sign) and ΔQ is the total charge on ΔV_{ph}. The volume ΔV_{ph} is small but not infinitely small in the mathematical sense. We speak of the volume ΔV_{ph} as an infinitely small volume in the physical sense. This means that it is very small and hence its position in space is characterized with a sufficiently high accuracy by the coordinate of a point lying inside this volume. In other words, for the argument of ρ on the left-hand side of (4.1) we can take the coordinates (x, y, z) of any point inside ΔV_{ph} and write $\rho(x, y, z)$. However, the volume ΔV_{ph} must contain a sufficiently large number of elementary charges so that a slight variation of this volume will not lead to a significant variation of density ρ calculated by formula (4.1). Consequently, ΔV_{ph} depends on specific conditions. *In some cases, a small volume ΔV may satisfy the required conditions and be considered as an infinitely small physical volume, but in other cases it may not.* Finally, *under some conditions there does not exist a volume ΔV which could be called an infinitely small physical volume. In this case, the concept of continuous distribution of charges cannot be used, and ρ in formula (4.1) cannot be defined as the volume density.* However, in most cases considered in the classical theory of electricity, the concept of continuous distribution of charge is valid.

When determining the volume density ρ with the help of formula (4.1), it can be considered as an ordinary mathematical function and the charge can be assumed to be continuously smeared over the volume. Then it follows from (4.1)

www.EngineeringBooksPDF.com

that the charge on the volume V is equal to

$$Q = \int_V \rho \, dV, \tag{4.2}$$

where dV is the differential of the volume.

Charge concentration. The *concentration of charges* of a certain sign is defined as the ratio of the number of charges to the volume occupied by them:

$$n_{\pm} = \frac{\Delta n_{\pm}}{\Delta V_{ph}}, \tag{4.3}$$

where $\Delta n_{\pm}$ is the number of charges of the appropriate sign in the volume ΔV_{ph}. Then [see (4.1)] we can write

$$\rho = \frac{1}{\Delta V_{ph}} \sum_{\Delta V_{ph}} e_i^{(+)} + \frac{1}{\Delta V_{ph}} \sum_{\Delta V_{ph}} e_i^{(-)}$$

$$= \frac{e^{(+)} \Delta n_{(+)}}{\Delta V_{ph}} + \frac{e^{(-)} \Delta n^{(-)}}{\Delta V_{ph}} = e^{(+)} n_{(+)} + e^{(-)} n_{(-)} = \rho^{(+)} + \rho^{(-)}, \tag{4.4}$$

where $e^{(+)}$ is the elementary point charge of the corresponding sign and $\rho^{(\pm)} = e^{(\pm)} n_{(\pm)}$ is the volume density of charges. An infinitely small physical volume must contain a sufficiently large amount of charges for the definition of concentration to be meaningful.

Surface charge density. Sometimes, charge is distributed in a very thin layer near a certain surface. If we are interested in the action of the charge at distances much longer than the thickness of the layer rather than in the processes within this layer, we can assume that the entire charge is concentrated on the surface. In other words, this very thin layer may be assumed to be the surface. The surface charge density is defined as

$$\boxed{\sigma = \frac{1}{\Delta S_{ph}} \sum_{\Delta S_{ph}} e_i = \frac{\Delta Q}{\Delta S_{ph}},} \tag{4.5}$$

where ΔS_{ph} is an infinitely small surface area in the physical sense, and ΔQ is the charge on the surface area ΔS_{ph} of a thin layer adjacent to it.

For the argument of σ we can take the coordinates of points of the surface and treat it as a function of these coordinates. The substantiation and the meaning of this procedure are the same as for the volume charge density ρ in (4.1). Consequently, the total charge on the surface S is

$$Q = \int_S \sigma \, dS, \tag{4.6}$$

where dS is the differential of the surface area.

Current density. The charges contained in a volume ΔV_{ph} move with velocities which differ in magnitude and in direction. *The motion of a charge results in a transport of the charge in the direction of its velocity.* Con-

www.EngineeringBooksPDF.com

sequently, various movements of charges contained in the volume ΔV_{ph} result in a certain average transport of the charge contained in this volume. The intensity of this transport is characterized by the current density defined as

$$\boxed{\mathbf{j} = \frac{1}{\Delta V_{ph}} \sum_{\Delta V_{ph}} e_i \mathbf{v}_i,} \tag{4.7}$$

where $\mathbf{v}_i$ is the velocity of the charge e_i.

Dividing the sum in (4.7) into the sums over positive and negative charges, we obtain

$$\mathbf{j} = \frac{1}{\Delta V_{ph}} \sum_i e_i^{(+)} \mathbf{v}_i^{(+)} + \frac{1}{\Delta V_{ph}} \sum_i e_i^{(-)} \mathbf{v}_i^{(-)} = \frac{e^{(+)}}{\Delta V_{ph}} \sum_i \mathbf{v}_i^{(+)} + \frac{e^{(-)}}{\Delta V_{ph}} \sum_i \mathbf{v}_i^{(-)}. \tag{4.8}$$

Formula (4.8) becomes more clear if we express the quantities appearing in it in terms of average velocities and concentrations of charges:

$$\sum_i \mathbf{v}_i^{(+)} = \Delta n^{(+)} \frac{1}{\Delta n^{(+)}} \sum_i \mathbf{v}_i^{(+)} = \Delta n^{(+)} \langle \mathbf{v}^{(+)} \rangle,$$

where

$$\langle \mathbf{v}^{(+)} \rangle = \frac{1}{\Delta n^{(+)}} \sum_i \mathbf{v}_i^{(+)},$$

since $\Delta n^{(+)}$ is the number of charges the sum of whose velocities appears under the summation sign Σ. The sum over the velocities of negative charges is transformed in a similar way. As a result, formula (4.8) becomes

$$\mathbf{j} = e^{(+)} \frac{\Delta n^{(+)}}{\Delta V_{ph}} \langle \mathbf{v}^{(+)} \rangle + e^{(-)} \frac{\Delta n^{(-)}}{\Delta V_{ph}} \langle \mathbf{v}^{(-)} \rangle = e^{(+)} n^{(+)} \langle \mathbf{v}^{(+)} \rangle + e^{(-)} n^{(-)} \langle \mathbf{v}^{(-)} \rangle$$
$$= \rho^{(+)} \langle \mathbf{v}^{(+)} \rangle + \rho^{(-)} \langle \mathbf{v}^{(-)} \rangle, \tag{4.9}$$

where we took into account relations (4.3) and (4.4). Thus, negative and positive charges generate their own current densities

$$\mathbf{j}^{(+)} = \rho^{(+)} \langle \mathbf{v}^{(+)} \rangle, \quad \mathbf{j}^{(-)} = \rho^{(-)} \langle \mathbf{v}^{(-)} \rangle,$$
$$\mathbf{j} = \mathbf{j}^{(+)} + \mathbf{j}^{(-)}. \tag{4.10}$$

The direction of current density of positive charges coincides with the direction of their average velocity, while for negative charges, the current density has a direction opposite to that of the average velocity.

For the sake of simplification, formulas (4.10) are usually represented in the form

$$\boxed{\mathbf{j} = \rho \mathbf{v},} \tag{4.11}$$

where ρ and $\mathbf{v}$ are the volume density and the velocity of the charges of the corresponding sign. If current is generated by charges of both signs, then the right-

www.EngineeringBooksPDF.com

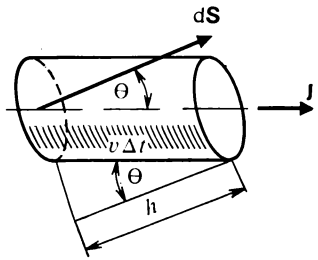


Fig. 9. Calculation of current through a surface area element.

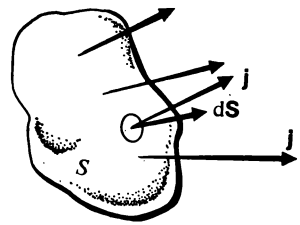


Fig. 10. Electric current through a surface.

hand side is assumed to contain the sum of two terms corresponding to positive and negative charges. However, in most cases considered in the theory of electricity, the current is due only to the motion of negative charges (electrons), and hence the right-hand side of (4.11) contains only the product of the negative volume charge density of electrons and their average velocity. *The transfer of a negative charge against the velocity is equivalent to the transfer of an equal positive charge in the direction of the velocity.* While analyzing various situations, it is more convenient to assume that the current is due to the motion of positive charges since their displacement in space coincides with the direction of the current density.

Current through a surface. An infinitely small surface element is characterized by the vector dS whose magnitude is equal to the area of the surface element and which is directed along the positive normal to the surface.

Let us calculate the charge which crosses the surface element dS during the time dt (Fig. 9). The displacement of the charge during this time is equal to $\mathbf{v}\,dt$. Consequently, the charge crossing dS is equal to the volume charge density multiplied by the volume of the oblique cylinder (Fig. 9). The area of the base and the height of this cylinder are equal to dS and $h = v\,\Delta t \cos\theta$. Consequently, the charge crossing the surface dS is equal to

$$dq = \rho v\,dt\,dS\cos\theta = dt\,j\,dS\cos\theta = dt\,\mathbf{j}\cdot d\mathbf{S}, \tag{4.12}$$

where $\mathbf{j}\cdot d\mathbf{S} = j\,dS\cos(\widehat{\mathbf{j}, d\mathbf{S}})$. The *current through* a surface is defined as the charge crossing the surface per unit time. Hence an infinitely small current dI crossing the surface element dS [see (4.12)] is given by

$$dI = dQ/dt = \mathbf{j}\cdot d\mathbf{S}. \tag{4.13}$$

The current flowing through a finite surface S (Fig. 10) is equal to the integral of the current elements (4.13) over this surface

$$\boxed{I = \int_S dI = \int_S \mathbf{j}\cdot d\mathbf{S}.} \tag{4.14}$$

www.EngineeringBooksPDF.com

If a direct electric current flows in a conductor, formula (4.14) defines the current as the quantity of electricity flowing per second through the conductor cross section.

Most of the macroscopic phenomena investigated in electricity involve an enormous number of electric charges and their discreteness is not manifested in any way. In some cases, a certain small volume can be considered as an infinitely small physical volume, while in some other cases this assumption is not true. Under certain conditions, no volume can be treated as an infinitely small physical volume. In this case, we cannot go over to a representation of a continuous charge distribution in a volume.

Sec. 5. Law of Charge Conservation

Two aspects of the concept of charge conservation are discussed. Integral and differential formulations of the law of charge conservation are given.

Two aspects of the concept of charge conservation. The concept of "charge conservation" includes two groups of entirely different facts: (1) *electrons and protons are material particles with an infinite lifetime, their elementary electric charges are invariant and do not depend on velocity.* Consequently, these charges remain unchanged as long as an electron and a proton exist, irrespective of the way in which they move. In other words, the charge is conserved under any type of motion. *In this aspect the law of charge conservation is just a consequence of the indestructibility of charge carriers as physical objects, and of the invariance of charge*; (2) *besides protons and electrons there exist a large number of other charged elementary particles. All these particles are created, create other particles and are annihilated in various interconversion processes. The entire multitude of experimental data indicates that whatever the process of interconversion of particles, the total charge of the particles before interconversion is equal to the total charge of the particles after the conversion.* For example, in the case of β-decay, the nucleus has a certain positive charge $Ze^{(+)}$. After the emission of an electron, the positive charge of the nucleus increases by one elementary positive charge and becomes equal to $(Z + 1)\, e^{(+)}$. However, together with the negative charge of the emitted electron, the system "nucleus + electron" has the same charge as before:

$$(Z + 1)e^{(+)} - |\, e^{(-)}\, | = Ze^{(+)}.$$

By way of another example, we can consider the creation of an electron-positron pair by a gamma-ray photon. The initial particle, viz. the gamma-ray photon, is electrically neutral and is transformed into a pair of particles whose total charge is again equal to zero. This has been proved to a high degree of accuracy during the measurement of positive charge on a positron. A vast number of cases of interconversion of elementary particles have been investigated and the

www.EngineeringBooksPDF.com

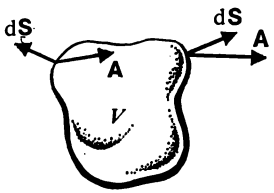


Fig. 11. The outward normal is the positive normal to a closed surface.

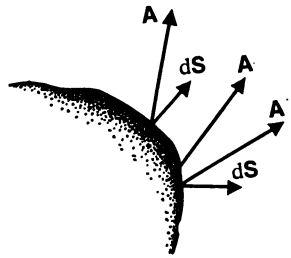


Fig. 12. Flux of vector A through a surface.

total charge in each case before and after the process remains the same. In other words, the law of conservation of charge is obeyed. Consequently, the charge acquires its individual existence in a certain sense independently of its carrier, and the law of its conservation can be formulated as follows: *In all processes associated with the motion of charge carriers, the charge is always conserved.*

In spite of its relative independence, however, a charge cannot exist without a charge carrier, or beyond space and time. This means that a *charge is not an independent entity, capable of existence without matter. Rather, it is a property of matter.* Finding the nature of this connection is one of the most difficult problems of modern physics. It is not yet clear as to why there exists just one elementary charge and why it is equal to $|e|$ and not to some other value.

Integral form of charge conservation law. Considering charge conservation as an experimental fact, we can express it as the statement that the charge in a certain volume V can change only if charge flows into, or out of, a closed contour S bounding the volume V:

$$\boxed{\frac{\partial}{\partial t}\int\limits_V \rho\, dV = -\oint\limits_S \mathbf{j}\cdot d\mathbf{S}.} \tag{5.1}$$

The left-hand side of this equation defines the rate of variation of charge in the volume, and the right-hand side is the strength of current flowing through the surface bounding this volume. *The negative sign indicates that if the positive charge inside the volume V decreases, the current density is directed outwards from the volume.* It should be recalled that the outward normal is considered positive for closed surfaces. Consequently, vector dS in (5.1) is directed along the outward normal to the surface (Fig. 11).

Divergence. The mathematical concept of divergence plays an important role in the description of processes associated with the creation, annihilation and conservation of physical quantities.

Suppose that a certain vector $\mathbf{A}(x, y, z)$ is defined at all points in space. We consider a surface S (Fig. 12). The integral

$$\Phi_A = \int\limits_S \mathbf{A}\cdot d\mathbf{S} \tag{5.2}$$

www.EngineeringBooksPDF.com

is called the **flux of the vector A across the surface** S. This term is due to the following reason. Suppose that we have a fire whose smoke has a density ρ and a velocity v at different points in space. We choose the quantity $\rho\mathbf{v}$ as the vector A. In this case, integral (5.2), together with Fig. 10, gives the mass of smoke passing through surface S per second. A similar concept was applied to an electric charge in Eq. (4.14). In analogy with (5.1), we conclude that the flux of vector A through a closed surface characterizes the intensity of creation or annihilation of A inside the volume bounded by the surface. Thus, the vector flux $\rho\mathbf{v}$ across the closed surface characterizes the intensity of smoke created within the volume bounded by the closed surface. When applied to electric charge, Eq. (5.1) can be interpreted in the same way. It can be stated that the integral (5.2) describes the total power of the sources of vector A inside the volume.

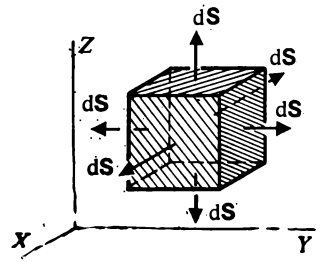


Fig. 13. The flux of a vector through the surface of a cube is the sum of the fluxes through its faces.

Divergence characterizes the power of sources and is defined by the formula

$$\operatorname{div}\mathbf{A} = \lim_{\Delta V \to 0} \frac{\oint_{\Delta S} \mathbf{A}\cdot d\mathbf{S}}{\Delta V} \tag{5.3}$$

where ΔS is an infinitely small closed surface bounding an infinitely small volume ΔV.

Let us find an expression for div A in Cartesian coordinates. For this purpose, we calculate the flux of vector A across the surface of a cube with sides Δx, Δy, Δz (Fig. 13) having its centre at the point (x, y, z). The coordinates of the midpoints of the faces are $(x + \Delta x/2, y, z)$, $(x - \Delta x/2, y, z)$, $(x, y + \Delta y/2, z)$, $(x, y - \Delta y/2, z)$, $(x, y, z + \Delta z/2)$, $(x, y, z - \Delta z/2)$. The integrand of Eq. (5.3) in coordinates has the form

$$\mathbf{A}\cdot d\mathbf{S} = A_x dS_x + A_y dS_y + A_z dS_z, \tag{5.4}$$

where

$$dS_x = \pm dy\, dz,\ dS_y = \pm dz\, dx,\ dS_z = \pm dx\, dy. \tag{5.5}$$

The sign of these quantities is determined by the direction of the outward normal to the face with respect to the positive direction of the corresponding axis. For example, dS_y has a positive value over the right face $(x, y + \Delta y, z)$ and a negative value over the left face. The integral over the surface of the cube is reduced to the sum of integrals over its faces.

Let us calculate, for example, the integral over the faces parallel to Y-axis. On these faces, $dS_x = 0$, $dS_y = \pm dz\, dx$, $dS_z = 0$ and, consequently, the sum on the right-hand side of (5.4) is reduced to a single term $A_y dS_y$. Denoting the surface area of the left and right faces by ΔS_{y1} and ΔS_{y2}, respectively, we can

www.EngineeringBooksPDF.com

write

$$I_y = \oint\limits_{\Delta S_{y1}+\Delta S_{y2}} \mathbf{A}\cdot d\mathbf{S} = \int\limits_{\Delta S_{y1}} A_y\, dS_y + \int\limits_{\Delta S_{y2}} A_y\, dS_y$$

$$= \int\limits_{\Delta S_{y1}} -A_y(x,\ y-\Delta y/2,\ z)\, dx\, dz + \int\limits_{\Delta S_{y2}} A_y(x,\ y+\Delta y/2,\ z)\, dx\, dz. \quad (5.6)$$

The negative sign of the first integral on the right-hand side of this equation takes into account the fact that the outward normal to the left face ΔS_{y1} is directed toward the negative values of y. For further calculations, we express A_y in the form of a Taylor series in Δy:

$$A_y(x,\ y+\Delta y/2,\ z) = A(x,\ y,\ z)$$
$$+ (\Delta y/2)\, \partial A_y(x,\ y,\ z)/\partial y + 0[(\Delta y)^2],$$
$$A_y(x,\ y-\Delta y/2,\ z) = A(x,\ y,\ z)$$
$$-(\Delta y/2)\, \partial A_y(x,\ y,\ z)/\partial y + 0[(\Delta y)^2], \quad (5.7)$$

where $O[(\Delta y)^2]$ are terms of highest order of infinitesimals in Δy. Substituting (5.7) into (5.6), we get

$$I_y = \Delta y \int\limits_{\Delta S_y} \frac{\partial A_y(x,\ y,\ z)}{\partial y}\, dx\, dz + 0[(\Delta y)^2], \quad (5.8)$$

where we have taken into account the fact that the surface areas ΔS_{y1} and ΔS_{y2} are equal and have the same coordinates along X- and Z-axes.

The integral in (5.8) can be calculated by expanding the integrand into a series assuming x and z as variables of integration rather than the coordinates of centres of the cube faces. If x and z denote the coordinates of faces of the cube, it is convenient to replace the variables according to formulas

$$x \to x + \xi,\ z \to z + \eta,\ dx\ dz \to d\xi\, d\eta, \quad (5.9)$$

$$\int\limits_{\Delta S_y} \frac{\partial A_y(x,\ y,\ z)}{\partial y}\, dx\, dz = \int\limits_{\Delta S_y} \frac{\partial A_y(x+\xi,\ y,\ z+\eta)}{\partial y}\, d\xi\, d\eta, \quad (5.10)$$

where x, y and z on the right-hand side of (5.10) are the coordinates of centres of the faces and remain constant in calculations of (5.10). The expression $\partial A_y/\partial y$ can be expanded into a series in ξ and η:

$$\frac{\partial A_y(x+\xi,\ y,\ z+\eta)}{\partial y} = \frac{\partial A_y(x,\ y,\ z)}{\partial y} + \xi \frac{\partial^2 A_y(x,\ y,\ z)}{\partial x\, \partial y}$$
$$+ \eta \frac{\partial^2 A_y(x,\ y,\ z)}{\partial z\, \partial y} + 0(\xi^2,\ \eta^2), \quad (5.11)$$

where ξ and η vary from 0 to $\pm\Delta x/2$ and $\pm\Delta z/2$ during integration and, consequently, have the same order of infinitesimals as Δx and Δz. Substituting (5.11)

www.EngineeringBooksPDF.com

into (5.10), we get

$$\int_{\Delta S_y} \frac{\partial A_y(x+\xi,\ y+z+\eta)}{\partial y}\,d\xi\,d\eta = \frac{\partial A_y}{\partial y}\int_{\Delta S_y} d\xi\,d\eta + \frac{\partial^2 A_y}{\partial x\,\partial y}\int_{\Delta S_y} \xi\,d\xi\,d\eta$$

$$+\frac{\partial^2 A_y}{\partial z\,\partial y}\int_{\Delta S_y} \eta\,d\xi\,d\eta + \ldots = \frac{\partial A_y}{\partial y}\Delta x\,\Delta y\,\Delta z + 0\,[(\Delta x)^2,\ (\Delta z)^2]. \quad (5.12)$$

This gives the following expression for (5.8):

$$I_y = \frac{\partial A_y(x,\ y,\ z)}{\partial y}\,\Delta x\,\Delta y\,\Delta z + 0\,[(\Delta x\,\Delta y\,\Delta z)^2]. \quad (5.13)$$

The flux across other pairs of faces is calculated in a similar manner:

$$\oint_S \mathbf{A}\cdot d\mathbf{S} = \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right)\Delta x\,\Delta y\,\Delta z + 0\,[(\Delta x\,\Delta y\,\Delta z)^2]. \quad (5.14)$$

Substituting (5.14) into (5.3) and considering that the volume ΔV of the cube is equal to $\Delta x\,\Delta y\,\Delta z$, we obtain

$$\operatorname{div}\mathbf{A} = \lim_{\Delta V\to 0}\left\{\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} + 0\,[(\Delta x\,\Delta y\,\Delta z)^2]/(\Delta x\,\Delta y\,\Delta z)\right\}$$

$$= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}, \quad (5.15)$$

since the term depending on $(\Delta x\ \Delta y\ \Delta z)$ vanishes as we proceed to the limit. The formula

$$\boxed{\operatorname{div}\mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}} \quad (5.16)$$

allows us to calculate the divergence in Cartesian coordinates.

Gauss' theorem. This theorem relates the *power of the sources to the fluxes of vectors generated by them*, and plays an important role in the theory of electricity. We divide the volume V bounded by the surface S (Fig. 14*a*) into a large number of volumes ΔV_i with surfaces ΔS_i.

Formula (5.3) can be written in the form

$$(\operatorname{div}\mathbf{A})_i\,\Delta V_i \approx \oint_{\Delta S_i} \mathbf{A}\cdot d\mathbf{S}, \quad (5.17)$$

where $(\operatorname{div}\mathbf{A})_i$ denotes div **A** in the ith volume. The approximate equality used in (5.17) indicates that although ΔV_i is small, it has a finite value. Upon an indefinite decrease in ΔV_i, the relation (5.17) becomes an exact equality.

www.EngineeringBooksPDF.com

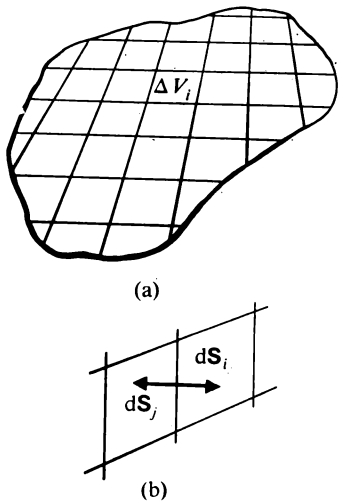


Fig. 14. To the derivation of the Gauss theorem.

Summing both sides of (5.17) over all the cells of volume V, we get

$$\sum_i (\operatorname{div} \mathbf{A})_i \Delta V_i \approx \sum \oint_{\Delta S_i} \mathbf{A} \cdot d\mathbf{S}. \qquad (5.18)$$

The sum on the right-hand side can be transformed as follows. The adjacent cells have a common surface of contact. The entire surface of inner cells is in contact with the adjoining cells. Hence, each surface integral in the volume V appears twice as the integral over adjacent parts of the neighbouring cells (in the sum on the right-hand side of (5.18), see Fig. 14*b*; $d\mathbf{S}_i$ is opposite to $d\mathbf{S}_j$). Since the normals in each pair of these integrals are in opposite directions, and the vector $\mathbf{A}$ has the same magnitude, these integrals are equal in absolute value and opposite in sign. Consequently, their sum is equal to zero, and hence the sum of all integrals on the right-hand side of (5.18) over the contact surface of cells within the volume V is equal to zero. This leaves only the sum of the integrals over those parts of cells on the boundary of volume V which are not in contact with other cells. The sum of the areas of these outer surfaces of cells lying on the boundary of volume V is equal to the surface area S bounding the volume V. Consequently,

$$\sum_i \oint_{\Delta S_i} \mathbf{A} \cdot d\mathbf{S} = \int_S \mathbf{A} \cdot d\mathbf{S}. \qquad (5.19)$$

This is an exact equality which is valid for any division of the volume V into cells ΔV_i.

For $\Delta V_i \to 0$, the left-hand side of (5.18) can be expressed in the form of the integral:

$$\lim_{\Delta V_i \to 0} \sum_{\Delta V_i} (\operatorname{div} \mathbf{A})_i \Delta V_i = \int_V \operatorname{div} \mathbf{A}\, dV. \qquad (5.20)$$

Substituting (5.19) into (5.18) and proceeding to the limit, we obtain the formula

$$\boxed{\int_V \operatorname{div} \mathbf{A}\, dV = \oint_S \mathbf{A} \cdot d\mathbf{S}.} \qquad (5.21)$$

This is the formulation of **Gauss' theorem.** *It connects the volume integral of divergence of a vector with the flux of this vector across the closed surface bounding this volume.* The conditions of applicability of this theorem are indicated in mathe-

www.EngineeringBooksPDF.com

matics and will not be specified here since they are automatically satisfied in most real physical systems.

Differential form of charge conservation law. Volume V and surface area S in formula (5.1) do not change with time. Consequently, the time derivative on the left-hand side of (5.1) can be included in the integral. On the other hand, the right-hand side of this equality can be transformed into a volume integral in accordance with Gauss' theorem:

$$\frac{\partial}{\partial t}\int_V \rho\, dV_i = \int_V \frac{\partial \rho}{\partial t}\, dV, \quad \oint_S \mathbf{j}\cdot d\mathbf{S} = \int_V \operatorname{div}\mathbf{j}\, dV. \tag{5.22}$$

Transposing all terms in (5.1) to the left-hand side and taking into consideration (5.22), we get

$$\int_V \left(\frac{\partial \rho}{\partial t} + \operatorname{div}\mathbf{j}\right) dV = 0. \tag{5.23}$$

This equality is valid for any volume. The integrand is obviously equal to zero. The proof is obtained by contraction. If the integrand is not equal to zero at some point, we can take for V a small volume around this point in which the sign of the integrand remains unchanged. This, however, is in contradiction to Eq. (5.23). Consequently, the integrand is equal to zero at all points. In this case,

$$\boxed{\frac{\partial \rho}{\partial t} + \operatorname{div}\mathbf{j} = 0.} \tag{5.24}$$

This is the differential form of the law of charge conservation. It is also called the **continuity equation.**

Charge is conserved in all motions and interconversions of charge carriers.
The power of a source is characterized by divergence. The Gauss theorem connects the total power of sources in a volume with the flux of the vector generated by the sources through the surfaces bounding this volume.
Charge is not a concept independent of matter. Rather, it is a property of matter.

What requirements must an infinitely small physical volume meet?
Under what conditions can the concept of continuous charge distribution be used? Is it always possible to determine the volume charge density? Give examples.
Under what conditions can the concept of surface charge be used?
What is the relation between the direction of the current density vector and that of the charge velocity vector?

Which two groups of different facts are described by the concept of charge conservation?
What is the physical meaning of the equation expressed by the Gauss theorem?
What condition must be satisfied so that vanishing of an integral results in vanishing of the integrand?

www.EngineeringBooksPDF.com

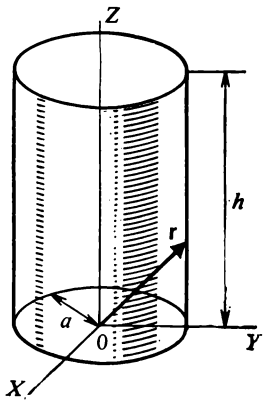


Fig. 15. Calculation of the flux of a radius vector through the surface of a right circular cylinder.

Example 5.1. *Calculate the flux of the radius vector across the surface of a right circular cylinder* (Fig. 15). *The calculations should be carried out directly and with the help of Gauss' theorem.*

We take the centre of the cylinder base as the origin and direct the Z-axis along the axis of the cylinder (Fig. 15). In this case,

$$\int_S \mathbf{r}\cdot d\mathbf{S} = \int_{S_l} \mathbf{r}\cdot d\mathbf{S} + \int_{S_u} \mathbf{r}\cdot d\mathbf{S} + \int_{S_{lat}} \mathbf{r}\cdot d\mathbf{S},$$

where S_l and S_u are the areas of the lower and upper bases of the cylinder, while S_{lat} is its lateral surface area. We have

$$\int_{S_l} \mathbf{r}\cdot d\mathbf{S} = 0, \quad \int_{S_u} \mathbf{r}\cdot d\mathbf{S} = h\pi a^2,$$

since

$$\mathbf{r}\cdot d\mathbf{S} = r\, dS \cos(\widehat{\mathbf{r}, d\mathbf{S}}) = 0,$$

and

$$\mathbf{r}\cdot d\mathbf{S} = r\, dS \cos(\widehat{\mathbf{r}, d\mathbf{S}}) = h\, dS$$

for points on the lower and upper bases. Finally, for the integral over the lateral surface, we have $\int_{S_{lat}} \mathbf{r}\cdot d\mathbf{S} = a2\pi ah$, since $\mathbf{r}\cdot d\mathbf{S} = a\, dS$ for points on the lateral surface. Consequently,

$$\int_S \mathbf{r}\cdot d\mathbf{S} = 3\pi a^2 h. \tag{5.25}$$

According to Gauss' theorem, we have

$$\int_S \mathbf{r}\cdot d\mathbf{S} = \int_V \operatorname{div} \mathbf{r}\, dV = 3\pi a^2 h, \tag{5.26}$$

where $\operatorname{div} \mathbf{r} = 3$ and $V = \pi a^2 h$ (volume of a right circular cylinder).

Sec. 6. Coulomb's Law

The accuracy of experimental verification of Coulomb's law is discussed.

Experimental verification of Coulomb's law. Coulomb's law for the force F of interaction between two point charges q_1 and q_2 separated by a distance r has the form

$$F = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r^2}, \tag{6.1}$$

www.EngineeringBooksPDF.com

where $\varepsilon_0 = 1/(4\pi \times 9 \times 10^9)$ F/m. This law was established by Coulomb (1736-1806) in 1785 from direct measurement of forces of interaction between charged bodies whose dimensions are much smaller than the distance between them. The accuracy of these experiments was not very high and it was only on the basis of general concepts emerging from an analogy with the forces of attraction that the absolute accuracy of this law could be believed.

Coulomb's law (6.1) is one of the fundamental experimental facts on which the study of electricity is based. The verification of its validity and the determination of the limits of its applicability are significant problems and considerable efforts were devoted to these investigations by experimental physicists.

A verification of the law (6.1) with a very high degree of accuracy is difficult through a direct measurement of the forces of interaction on account of the fact that the scientists do not have at their disposal point charges at rest. Consequently, the experimental results can be associated with corollaries of Coulomb's law, and this serves as the basis for ascertaining the limits of applicability and the accuracy of this law.

The first experimental verification of this law was made in 1772 by Cavendish (1731-1810) thirteen years before it was actually discovered by Coulomb. However, Cavendish did not publish his results and thus lost his claim to this discovery. The manuscript containing a description of his experiments was found in the archives only in 1860's. The Cavendish method was widely used and has recently led to the experimental verification of Coulomb's law with a high degree of accuracy.

The problem of experimental verification is formulated as follows. The law of interaction is expressed in the form

$$F = \frac{\text{const}}{r^{2+\alpha}} \tag{6.2}$$

It is required to find the order of smallness of α. The smaller the value of $|\alpha|$, the closer the law of interaction to Coulomb's law. Hence the experimental result is expressed in the form of a constraint on α: $|\alpha| \leqslant \delta$. The aim of the experiment is to find the value of δ.

The Cavendish method. Free charges in a homogeneous conductor are located on its surface. At first glance, it seems to be a consequence of the repulsion of like charges, which makes them move apart to the maximum possible distances, i.e. to the surface. This, however, is not true. *Such a situation arises due to the fact that the force of interaction between point charges decreases exactly in inverse proportion to the square of the distance between them, and in no other manner.*

It is known from the theory of gravitation that a spherical homogeneous layer of a substance does not create any force in a cavity surrounded by this layer. Consequently, if point charges interact in accordance with the inverse quadratic law, a spherical layer of charges does not create any force in this cavity.

Suppose that a charge is distributed uniformly over a sphere with a surface density σ (Fig. 16). At a point P inside the sphere, the charges on the surface elements dS_1 and dS_2 create forces $dF_1 = \sigma\, dS_1/(4\pi\varepsilon_0 r_1^2)$ and $dF_2 = \sigma\, dS_2/(4\pi\varepsilon_0 r_2^2)$ which are directed oppositely. It follows from the property of tangents

www.EngineeringBooksPDF.com

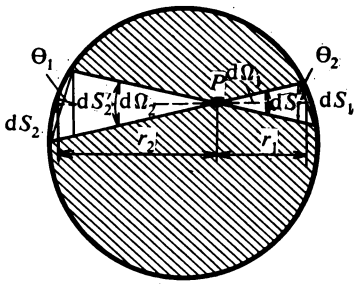


Fig. 16. To the theory of the Cavendish method.

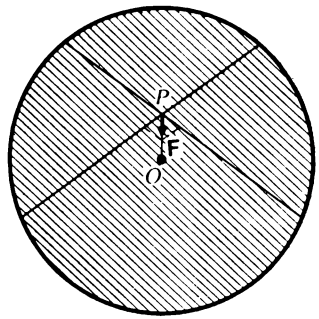


Fig. 17. Emergence of a force due to a spherical layer at points inside the sphere.

at the ends of a chord that angles θ_1 and θ_2 between the perpendiculars to the chord and the surface elements dS_1 and dS_2 are equal. In this case. $dS_1 = dS_2'/\cos\theta$ and $dS_2 = dS_2'/\cos\theta$. Consequently, $dF_1 = \sigma\, dS_1'/(4\pi\varepsilon_0 r_1^2 \cos\theta)$, and $dF_2 = \sigma\, dS_2'/(4\pi\varepsilon_0 r_2^2 \cos\theta)$, where $dS_1'/r_1^2 = d\Omega_1$ and $dS_2'/r_2^2 = d\Omega_2$ are the solid angles at which dS_1 and dS_2 are seen from the point P (they are equal by construction). Thus, forces dF_1 and dF_2 are equal in magnitude but are directed oppositely, since the charge on dS_1 is the same as the charge on dS_2. This results in a mutual compensation of forces from all pairs of opposite surface elements, and the total force acting on a test charge at the point P is equal to zero.

A charge imparted to a conducting sphere is distributed uniformly over its surface due to spherical symmetry. The absence of charge inside the volume is proved in the following manner. Suppose that certain charges exist inside the sphere. In view of spherical symmetry, the distribution of these charges must be spherically symmetric. Let us consider a spherical layer of charges. These charges are not acted upon by any force from charges located outside the cavity bounded by the spherical layer but they are influenced by the repulsive forces from the charges located inside such a cavity. This means that the spherical layer of charges starts moving from the centre to the periphery. Consequently, under equilibrium distribution, there are no charges in a conducting sphere.

The situation is quite different if the interaction between charges does not obey Coulomb's law. In this case, the following forces act at point P due to charges $\sigma\, dS_1$ and $\sigma\, dS_2$ located on the surface elements dS_1 and dS_2:

$$dF_1 = \text{const}\,\frac{dS_1\sigma}{r_1^{2+\alpha}} = \frac{\text{const}\cdot\sigma}{\cos\theta}\, d\Omega_1 \frac{1}{r_1^{\alpha}},$$
$$dF_2 = \text{const}\,\frac{dS_2\sigma}{r_2^{2+\alpha}} = \frac{\text{const}\cdot\sigma}{\cos\theta}\, d\Omega_2 \frac{1}{r_2^{\alpha}}. \qquad (6.3)$$

www.EngineeringBooksPDF.com

The resultant

$$\Delta F = A\left(\frac{1}{r_2^{\alpha}} - \frac{1}{r_1^{\alpha}}\right) \tag{6.4}$$

of these forces is not equal to zero. In formula (6.4) A denotes equal coefficients of $1/r_1^{\alpha}$ and $1/r_2^{\alpha}$ in (6.3).

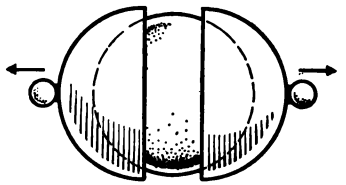


Fig. 18. The Cavendish method for verifying Coulomb's law.

The existence of the force ΔF facilitates a uniform distribution of charges over the entire volume of the conducting sphere. This is so because the charge inside the sphere is subjected not only to forces from the inner spherical layers but also to external forces whose nature depends on the sign of α.

Let us consider the case when $\alpha > 0$. Here, the force due to a charge ($\sigma > 0$) located at a more distant surface element from point P (Fig. 16) is smaller than the force due to a charge located at a nearer surface element. Consequently the force is directed towards the more distant surface element. Adding possible pairs of surface elements, we conclude that the resultant force $\mathbf{F}$ is directed towards the centre O (Fig. 17). Consequently, we can distribute a charge inside a sphere of radius OP so that the force due to this distribution at point P compensates the force due to charges in the outer spherical layers. As a result, the layer of charges on a sphere of radius OP may be in equilibrium. We should choose the radial distribution of charge density so that the force at each point inside the sphere is equal to zero. This is the equilibrium distribution. Hence, for $\alpha > 0$, charges in a charged conducting sphere are located not only on the surface as for $\alpha = 0$, but also inside the sphere. A similar result is obtained for $\alpha < 0$. We can carry out a more detailed numerical analysis and determine the charge in the volume of the sphere as a function of α. The Cavendish method involves the measurement of charge in the volume of a sphere and a subsequent calculation of the value of α.

A conducting spherical shell consisting of two hemispheres tightly clasps a conducting sphere (Fig. 18) thus imparting an electric charge to the system. The shell is then detached from the sphere with the help of insulated handles, and the charge remaining on the sphere is investigated.

If Coulomb's law holds, the entire charge is located on the shell and is removed together with it. The charge remaining on the sphere is equal to zero.

If Coulomb's law is violated, a part of the charge is concentrated in the volume of the sphere, while the remaining charge is located on the shell. After the removal of the shell, some charge is left on the sphere. By finding the value of this charge, we can estimate α. Of course, we can measure the potential instead of the charge in actual experiments. This, however, does not alter the state of affairs.

Cavendish found that $|\alpha| \leqslant 0.02$. Similar experiments carried out by Maxwell about a hundred years later gave a value of $|\alpha| \leqslant 5 \times 10^{-5}$. The Cavendish method was perfected in 1971. The experiment was carried out not under static conditions but with potentials varying in time. The apparatus consisted of two concentric conducting spheres. An alternating voltage of ± 10 kV with respect to the earth was supplied to the outer sphere. If Coulomb's law were violated, the

www.EngineeringBooksPDF.com

potential on the inner sphere would vary relative to the earth. The researchers were able to detect potential differences less than 1 pV. Since no potential oscillations were observed on the inner sphere, it could be assumed that $|\alpha| \leqslant |2.7 \pm 3.1| \times 10^{-16}$.

These experiments *confirmed the validity of Coulomb's law, for distances from several millimetres to tens of centimetres with an extremely high degree of accuracy indicated above.*

Verification of Coulomb's law for large distances. It is difficult to apply the Cavendish method to verify Coulomb's law for distances of several metres and larger. In this case, indirect methods are used, whose substantiation is beyond the scope of the classical theory of electricity. They are based on quantum-mechanical concepts about the interaction between particles, taking into account their wave properties. Each interaction involves a certain type of particles. The law of interaction depends on the properties of particles responsible for this interaction and, above all, on their mass. If the rest mass of particles responsible for an interaction is zero, the force of interaction is inversely proportional to the square of the distance, while the interaction potential is inversely proportional to the distance. If, however, the rest mass of interacting particles differs from zero, the interaction potential varies in proportion to $\sim(1/r)\exp(-\mu r)$, where μ depends on the rest mass of the particles. For the zero rest mass, μ is equal to zero, and the potential varies in inverse proportion to the distance, as it should be when the Coulomb law and Newton's gravitation law are valid. *According to modern concepts, electromagnetic interactions are due to photons. Consequently, the verification of the validity of Coulomb's law is reduced to proving that the rest mass of photons is equal to zero.*

Besides corpuscular properties, all particles possess wave properties. The energy ε_{ph} of photons is connected with their frequency and mass through the relations $\varepsilon_{ph} = \hbar\omega$ and $\varepsilon_{ph} = m_\gamma c^2$, where $\hbar = 1.05 \times 10^{-34}$ J·s is Planck's constant and m_γ is the photon mass. This mass is larger than the rest mass, if a photon had one. Consequently, having found the upper limit for m_γ, we obtain the constraint for the photon rest mass. Having proved the existence of electromagnetic waves with a sufficiently large wavelength experimentally, we can state that the value of m_γ is quite small. If we could verify the existence of electromagnetic waves with an infinite wavelength, we would be able to state that the photon rest mass is equal to zero, and that Coulomb's law is valid.

The longest electromagnetic waves that can be observed at present are formed as standing waves in the space between the surface of the Earth and the ionosphere. They are called the **Schumann resonances**. The smallest Schumann resonance corresponds to the frequency $\nu_0 = 8$ Hz. Hence, from this fact and on account of the distance from the surface of the Earth to the ionosphere and from the conditions of formation of standing waves, we obtain that the photon mass $m_\gamma < 10^{-48}$ kg. This estimate shows that Coulomb's law is observed to an extremely high degree of accuracy, since the inequality $|\alpha| \leqslant 10^{-16}$ is equivalent to $m_\gamma \leqslant 10^{-50}$ kg.

Experiments on the investigation of the magnetic field in Earth's atmosphere have been carried out with the help of satellites. These experiments made it

www.EngineeringBooksPDF.com

possible to determine the accuracy with which Coulomb's law is satisfied at large distances. It was established that *Coulomb's law is valid to a very high degree of accuracy up to distances of the order of* 10^7 m. *Undoubtedly, Coulomb's law is satisfied for larger distances as well, although there is no direct experimental evidence for it so far.*

Verification of Coulomb's law for small distances. The validity of Coulomb's law for small distances is verified in experiments on interaction between elementary particles. Even Rutherford's experiments have led to the conclusion that Coulomb's law is valid to a high degree of accuracy down to distances of the order of 10^{-15} m. *Subsequent experiments on elastic scattering of electrons having energy of the order of several billion electron-volts have shown that Coulomb's law is valid down to distances of the order of* 10^{-17} m.

These experiments are interpreted with the help of quantum electrodynamics.

Field form of Coulomb's law. Before Faraday's experiments, Coulomb's law was interpreted as a **long-range interaction**, i.e. *it was assumed that one body acts on another as if without intermediaries. For this reason, this phenomenon was termed a long-range interaction.* Another point of view on the mechanism of interaction appeared in the second half of the 19th century, according to which bodies interact only due to a continuous "transfer of forces" in the space between them. This phenomenon was called a **short-range interaction.** It was introduced by Faraday (1791-1867) in a number of works published between 1831 and 1855. *Together with the idea of short-range interaction, the concept of the field as an intermediary in interactions was introduced.* Initially, the role of an intermediary was assigned to the medium pervading the entire space. This medium was called the **ether**. The state of the ether was characterized by certain mechanical properties such as elasticity, tension, motion of some parts of the medium relative to others, and so on. According to this treatment, the force acting on a body is the result of interaction of the body with the medium at the point where the body is located. Thus, *the mechanism of interaction was formulated in the form of local relations.* An attempt to provide a mathematical interpretation of this mechanical mode of interaction was made by Maxwell in 1861-1862. Maxwell (1831-1879) endeavoured to represent the forces of electromagnetic interaction in the form of mechanical forces caused by stress and pressure in the ether. He then went over to a phenomenological formulation of the interaction, characterizing the state of the medium with the help of vectors **E**, **D**, **H**, and **B** without giving, however, any mechanical interpretations to these vectors. It should be noted that Maxwell did not exclude a possibility of the mechanical interpretation of phenomenological equations. In 1864, he formulated the equations for an electromagnetic field, viz. Maxwell's equations. Later on it was found that mechanical properties could not be ascribed to the ether and motion relative to it was also ruled out. *The hope of the mechanical interpretation of electromagnetic interaction was lost, but the ideas of local formulation of interaction and of the existence in space of a field through which this interaction could be realized were retained. The field is considered to be a primary concept and is characterized by the quantities that cannot be interpreted within the framework of mechanical concepts.*

www.EngineeringBooksPDF.com

This statement was formulated in the most clearcut form in 1889 by Hertz (1857-1894) who experimentally discovered electromagnetic waves and formulated the Maxwell equations for vacuum in the modern form. Obviously, the field, along with matter, exists in space and time, in the form of atoms, molecules, etc.

Consequently, the field is also a form of matter, which possesses properties like momentum and energy, characteristic of all types of matter.

Electric field. Let us denote by $\mathbf{F}_{12}$ the force exerted by charge q_1 on charge q_2 and by $\mathbf{F}_{21}$ the force exerted by charge q_2 on charge q_1. Correspondingly, $\mathbf{r}_{12}$ and $\mathbf{r}_{21}$ are the vectors drawn from the point of location of the first charge to that of the second charge and vice versa. Accordingly, Coulomb's law can be written in the form

$$\mathbf{F}_{12} = \frac{1}{4\pi\varepsilon_0} \frac{q_1}{r_{12}^2} \frac{\mathbf{r}_{12}}{r_{12}} q_2, \tag{6.5a}$$

$$\mathbf{F}_{21} = \frac{1}{4\pi\varepsilon_0} \frac{q_2}{r_{21}^2} \frac{\mathbf{r}_{21}}{r_{21}} q_1. \tag{6.5b}$$

These formulas differ in their physical content and define the forces acting on the second and first charges at the points of their location, i.e. they describe the forces at different points in space. However, *the mechanism of emergence of these forces is the same. The charges q_1 and q_2 create in the space surrounding them a field which is characterized by the strength* $\mathbf{E}$. *The field strength is a local concept and has a definite value at each point in space.* **The electric field strength** *at a point is the quantity defined as the ratio of the force with which the field acts on a positive charge placed at this point to the charge itself.* This, however, does not mean that the field at any point in space can be measured just by placing a positive charge at this point and measuring the force acting on it.

Frequently, imparting a charge to a given point entails a sharp change in the electric field at this point, and the results of measurement appear to be considerably distorted (see Sec. 7).

Taking this into account, we can represent formulas (6.5) in the form

$$\mathbf{E}_2 = \frac{1}{4\pi\varepsilon_0} \frac{q_1}{r_{12}^2} \frac{\mathbf{r}_{12}}{r_{12}}, \tag{6.6a}$$

$$\mathbf{F}_{12} = \mathbf{F}_2 = q_2\mathbf{E}_2, \tag{6.6b}$$

$$\mathbf{E}_1 = \frac{1}{4\pi\varepsilon_0} \frac{q_2}{r_{21}^2} \frac{\mathbf{r}_{21}}{r_{21}}, \tag{6.7a}$$

$$\mathbf{F}_{21} = \mathbf{F}_1 = q_1\mathbf{E}_1. \tag{6.7b}$$

Formula (6.6a) describes the strength of the electric field created by the point charge q_1, while formula (6.6b) characterizes the force with which a field of strength $\mathbf{E}_2$ acts on a charge located at a point in this field. Formulas (6.7a) and (6.7b) have a similar meaning.

Thus, the action of one charge on another can be divided into two stages.

1. A point charge q creates in the space surrounding it an electric field

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \frac{q}{r^2} \frac{\mathbf{r}}{r}, \tag{6.8}$$

www.EngineeringBooksPDF.com

where r is the radius vector drawn from the point of location of the charge to the point where the field strength is measured (Fig. 19).

2. The point charge q located at a point where the field has a strength E is subjected to a force

$$\boxed{\mathbf{F} = q\mathbf{E},} \qquad (6.9)$$

exerted by this field.

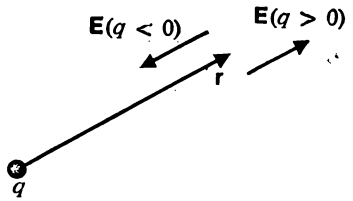


Fig. 19. Field form of Coulomb's law.

The statement (6.9) *for the second stage of interaction is of a local nature*: field strength E, charge q, and force F are determined at the same point. *The statement* (6.8) *for the first stage of interaction, however, is not local*: field strength E on the left-hand side depends not only on the point where it is determined, but also on the point where the field source is located. In other words, Eq. (6.8) is a relation between quantities pertaining to different points in space, and thus is of a nonlocal nature. The local formulation of this interaction will be given in Sec. 13.

On the limits of applicability of the classical concept of field. It was assumed above that the field strength E varies continuously and quite smoothly in space and time. However, in accordance with quantum concepts, the force of interaction between charged bodies appears as a result of an exchange of photons. *This leads to the discreteness of interaction*, and hence the field E cannot be represented as a continuous quantity which varies smoothly in space and time. It can be asked: under what conditions can the field be still treated as a continuously varying quantity? Clearly, this is possible only if the action of individual quanta is weak in comparison to their combined action, i.e. if the phenomenon under consideration is determined by the simultaneous action of a huge number of quanta. Such a situation is encountered most frequently. For example, the photon flux of visible light from a 200 W electric lamp at a distance of 2 m is about 10^{15} photons/(cm^2·s). The pupil of the eye has an area much smaller than 1 cm^2, and yet the number of photons impinging per second at the eye is quite large. Hence the photon flux is perceived as continuous. However, by decreasing the intensity of light we can arrive at a situation when a small number of photons is incident on the eye per second. Under special conditions, a human eye is capable of perceiving photons as separate flashes. In this case, we cannot use the concept of continuous light flux. Short-wave transmitters in the USSR operate at frequencies of 60-70 MHz. The electromagnetic flux from a 200 W transmitter of this type at a distance of 10 km is about 4×10^{14} quanta/(cm^2·s), which corresponds to a density of 10^4 quanta/cm^3. Consequently, more than 10^{11} radiation quanta are present in a volume equal to the cube of the wavelength ($\sim$64 m^3). Under such conditions, the detection of the field of a single quantum is quite difficult. Classical description can be used in cases where the action of individual quanta is not manifested. This is possible when the number of quanta is large and the momentum of an individual quantum is small in comparison with the momentum of a material system. For example, the radiation from an individual atom cannot be analyzed classically, since there are no photons before the emission, while only one photon is present after the emission.

www.EngineeringBooksPDF.com

If Coulomb's law is strictly satisfied, the charge of a conducting sphere is distributed over its surface. If this law is violated, a charge also exists inside the sphere.

The concept of a classical continuous interaction is valid only when the effect of individual quanta is small in comparison with their collective effect, i.e. when the phenomenon under consideration depends on the simultaneous action of a huge number of quanta and the action of individual quanta is not manifested.
The determination of electric field strength is not connected with the smallness of test charges.

What physical principle lies behind the Cavendish method of verification of Coulomb's law? What is the accuracy with which Coulomb's law can be verified by the Cavendish method with the help of modern facilities? For what distances are the verifications valid?
What is the essence of the method for verifying Coulomb's law for large distances? Up to what distances are the results of direct verification available? What are these results?
How is the validity of Coulomb's law verified for very small distances? What are the results of these verifications?
What is the difference between the concept of an electromagnetic field and that of ether?

Sec. 7. Superposition Principle

The physical meaning of the superposition principle is analyzed and the limits of its applicability are discussed.

Superposition principle for interaction of point charges. The forces of interaction between two isolated point charges are defined by Coulomb's law (6.1). Does this force change if a third point charge is brought in the vicinity of two interacting point charges? In order to give a unique meaning to this question, we must specify the forces of interaction between two point charges in the presence of a third charge (all charges are assumed fixed).

If by the forces of interaction we mean the forces directed along the line connecting the two interacting charges, these forces will depend on the third charge and will not fulfill the requirement that action and reaction be equal. The difficulty stems from the fact that we can measure the force acting on a charge, but we cannot distinguish between the forces due to separate charges. However, the third charge does not differ in any way from the two point charges under consideration and all three charges are equivalent. Hence the question can be formulated in a different way. Suppose that we have three interacting point charges. The experimentally measurable quantities are the forces acting on each of these charges. We know the law of addition of forces according to the parallelogram rule. It can be asked whether the measured force acting on each charge is equal to the sum of the forces exterted by the other two charges, if these forces are calculated in accordance with the Coulomb law (6.1). It should be noted that we are speaking of experimental measurement of a force and the mathematical calculation of forces in accordance with the law (6.1), as well as their addition

www.EngineeringBooksPDF.com

in accordance with the parallelogram rule. Under such a formulation, the question has a quite definite meaning and the answer can be obtained experimentally. Investigations have shown that *the force being measured is always equal to the sum of forces exerted by the two separate charges and calculated in accordance with the Coulomb law*. This experimental result can be expressed in the form of the following statements:

(a) *the force of interaction between two point charges does not change in the presence of other charges*;

(b) *the force exerted on a point charge by two point charges is equal to the sum of forces exerted by each point charge separately in the absence of the other*.

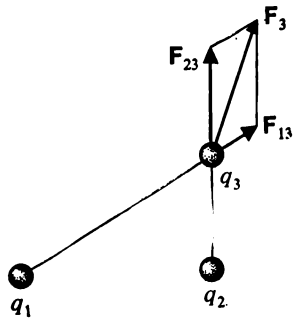


Fig. 20. Principle of superposition.

This statement is called the **superposition principle**, and reflects an experimental fact which is one of the fundamental principles of electricity. This principle is as significant in electricity as, say, the Coulomb law. Obviously, this rule can be generalized to the case of many charges.

Field form of the superposition principle. Let us consider a force $\mathbf{F}_3$ acting on a point charge q_3 in the presence of two other point charges q_1 and q_2 (Fig. 20). We denote the forces exerted on charge q_3 by charges q_1 and q_2 by $\mathbf{F}_{13}$ and $\mathbf{F}_{23}$ respectively, assuming q_3 to be the only charge present. In accordance with the superposition principle, we have

$$\mathbf{F}_3 = \mathbf{F}_{13} + \mathbf{F}_{23}. \tag{7.1}$$

Let us denote by $\mathbf{E}_{13}$ and $\mathbf{E}_{23}$ the strengths of the field created by charges q_1 and q_2 at the point charge q_3 assuming it to be the only charge present. According to formula (6.9), we get

$$\mathbf{F}_{13} = q_3\mathbf{E}_{13}, \quad \mathbf{F}_{23} = q_3\mathbf{E}_{23}. \tag{7.2}$$

Expression (7.1) can be written in the form

$$\mathbf{F}_3 = q_3\mathbf{E}_{13} + q_3\mathbf{E}_{23}. \tag{7.3}$$

The force in an electric field appears as a result of the action of the field on a charge. Consequently, force $\mathbf{F}_3$ in (7.3) shows that an electric field of strength $\mathbf{E}_3$ is set up at the point of location of charge q_3 [see (6.9)], i.e.

$$\mathbf{F}_3 = q_3\mathbf{E}_3. \tag{7.4}$$

Substituting (7.4) into (7.3) and cancelling out the common factor q_3, we get

$$\mathbf{E}_3 = \mathbf{E}_{13} + \mathbf{E}_{23}. \tag{7.5}$$

This expression is the field form of the superposition principle: *the field strength of two point charges is equal to the sum of the field strengths due to each separate charge if it acted alone*. This formulation is of a local nature, since all the quantities involved are referred to one point in space.

www.EngineeringBooksPDF.com

The generalization of this principle to the case of many charges is obvious:

$$\boxed{\mathbf{E} = \sum \mathbf{E}_i,} \tag{7.6}$$

i.e. *the field strength of any number of point charges is equal to the sum of field strengths due to each separate charge if it acted alone.*

Test charges. It follows from the definition of the electric field that its measurement is reduced to the measurement of the force acting on a point charge. The point charge used for measuring the strength of an electric field is called a **test charge.** The magnitude of a test charge remains to be specified. *If we assume that all point charges whose total field strength is calculated are fixed at certain points in space the test charge can be of any magnitude.* If, however, the positions of the point charges are not fixed in space, the action of the test charge on these charges may lead to their displacement in space. In this case, instead of the field at the point of location of the test charge for the initial position of all charges, we shall obtain the field created due to the displacement of charges to their new positions under the influence of the test charge. In order to avoid this, we must decrease the effect of the test charge on the charges creating the field under investigation. Hence the *test charge must be quite small.* It should, however, be noted that *this requirement has nothing to do with the principle of superposition, but simply ensures the conditions under which the strength of the field under consideration does not alter due to measurement.*

Limits of applicability of the principle of superposition. The agreement between the results obtained by applying the principle of superposition and the experimental results serves as an experimental verification of the principle. It has been established that *the principle of superposition is valid up to very high fields.* Its accuracy is verified in engineering for field strengths up to several million volts per metre (electrical engineering, particle accelerators, high-voltage discharge, etc.). Atoms and nuclei possess stronger fields. The strength of fields on electron orbits of atoms attains values $E \cong 10^{11}$-10^{17} V/m. The differences in atomic energy levels calculated by using the superposition principle are confirmed experimentally to a very high degree of accuracy (the relative error is not higher than 10^{-6}). This means that the principle of superposition is also obeyed to a high degree of accuracy for intra-atomic field strengths. Extremely high field strengths are attained ($E \simeq 10^{22}$ V/m) at the surfaces of heavy nuclei. Experimental results indicate that the principle of superposition is obeyed even for such strong fields. However, this is accompanied by the appearance of some other effects; to be more precise, the creation of electron-positron pairs leads to a polarization of vacuum at field strengths of about 10^{20} V/m. This brings about a nonlinearity in quantum-mechanical interactions.

The force of interaction between two point charges does not vary in the presence of other charges but, generally speaking, the force of interaction between charged bodies varies in the presence of other charged bodies. A test charge is assumed to be quite small. This requirement, however, has nothing to do with the principle of superposition which remains valid for all values of the test charge.

www.EngineeringBooksPDF.com

Why does the force of interaction between two charged bodies generally vary in the presence of a third charged body? Is it a violation of the superposition principle? What experimental facts help judge about the validity of the principle of superposition for very strong electric fields?

Sec. 8. Magnetic Field

The relativistic nature of magnetic field is analyzed. The law of interaction between parallel conductors is derived from Coulomb's law with the help of relativistic transformations.

Inevitability of magnetic field generation due to motion of charges. The interaction between fixed point charges is defined completely by Coulomb's law. This law, however, *is incapable of describing the interaction between moving charges.* Such a conclusion is based on relativistic properties of space and time and the relativistic equation of motion rather than on the specific features of Coulomb's interaction.

This statement stems, in principle, from the following considerations. The relativistic equation of motion

$$\mathrm{d}\mathbf{p}/\mathrm{d}t = \mathbf{F} \quad (8.1)$$

is invariant and has the same form in all inertial coordinate systems, including the system K' which moves uniformly and rectilinearly relative to the system K:

$$\mathrm{d}\mathbf{p}'/\mathrm{d}t' = \mathbf{F}'. \quad (8.2)$$

The quantities with primes pertain to system K'. The left-hand sides of these equations contain purely mechanical quantities whose behaviour under a transformation from one coordinate system to another is known. Consequently, the left-hand sides of Eqs. (8.1) and (8.2) can be interrelated through a certain formula. But this means that the forces on the right-hand sides of these equations are also interrelated. Such a relation stems from the requirement of relativistic invariance of the equation of motion. Since the left-hand sides of Eqs. (8.1) and (8.2) contain velocities, we can conclude that the force of interaction between moving charges is velocity-dependent and cannot be reduced to a Coulomb force. This proves that *the interaction between moving charges is due not only to a Coulomb force but also to another kind of force called the magnetic force.* The existence of such a force can be revealed by considering the following example of charge interaction.

Interaction between a point charge and an infinitely long charged filament. The Coulomb interaction between two point charges which are at rest in the coordinate system K' is, of course, the simplest type of interaction. In another coordinate system K moving relative to K', however, these charges move with the same velocity, and their interaction becomes more complicated since the

www.EngineeringBooksPDF.com

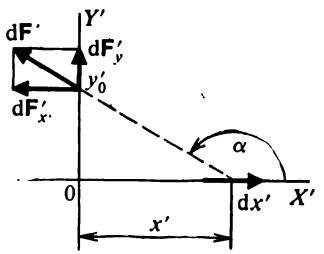


Fig. 21. On the calculation of force of interaction between a point charge and an infinitely long charged filament.

electric field at each point in space varies due to the motion of charges. Hence it is expedient to choose a situation which is quite simple for the coordinate system K', where the charges are at rest, and for the system K, where the charges are in motion. A relatively simple interaction is that between a point charge and an infinitely long charged filament.

In the system K' the filament is at rest and directed along the X'-axis (Fig. 21). A point charge q is located on the Y'-axis at a distance y_0' from the filament. We denote the cross-sectional area of the filament by S_0' and assume that its linear dimensions are very small in comparison with the distance from the point charge. If ρ' is the volume charge density, a charge $dq' = \rho' S_0' \, dx'$ exists on the element dx' of the filament. For the sake of definiteness, we assume that the charge on the filament as well as the point charge is positive. In this case, the forces exerted on the point charge by the charge on the element of the filament are directed as shown in Fig. 21. In accordance with Coulomb's law, we have

$$dF_x' = \frac{q\rho' S_0' \, dx'}{4\pi\varepsilon_0 (y_0'^2 + x'^2)} \cos\alpha, \qquad (8.3)$$

$$dF_y' = \frac{q\rho' S_0' \, dx'}{4\pi\varepsilon_0 (y_0'^2 + x'^2)} \sin\alpha.$$

Considering that $\cos\alpha = -x'/(y_0'^2 + x'^2)^{1/2}$, $\sin\alpha = y_0'/(y_0'^2 + x'^2)^{1/2}$, we obtain the following expressions for the projections of the force:

$$F_x' = -\frac{q\rho' S_0'}{4\pi\varepsilon_0} \int_{-\infty}^{\infty} \frac{x' \, dx'}{(y_0'^2 + x'^2)^{3/2}}, \qquad (8.4)$$

$$F_y' = \frac{q\rho' S_0' y_0'}{4\pi\varepsilon_0} \int_{-\infty}^{\infty} \frac{dx'}{(y_0'^2 + x'^2)^{3/2}}.$$

The first integral is equal to zero since the integrand is an odd function, but to calculate the second integral, it is expedient to carry out the following substitution of variables: $x' = -y_0' \cot\alpha$, $dx' = y_0' \, d\alpha/\sin^2\alpha$, $1 + \cot^2\alpha = 1/\sin^2\alpha$. This gives

$$F_x' = 0, \quad F_y' = \frac{q\rho' S_0'}{4\pi\varepsilon_0 y_0'} \int_0^{\pi} \sin\alpha \, d\alpha = \frac{q\rho' S_0'}{2\pi\varepsilon_0 y_0'}. \qquad (8.5)$$

Besides, $F_z' = 0$. Considering that the charge is at rest at a given instant and denoting the mass of a charge carrier by m_0, we obtain the following expression for the acceleration of charge in the system K':

$$a_x' = 0, \quad a_y' = F_y'/m_0 = q\rho' S_0'/(2\pi\varepsilon_0 y_0' m_0), \quad a_z' = 0. \qquad (8.6)$$

www.EngineeringBooksPDF.com

Let us now consider this interaction in the coordinate system K moving relative to K' towards negative values of the X'-axis with velocity v. We direct the X-axis along the filament so that its positive direction coincides with the positive direction of the X'-axis and assume that this system is fixed. In the coordinate system K, the filament, the charge, and the system K' move with velocity v in the direction of positive values of the X-axis.

Let us calculate the repulsive Coulomb force exerted by the moving filament on the moving charge. The point charge q is constant in view of the charge invariance. Due to a reduction of moving scales, the number of charges per metre length of the moving filament is larger than that of the fixed filament. In other words, the charge density of the moving filament is higher than that of the fixed filament. The charge density of the fixed filament was denoted by ρ' in the above calculations. Hence, the charge density of the moving filament in the coordinate system K is

$$\rho = \rho' / \sqrt{1 - v^2/c^2}, \tag{8.7}$$

where $\sqrt{1 - v^2/c^2}$ takes into account the relativistic variation of moving scales. All further calculations are exactly the same as for the fixed filament. Since the dimensions in a direction perpendicular to the velocity v remain unchanged, the cross-sectional area of the moving filament and the distance between the filament and the point charge remain unchanged. Hence, we obtain instead of (8.5)

$$f_x = 0, \quad f_y = q\rho S_0/(2\pi\varepsilon_0 y_0), \quad f_z = 0. \tag{8.8}$$

Here, the Coulomb force is denoted by a small letter in order to distinguish it from the total force which acts on the charge and cannot be reduced to a Coulomb force. Substituting (8.7) into the second of Eqs. (8.8), we obtain

$$f_y = q\rho' S_0/(2\pi\varepsilon_0 y_0 \sqrt{1 - v^2/c^2}) = q\rho' S_0/(2\pi\varepsilon_0 y_0' \sqrt{1 - v^2/c^2}) = F_y'/\sqrt{1 - v^2/c^2}, \tag{8.9}$$

where $S_0 = S_0'$, $y_0 = y_0'$, and formula (8.5) is taken into account.

Let us find the total force acting on a point charge in the coordinate system K. Due to symmetry, the force is directed along the Y-axis and is connected with the momentum through the following equation of motion:

$$F_y = \mathrm{d}p_y/\mathrm{d}t. \tag{8.10}$$

In the coordinate system K', this relation has the form

$$F_y' = \mathrm{d}p_y'/\mathrm{d}t'. \tag{8.11}$$

Using the transformation formulas of the theory of relativity, we get

$$p_y' = p_y, \quad \frac{\mathrm{d}t'}{\mathrm{d}t} = \frac{\sqrt{1 - \beta^2}}{1 + v u_x'/c^2} \quad (\beta = v/c), \tag{8.12}$$

where u_x' is the component of the particle velocity in the coordinate system K'. In our case, $u_x' = 0$. Taking (8.12) into account, we obtain from (8.10)

$$F_y = \mathrm{d}p_y/\mathrm{d}t = (\mathrm{d}p_y'/\mathrm{d}t')(\mathrm{d}t'/\mathrm{d}t) = F_y' \sqrt{1 - \beta^2}. \tag{8.13}$$

www.EngineeringBooksPDF.com

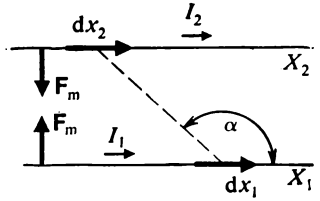


Fig. 22. Interaction between two parallel currents.

A comparison of (8.13) with (8.9) shows that

$$F_y = (1 - \beta^2) f_y, \tag{8.14}$$

i.e. *the Coulomb repulsive force f_y is larger than the force F_y exerted on the moving charge by the moving filament. Consequently, besides the Coulomb repulsive force, a non-Coulomb force, which in the present case is attractive, also acts on the charge.* This force appears as a result of the motion of charges and is called **the magnetic force**. The field form of interaction for the magnetic force is similar to that for the electric interaction: *a moving charge creates a magnetic field in the surrounding space. The magnetic field exerts a force on the moving charge.*

Relativistic nature of magnetic field. Formula (8.14) shows that the magnetic force is

$$F_{ym} = F_y - f_y = -v^2 f_y / c^2. \tag{8.15}$$

The 'minus' sign indicates that the force is directed towards the charged filament, i.e. is attractive in nature. It can be seen from this formula that this force is described by a quantity of the second order of smallness in v/c relative to the Coulomb interaction. Consequently, *the magnetic interaction is comparable to the electric interaction only for quite large velocities of particles. Nevertheless, it is also perceptible for small charge velocities if Coulomb interaction does not manifest itself for some reason or other.* Such a situation may arise, for example, when an electric current flows in a conductor. In this case, the electric field of moving charges is neutralized by the electric field of the opposite charges of the conductor, i.e. it is screened. As a result, only the magnetic force remains, which is negligibly small in comparison with the Coulomb force if the latter were not screened. For instance, for typical electron drift velocities in a metallic conductor (see Sec. 31), the magnetic force is less than Coulomb's force by a factor of 10^{20}, but nevertheless it is sufficiently large and is manifested in the form of interaction between current-carrying conductors. Therefore, a *purely relativistic effect of the emergence of magnetic field is exhibited not only at very high velocities but at any velocity.*

Forces of interaction between parallel current-carrying conductors. Suppose that charges move in a thin cylindrical wire which is electrically neutral as a whole. Then Coulomb's forces exerted by moving charges which create an electric current are screened by the opposite charges of the wire, and only a magnetic force (8.15) is acting outside the wire. Consequently, in the space surrounding the current-carrying conductor only the magnetic force acts on moving charges that generate an electric current. This leads to a magnetic interaction between currents. This result was obtained from the relativistic analysis of interaction of moving charges. However, the magnetic interaction between currents was discovered long before the theory of relativity was developed.

Let us assume that moving charges create a linear current in a conductor parallel to the initial current flowing along the X-axis and arranged at a distance r from it (Fig. 22). We shall use subscripts 1 for the quantities pertaining

www.EngineeringBooksPDF.com

to the initial current and 2 for those pertaining to the linear current. Each charge of current I_2 is acted upon by the magnetic attractive force F_m (8.15) due to current I_1. Taking into account (8.8), it is convenient to represent this force in the form

$$F_{my} = -\frac{v^2}{c^2}\frac{q\rho_1 S_{01}}{2\pi\varepsilon_0 r} = -\frac{1}{2\pi\varepsilon_0 c^2}qv\frac{\rho_1 v S_{01}}{r} = -\frac{1}{2\pi\varepsilon_0 c^2}qv\frac{I_1}{r}, \quad (8.16)$$

where $\rho_1 v S_{01} = I_1$ [see (4.11) and (4.14)] and $r = y_0$ [see (8.8)].

Let us denote the linear charge concentration on the second conductor by n_2. The element of length dx_2 contains $n_2 dx_2$ charges acted upon by the magnetic force

$$dF_m = F_{my} n_2 dx_2. \quad (8.17)$$

Substituting (8.16) into this expression, we find

$$dF_m = -\frac{1}{2\pi\varepsilon_0 c^2}\frac{I_1 q v n_2\, dx_2}{r}, \quad (8.18)$$

where $qvn_2 = I_2$. Besides, in the theory of magnetism, the magnetic constant $\mu_0 = 1/(\varepsilon_0 c^2)$ is commonly used instead of the constant ε_0. Then we obtain [see (8.18)]

$$dF_m = -\frac{\mu_0}{2\pi}\frac{I_1 I_2}{r}dx_2. \quad (8.19)$$

This force characterizes the interaction between linear currents in infinitely long parallel conductors. It should be noted that formula (8.19) can only be applied when the cross sections of the conductors are small in comparison with the distance between them (thin conductors and linear currents).

Unit of current. It immediately follows from formula (8.19) that the force acting per unit length of a conductor is

$$F_{ml} = -\frac{\mu_0}{2\pi}\frac{I_1 I_2}{r}l_2. \quad (8.20)$$

The 'minus' sign indicates that when currents I_1 and I_2 have the same direction, the force acting between the conductors is attractive. If, however, the currents I_1 and I_2 have opposite directions, a repulsive force appears.

Equation (8.20) can be used for defining the unit of current: *the ampere* (A) *is the constant current which, if maintained in two very long parallel conductors of negligibly small cross-section one metre apart in vacuum, produces a force of* 2×10^{-7} N *on each metre of the conductors.* Putting $I_1 = I_2 = 1$ A, $r = 1$ m, $l_2 = 1$ m, and $F_{ml} = -2 \times 10^{-7}$ N in (8.20), we obtain

$$\mu_0 = 4\pi \cdot 10^{-7}\ \text{N/A}^2. \quad (8.21)$$

It has been noted that [see (8.19)]

$$\mu_0 \varepsilon_0 = 1/c^2, \quad (8.22)$$

where c is the velocity of light in a vacuum. This expression reflects a relation existing between electric and magnetic fields which is characterized by the

www.EngineeringBooksPDF.com

fundamental physical constant c equal to the velocity of light. The nature of this relation will become clear from the analysis of electromagnetic waves (see Chap. 9).

Magnetic field. In complete analogy with the field form of Coulomb's interaction (see Sec. 6), we can represent the process of creation of force (8.18) in two stages: generation of a magnetic field by current I_1 in the space surrounding it, and the action of the magnetic field on a moving charge or current. However, the laws governing the generation of a magnetic field and the action of force turn out to be more complicated than Coulomb's law, since they depend on the mutual orientation of currents and the velocity of the charge. Besides, current I_1 flowing in a very long conductor cannot be considered as an elementary object whose interaction with a point charge could be considered as an elementary act. For this reason, we should return to the analysis of the action of forces on moving point charges and current elements.

Coulomb's law is insufficient for describing the interaction between moving charges. This conclusion is based on the relativistic properties of space and time and the relativistic equation of motion rather than on the specific features of Coulomb's interaction.
Magnetic interaction can be compared with electric interaction only at sufficiently high velocities of charged particles. If, however, Coulomb's interaction is absent due to some reasons, magnetic interaction can manifest itself at very low velocities.

Sec. 9. Lorentz Force. Ampère Force

Relativistic properties of Lorentz and Ampère forces are discussed.

Transformation of forces. It was shown in Sec. 8 on the basis of a specific example that if we assume the relativistic invariance of the equation of motion, it is possible to define the law of transformation of a force upon a transformation from one system of coordinates to another. Let us extend this method to a more general case.

As usual, the coordinate system K' moves relative to the system K with a velocity v in the positive direction of the X-axis. Let us consider the motion of a material point under the action of given forces. Let (F'_x, F'_y, F'_z) and (F_x, F_y, F_z) be the components of forces in the coordinate systems K' and K respectively. In the general case, the corresponding components of these forces in different coordinate systems are not equal. However, these components are connected through quite definite relations which ensure the invariance of equations of motion, i.e. the identity of their forms in different coordinate systems:

$$dp_x/dt = F_x, \quad dp_y/dt = F_y, \quad dp_z/dt = F_z \tag{9.1}$$

$$dp'_x/dt' = F'_x, \quad dp'_y/dt' = F'_y, \quad dp'_z/dt' = F'_z. \tag{9.2}$$

www.EngineeringBooksPDF.com

We transform the left-hand sides of these equations by using formulas from the theory of relativity for momentum, and the Lorentz transformations:

$$p_x = \frac{p'_x + (E'/c^2)\,v}{\sqrt{1-\beta^2}}, \quad p_y = p'_y, \quad p_z = p'_z, \tag{9.3}$$

where $E' = m'c^2$ is the total energy of the material point, and $\beta = v/c$. Formulas (9.1) are reduced to the form

$$F_x = \frac{dp_x}{dt} = \frac{dp_x}{dt'}\frac{dt'}{dt} = \frac{d}{dt'}\left[\frac{p'_1 + (E'/c^2)\,v}{\sqrt{1-\beta^2}}\right]\frac{dt'}{dt}$$

$$= F'_x + \frac{vu'_y/c^2}{1+vu'_x/c^2}F'_y + \frac{vu'_z/c^2}{1+vu'_x/c^2}F'_z, \tag{9.4}$$

$$F_y = \frac{dp_y}{dt} = \frac{dp'_y}{dt'}\frac{dt'}{dt} = \frac{\sqrt{1-\beta^2}}{1+vu'_x/c^2}F'_y, \tag{9.5}$$

$$F_z = \frac{dp_z}{dt} = \frac{dp'_z}{dt'}\frac{dt'}{dt} = \frac{\sqrt{1-\beta^2}}{1+vu'_x/c^2}F'_z, \tag{9.6}$$

where (u'_x, u'_y, u'_z) are the components of the velocity of the point in the system K'. The quantities F'_x, F'_y, F'_z appear on the right-hand sides of Eqs. (9.4)-(9.6) as a result of the application of the equation of motion (9.2). While calculating Eq. (9.4), we have taken into consideration the formula

$$\frac{dE'}{dt'} = \mathbf{F}' \cdot \mathbf{u}', \tag{9.7}$$

which expresses the law of conservation of energy in the system K'. With the help of the formulas for summation of velocities

$$u_y = \frac{u'_y\sqrt{1-\beta^2}}{1+vu'_x/c^2}, \quad u_z = \frac{u'_z\sqrt{1-\beta^2}}{1+vu'_x/c^2}, \tag{9.8}$$

we can transform Eq. (9.4) as follows:

$$F_x = F'_x + \frac{vu_y/c^2}{\sqrt{1-\beta^2}}F'_y + \frac{vu_z/c^2}{\sqrt{1-\beta^2}}F'_z. \tag{9.9}$$

In order to simplify (9.5) and (9.6), we require an important relation which is obtained from formulas for velocity transformation. By way of an example, let us write the direct and inverse transformations of the y-components of the velocity:

$$u_y = \frac{u'_y\sqrt{1-\beta^2}}{1+vu'_x/c^2}, \quad u'_y = \frac{u_y\sqrt{1-\beta^2}}{1-vu_x/c^2}.$$

Multiplying the left- and right-hand sides of these equations termwise and dividing the equations thus obtained by the common multiplier $u_y u'_y$, we obtain

$$\left(1 + \frac{vu'_x}{c^2}\right)\left(1 - \frac{vu_x}{c^2}\right) = 1 - \beta^2. \tag{9.10}$$

www.EngineeringBooksPDF.com

With the help of this equation, we can transform formulas (9.5) and (9.6):

$$F_y = \frac{1 - vu_x/c^2}{\sqrt{1-\beta^2}} F'_y, \tag{9.11}$$

$$F_z = \frac{1 - vu_x/c^2}{\sqrt{1-\beta^2}} F'_z. \tag{9.12}$$

Thus, a force in the system of coordinates K can be expressed in terms of a force in the system K' with the help of equations (9.9), (9.11) and (9.12). The formulas for inverse transformation can be easily obtained by using the relativity principle.

In the derivation of the above formulas, no assumptions were made about the properties of the initial forces which may depend on coordinates, time and velocity. Besides, it was not assumed that in some coordinate system the particle will be at rest, since no constraints were imposed on the particle velocities. The formulas thus obtained show that *in the relativistic theory, the dependence of forces on velocity is unavoidable*: even if the force is independent of velocity in some coordinate system (say, F'_x, F'_y, F'_z), the dependence inevitably appears in other coordinate systems (in the present case, F_x, F_y, F_z depend on the particle velocity (u_x, u_y, u_z).

Let us write the formulas for the transformation of forces in vector form. For this purpose, we introduce the notation

$$\Phi = (F'_x,\ F'_y/\sqrt{1-\beta^2},\quad F'_z/\sqrt{1-\beta^2}), \tag{9.13}$$

$$\mathbf{G} = [0,\ -(v/c^2)\,F'_z/\sqrt{1-\beta^2},\quad (v/c^2)\,F'_y/\sqrt{1-\beta^2}]. \tag{9.14}$$

It can be easily verified that in terms of these quantities, formulas (9.9), (9.11) and (9.12) can be written in the form of the vector equation

$$\mathbf{F} = \Phi + \mathbf{u} \times \mathbf{G}. \tag{9.15}$$

Since $\mathbf{F}$ is a vector, the entire right-hand side is a vector. This equality is valid for any $\mathbf{u}$; consequently, each term on the right-hand side is a vector. Since $\mathbf{u} \times \mathbf{G}$ and $\mathbf{u}$ are vectors, we conclude that $\mathbf{G}$ is also a vector. This shows that the quantities Φ and $\mathbf{G}$ defined by equalities (9.13) and (9.14) are vectors.

Lorentz force. We assume that only an electric field exists in the coordinate system K', and hence the force $(F'_x,\ F'_y,\ F'_z)$ is independent of the particle velocity $\mathbf{u}'$. In this case, Φ [see (9.13)] is independent of the particle velocity $\mathbf{u}$ and represents the electric force in the coordinate system K.

Similarly, it can be concluded that the vector $\mathbf{G}$ is also independent of the particle velocity $\mathbf{u}$ and can depend only on coordinates and time. Consequently, the dependence of force on particle velocity is expressed by the second term of (9.15):

$$\mathbf{F}_m = \mathbf{u} \times \mathbf{G}. \tag{9.16}$$

This is a magnetic force directed perpendicular to the particle velocity and vector $\mathbf{G}$ which represents the magnetic field acting on the moving particle.

www.EngineeringBooksPDF.com

Since Φ in formula (9.15) is the electric force acting on a point charge q, we can write the electric field strength in the form

$$\mathbf{E} = \Phi/q. \tag{9.17}$$

Similarly, the magnetic induction is given by the formula

$$\mathbf{B} = \mathbf{G}/q. \tag{9.18}$$

Taking these formulas into account, we can rewrite formula (9.15) for the force acting on a point charge in the following form:

$$\boxed{\mathbf{F} = q\mathbf{E} + q\mathbf{u} \times \mathbf{B}.} \tag{9.19}$$

This is the **Lorentz force.** *The first term on the right-hand side characterizes the force exerted on a point charge by an electric field, while the second term expresses the force exerted by a magnetic field.*

Magnetic induction. Since the force exerted on a moving charge by a magnetic field is denoted by **B**, it is but natural to term this vector as the magnetic field strength. However, historically the term magnetic field strength was used to describe another vector which is denoted by **H**. This vector is not a property of the magnetic field, but just takes into account the properties of the material medium in which the field exists. In particular, for a given **H**, the vector **B**, and hence the force acting on a moving charge, can have quite different values (see Sec. 38). The vector **B** is called the **magnetic induction.**

Ampère force. Suppose that we have an aggregate of point charges with a concentration n. Then there will be $n\,dV$ charges in a volume element dV. If the velocity of each of these charges is $\mathbf{u}$, and if each charge is acted by a magnetic force defined by the second term in (9.19), the force acting on the charges in the volume element dV will be

$$d\mathbf{F}_m = nq\,dV\mathbf{u} \times \mathbf{B}. \tag{9.20}$$

Henceforth, we shall omit the subscript "m" on the force, which only indicates the magnetic nature of the force. The force exerts the same action on a charge irrespective of its origin. Considering that

$$nq = \rho, \quad nq\mathbf{u} = \rho\mathbf{u} = \mathbf{j}, \tag{9.21}$$

where ρ and $\mathbf{j}$ represent the charge and current densities [see (4.4) and (4.11)], we can write formula (9.20) in the form

$$d\mathbf{F} = \rho\mathbf{u} \times \mathbf{B}\,dV, \tag{9.22}$$

or

$$\boxed{d\mathbf{F} = \mathbf{j} \times \mathbf{B}\,dV.} \tag{9.23}$$

This relation is called the Ampère law and defines the force acting on an element of electric current with density $\mathbf{j}$, enclosed in volume dV.

www.EngineeringBooksPDF.com

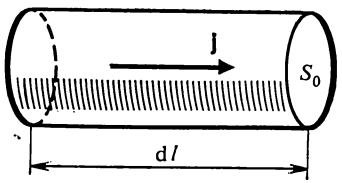


Fig. 23. Transformation from steady volume currents to linear currents: $\mathbf{j}\,dV = jS_0\,d\mathbf{l} = I\,d\mathbf{l}$.

Transformation from steady volume currents to linear currents. Formula (9.23) can be represented in another form also. Suppose that an electric current flows in a thin conductor of cross-sectional area S_0. Consider a line element dl of the conductor (Fig. 23). The volume dV of this element is $S_0\,dl$. Since the cross-sectional area of the conductor is small, we assume that the density $\mathbf{j}$ of the current flowing through it is constant, and hence

$$I = S_0 j. \tag{9.24}$$

Suppose that the direction of $d\mathbf{l}$ coincides with the density vector of the current flowing through this region of the conductor. In this case,

$$\mathbf{j}\,dV = \mathbf{j}S_0\,dl = I\,d\mathbf{l}. \tag{9.25}$$

Generally speaking, the electric current flowing through each point in space has a different density and is therefore called the volume current. The force acting on such a current in the volume element dV is defined by formula (9.23). If, however, the current passes through thin conductors (having infinitely small thickness in the limit, in the physical sense), it is called linear current. In this case, we can speak of a current element on the length $d\mathbf{l}$ of the conductor. A transformation from formulas derived for volume currents to formulas for linear currents is accomplished through relation (9.25) which can be represented in the form

$$\boxed{\mathbf{j}\,dV \rightleftarrows I\,d\mathbf{l}.} \tag{9.26}$$

The arrows indicate that the transformation can be made from formulas for volume currents to formulas for linear currents, and vice versa.

In particular, formula (9.23) for linear currents assumes the form

$$\boxed{d\mathbf{F} = I\,d\mathbf{l} \times \mathbf{B}.} \tag{9.27}$$

This formula reflects the basic idea of Ampère, viz. to reduce the interaction between current circuits to the interaction between very small current elements.

Magnetic field of a rectilinear current. Comparing formulas (9.27) and (8.19), we conclude that *the current flowing through a long straight conductor generates a magnetic field whose lines of force are concentric circles around the current and lie in a plane perpendicular to it.* The magnetic induction at a distance r from the centre of a current-carrying conductor is given by the formula

$$B = \frac{\mu_0}{2\pi}\frac{I}{r}. \tag{9.28}$$

This formula is obtained from Coulomb's law with the help of the theory of relativity and by taking into account the principle of superposition for the electric field strength and the invariance of charge. *From the superposition principle*

www.EngineeringBooksPDF.com

for the electric field strength, we can conclude that the superposition principle is also valid for magnetic induction.

Formula (9.28) can be expressed in vector form as follows. We direct the X-axis of the Cartesian system of coordinates along the line current. Denoting by

$$\mathbf{r} = \mathbf{i}_x (x - x') + \mathbf{i}_y y + \mathbf{i}_z z$$

the radius vector directed from point $(x', 0, 0)$ to the point (x, y, z), and taking into account the value of the integral in the expression (8.4) for F'_y, i.e.

$$\int_{-\infty}^{\infty} \frac{dx'}{(y_0'^2+x_0'^2)^{3/2}} = \int_{-\infty}^{\infty} \frac{dx'}{[y_0'^2+(x-x')^2]^{3/2}} = \frac{2}{y_0'^2} \tag{9.29}$$

we can express (9.28) in the following form:

$$\mathbf{B} = \frac{\mu_0 I}{4r} \int_{-\infty}^{\infty} \frac{\mathbf{i}_x \times \mathbf{r}}{r^3} dx', \tag{9.30}$$

where we have considered that

$$|\mathbf{i}_x \times \mathbf{r}| = \sqrt{y^2+z^2}.$$

The transformation formulas for force are obtained from the requirement of invariance of the relativistic equation of motion.

The dependence of force on velocity is inevitable in the relativistic theory. Even if the force is independent of velocity in some coordinate system, this dependence appears in another coordinate system moving with respect to the first system.

The formulas for force transformation are obtained from the requirement of invariance of the relativistic equation of motion. Does this mean that the law of force transformation is a statement devoid of any physical meaning, and is just the tautology of the requirement of relativistic invariance?

Why are we unable to conclude directly from formulas (9.13) and (9.14) that Φ and $\mathbf{G}$ are vectors?

Sec. 10. Biot-Savart Law

The field form of the interaction between currents and the Biot-Savart law are discussed.

Interaction between current elements. The law of interaction between currents was discovered experimentally long before the theory of relativity had been created. It is much more complicated than Coulomb's law describing the interaction between fixed point charges. This explains why so many scientists took part in the investigation of this phenomenon. A considerable contribution to the

www.EngineeringBooksPDF.com

discovery of this law was made by Biot (1774-1862), Savart (1791-1841), Ampère (1775-1836), and Laplace (1749-1827).

In 1820, Oersted (1777-1851) discovered the action of an electric current on a magnetic needle. In the same year, Biot and Savart formulated this law for the force dF with which the current element $I\,dl$ acts on a magnetic pole at a distance r from this element:

$$dF \sim I\,dl\varphi\,(\alpha)\,f\,(r). \tag{10.1}$$

where α is the angle characterizing the mutual orientation of the current element and the magnetic pole. The function $\varphi\,(\alpha)$ was soon found experimentally. The function $f\,(r)$ was theoretically determined by Laplace in the form

$$f\,(r) \sim 1/r^2. \tag{10.2}$$

Thus, the efforts of Biot, Savart, and Laplace led to the formula describing the action of current on a magnetic pole. In the final form, the Biot-Savart-Laplace law was formulated in 1826 as an expression for the force acting on a magnetic pole, since the concept of field strength has not yet been introduced.

In 1820, Ampère discovered the interaction (attraction or repulsion) between parallel currents. He proved the equivalence of a solenoid and a permanent magnet. This made it possible to clearly formulate the problem, viz. to reduce all magnetic interactions to the interaction between current elements and to find the law of their interaction as a fundamental law which plays the same role in magnetism as Coulomb's law in electricity. By his education and inclinations, Ampère was a theoretician and mathematician. Nevertheless, while investigating the interaction between current elements, he fulfilled scrupulous experimental work and constructed a number of intricate devices. Ampère's bench for demonstrating the forces of interaction between current elements and their dependence on angles is still used as a teaching aid. As a result, Ampère discovered the law of interaction between current elements. Unfortunately, the way that has led him to this discovery is reflected neither in publications nor in his notes. However, Ampère's formula for the force differs from (10.3) in that it contains a total differential on the right-hand side. This difference is insignificant when the force of interaction between current loops is calculated, since the integral of the total differential around a closed path is equal to zero. Considering that the force of interaction between current loops is measured in experiments rather than the force of interaction between current elements, we have all grounds to assume that Ampère was the author of the law of magnetic interaction of currents. In its present form, the law of interaction between current elements was obtained in 1844 by Grassman (1809-1877). In modern notation, this law is written as follows:

$$\boxed{d\mathbf{F}_{12} = \frac{\mu_0}{4\pi} \frac{I_2\,d\mathbf{l}_2 \times (I_1\,d\mathbf{l}_1 \times \mathbf{r}_{12})}{r_{12}^3},} \tag{10.3}$$

where $d\mathbf{F}_{12}$ is the force with which the current element $I_1 d\mathbf{l}_1$ acts on the current element $I_2 d\mathbf{l}_2$, and $\mathbf{r}_{12}$ is the radius vector drawn from current element $I_1 d\mathbf{l}_1$

www.EngineeringBooksPDF.com

to $I_2 d\mathbf{l}_2$ (Fig. 24). The dashed circles in the figure denote closed loops in which the interaction of current elements is not considered.

The force $d\mathbf{F}_{21}$ with which current element $I_2 d\mathbf{l}_2$ acts on $I_1 d\mathbf{l}_1$ is naturally given by the same formula (10.3) in which subscripts 2 and 1 are interchanged:

$$d\mathbf{F}_{21} = \frac{\mu_0}{4\pi} \frac{I_1 d\mathbf{l}_1 \times (I_2 d\mathbf{l}_2 \times \mathbf{r}_{21})}{r_{21}^3}. \quad (10.4)$$

The unit vectors $\mathbf{n}_{21}$ and $\mathbf{n}_{12}$ in Fig. 24 show the directions of forces $d\mathbf{F}_{21}$ and $d\mathbf{F}_{12}$ perpendicular to the corresponding current elements. Generally, these forces are not collinear. Consequently, *the interaction of current elements does not obey Newton's third law*:

$$\boxed{d\mathbf{F}_{21} + d\mathbf{F}_{12} \neq 0.} \quad (10.5)$$

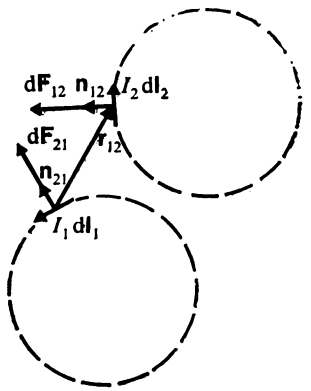


Fig. 24. Interaction between current elements.

The force with which current I_1 flowing in closed loop L_1 acts on closed loop L_2 carrying current I_2, according to (10.3), is

$$\mathbf{F}_{12} = \frac{\mu_0 I_1 I_2}{4\pi} \int_{L_1} \int_{L_2} \frac{d\mathbf{l}_2 \times (d\mathbf{l}_1 \times \mathbf{r}_{12})}{r_{12}^3}. \quad (10.6)$$

The currents I_1 and I_2 are taken out of the integral since they are constant at all points of integration paths L_1 and L_2. The formula for the force $\mathbf{F}_{21}$ acting on the closed loop carrying current I_1 has a similar form. *Newton's third law is observed for the forces of interaction between closed current loops* (see Sec. 39):

$$\boxed{\mathbf{F}_{21} + \mathbf{F}_{12} = 0.} \quad (10.7)$$

On experimental verification of the law of interaction. Strictly speaking, the *law* (10.3) *of interaction between current elements cannot be verified experimentally*, since there are no isolated current elements $I\,d\mathbf{l}$ the force of interaction between which could be measured. Each current element is a part of a closed current, loop, and hence only the law (10.6) of interaction between closed current loops can be verified experimentally. However, the validity of (10.6) does not imply the validity of (10.4), since we can add to (10.4) any function which, when substituted into (10.6) vanishes upon integration over closed path.

The electric current is due to motion of charges. Hence, formula (10.4) also expresses *the law of magnetic interaction between moving charges, which can be easily obtained from it and verified experimentally*, since the force of interaction between moving charges can be measured. The agreement between numerous corollaries of this formula and experimental results provides its most convincing verification.

www.EngineeringBooksPDF.com

Field form of interaction. In complete analogy with electrostatics, the interaction between current elements occurs in two stages: current element $I_1 d\mathbf{l}_1$ creates a magnetic field at the point of location of current element $I_2 d\mathbf{l}_2$. The interaction of element $I_2 d\mathbf{l}_2$ with this field leads to a force $d\mathbf{F}_{12}$. The action of the magnetic field with induction $\mathbf{B}$ on I dl is described by formula (9.27). Taking this formula into account, the two stages of interaction are described as follows:

(1) *current element $I_1 d\mathbf{l}_1$ creates at the point of location of current element $I_2 d\mathbf{l}_2$ a magnetic field characterized by induction*

$$\boxed{d\mathbf{B}_{12} = \frac{\mu_0}{4\pi} \frac{I_1 d\mathbf{l}_1 \times \mathbf{r}_{12}}{r_{12}^3};} \tag{10.8}$$

(2) *current element $I_2 d\mathbf{l}_2$, located at a point where the magnetic induction is $d\mathbf{B}_{12}$, is acted upon by the force*

$$\boxed{d\mathbf{F}_{12} = I_2 d\mathbf{l}_2 \times d\mathbf{B}_{12}.} \tag{10.9}$$

Biot-Savart law. Relation (10.8), describing the generation of a magnetic field by a current, is called the **Biot-Savart law**. For a closed loop with current I, we have

$$\boxed{\mathbf{B} = \frac{\mu_0}{4\pi} \oint \frac{I d\mathbf{l} \times \mathbf{r}}{r^3},} \tag{10.10}$$

where $\mathbf{r}$ is the radius vector drawn from the current element I dl to the point at which the magnetic induction $\mathbf{B}$ is being calculated. The integration in (10.10) is carried out over the closed path. The current is assumed to be linear. A transition to volume currents is accomplished in accordance with rule (9.26). For volume currents, the Biot-Savart law (10.10) assumes the form

$$\boxed{\mathbf{B} = \frac{\mu_0}{4\pi} \int_V \frac{\mathbf{j} \times \mathbf{r}}{r^3} dV.} \tag{10.11}$$

Here, the integration is carried out over all regions in space where volume currents exist and are characterized by the current density $\mathbf{j}$.

Force of interaction between rectilinear currents. The magnetic induction $d\mathbf{B}_{12}$ created by the current element $I_1 dx_1$ at the point of location of the current element $I_2 dx_2$ (Fig. 22) is directed along the outward normal to the plane of the figure. Its magnitude is

$$dB_{12} = \frac{\mu_0}{4\pi} \frac{I_1 dx_1 \sin\alpha}{r_{12}^2}. \tag{10.12}$$

www.EngineeringBooksPDF.com

Consequently, the magnetic induction of the field created by a rectilinear current I_1 flowing in an infinite conductor at the point of location of the current element $I_2 dx_2$ [see (10.10)] is expressed by the formula

$$B_{12} = \frac{\mu_0 I_1}{4\pi} \int_{-\infty}^{\infty} \frac{\sin\alpha\, dx_1}{r_{12}^2} = \frac{\mu_0}{2\pi} \frac{l}{r} \qquad (10.13)$$

where the substitution of variables for evaluating the integrals is the same as in formula (8.5).

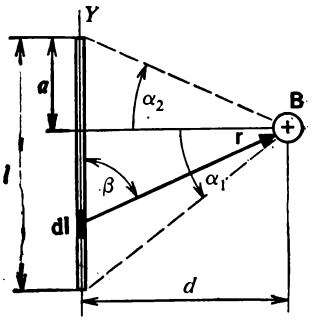


Fig. 25. Magnetic induction of a straight wire segment of finite length.

Ampère's formula leads to the conclusion that in a magnetic field with induction (10.13), the force dF_{12}, acting on the current element $I_2 dl_2$, is perpendicular to the conductor carrying current I_2 and is directed towards current I_1, i.e. is attractive:

$$dF_{12} = \frac{\mu_0}{2\pi} \frac{I_1 I_2}{r} dx_2. \qquad (10.14)$$

Formulas (10.13) and (10.14) coincide with (9.28) and (8.19) respectively.

The experimental verification of formulas for a magnetic field obtained from the corresponding formulas for an electric field through relativistic transformations not only serves as a proof of the existence of a magnetic field, but also confirms its relativistic nature.

The forces of interaction between current elements do not obey Newton's third law. The forces of interaction between closed current-carrying loops obey Newton's third law.

Current elements do not exist in an isolated form. What, then, is the idea behind a direct experimental verification of the formula for interaction between current elements?
What conclusion can be drawn from the fact that forces of interaction between current elements do not obey Newton's third law, while the forces of interaction between closed current-carrying loops obey this law?

Example 10.1. *Find the magnetic induction of the field created by a straight wire segment of length l carrying a current I* (Fig. 25).

The magnetic induction of the field created by each element of the conductor is perpendicular to the plane of the figure and, in accordance with the law (10.10), is

$$dB = \frac{\mu_0}{4\pi} I \frac{dl \times r}{r^3},$$

since $dl \times r$ is perpendicular to the plane of the figure. Then

$$|dl \times r| = dlr \sin(\widehat{dl, r}) = dlr \sin\beta = dyd,$$

and hence

$$B = \frac{\mu_0 Id}{4\pi} \int_{-(l-a)}^{a} \frac{dy}{(d^2+y^2)^{3/2}} = \frac{\mu_0 I}{4\pi d} (\sin\alpha_1 + \sin\alpha_2).$$

Using this formula, we can calculate the magnetic induction for any current loop consisting of rectilinear segments.

www.EngineeringBooksPDF.com

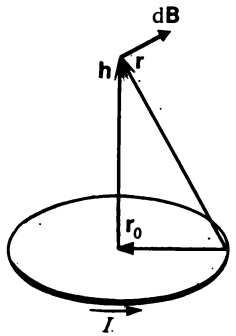


Fig. 26. Magnetic induction along the axis of a current-carrying loop.

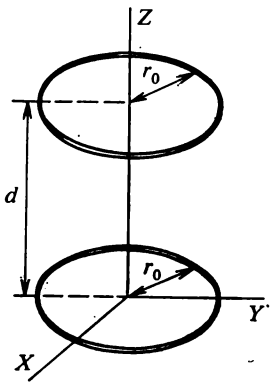


Fig. 27. To the calculation of interaction between two circular currents.

Example 10.2. *Find the magnetic induction on the axis of a circular current I of radius* r_0 (Fig. 26).

Let us use the law (10.11):

$$\mathbf{B}=\frac{\mu_0 I}{4\pi}\oint_L \frac{d\mathbf{l}\times\mathbf{r}}{r^3}.$$

where $\mathbf{r}=\mathbf{r}_0+\mathbf{h}$, $d\mathbf{l}\times\mathbf{r}=d\mathbf{l}\times\mathbf{r}_0+d\mathbf{l}\times\mathbf{h}$. The magnitude of r does not change in integration, and hence

$$\mathbf{B}=\frac{\mu_0 I}{4\pi r^3}\left(\oint_L d\mathbf{l}\times\mathbf{r}_0+\oint_L d\mathbf{l}\times\mathbf{h}\right). \tag{10.15}$$

Since h is a constant vector, we find

$$\oint_L d\mathbf{l}\times\mathbf{h}=\left(\oint_L d\mathbf{l}\right)\times\mathbf{h}=0,$$

because $\oint d\mathbf{l}=0$. The other integral in (10.15) is calculated as follows:

$$\oint_L d\mathbf{l}\times\mathbf{r}_0=\oint_L \mathbf{n}r_0\,dl=\mathbf{n}r_0\oint_L dl=\mathbf{n}r_0 2\pi r_0,$$

where n is a unit vector perpendicular to the plane containing the current loop.

Then we obtain

$$B_h=\frac{\mu_0 I}{2}\frac{r_0^2}{(r_0^2+h^2)^{3/2}}\mathbf{n}. \tag{10.16}$$

Example 10.3. *Helmholtz rings are two coaxial circular conductors of the same radius arranged in parallel planes so that the distance d between them is equal to their radius.*

Prove that the magnetic field on the axis of Helmholtz rings at the midpoint between them is uniform to a high degree of accuracy.

Let us place the origin of the Cartesian system of coordinates at the centre of one of the rings and direct the Z-axis along the axis of the rings (Fig. 27). The magnetic induction on the axis of the rings at a point with coordinate z is given, in accordance with (10.16), by

www.EngineeringBooksPDF.com

$$B_z = \frac{\mu_0 I r_0^2}{2}\left[\frac{1}{(z^2+r_0^2)^{3/2}} + \frac{1}{[(z-d)^2+r_0^2]^{3/2}}\right], \tag{10.17}$$

where I is the current in a ring.

In a first approximation, the nonuniformity of B_z is characterized by the first derivative

$$\frac{\partial B_z}{\partial z} = \frac{3\mu_0 I r_0^2}{2}\left[\frac{-z}{(z^2+r_0^2)^{5/2}} - \frac{z-d}{[(z-d)^2+r_0^2]^{5/2}}\right]. \tag{10.18}$$

For $z = d/2$, we obtain $\partial B_z/\partial z = 0$, which gives

$$\frac{\partial^2 B_z}{\partial z^2} = \frac{3\mu_0 I r_0^2}{2}\left\{\frac{5z^2}{(z^2+r_0^2)^{7/2}} - \frac{1}{(z^2+r_0^2)^{5/2}} + \frac{5(z-d)^2}{[(z-d)^2+r_0^2]^{7/2}} - \frac{1}{[(z-d)^2+r_0^2]^{5/2}}\right\}. \tag{10.19}$$

Since for Helmholtz rings $d = r_0$, for $z = d/2$ we obtain $\partial^2 B_z/\partial z^2 = 0$. This means that the field in the vicinity of the point $z = d/2$ on the axis of Helmholtz rings is actually uniform to a high degree of accuracy.

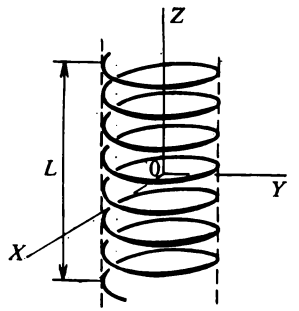


Fig. 28. A solenoid of finite length.

Example 10.4. *A straight circular solenoid of length L consists of n turns of a tightly wound thin wire. Find the magnetic induction on the solenoid axis assuming that current I flows through its turns.*

Since the turns are wound tightly, we can assume with a sufficiently high accuracy that each turn creates on the solenoid axis a field defined by formula (10.16). The number of turns per unit length is n/L. We can assume that the current over the length dz of the solenoid is $(In/L)\,dz$. Placing the origin of coordinates at the midpoint on the solenoid axis (Fig. 28), we find, using formula (10.16), that the magnetic induction at the point z on the solenoid axis is given by

$$B_z = \frac{\mu_0 n r_0^2 I}{2L}\int_{-L/2}^{L/2}\frac{dz'}{[(z-z')^2+r_0^2]^{3/2}}$$

$$= \frac{\mu_0 n I}{2L}\left\{\frac{-z+L/2}{[(z-L/2)^2+r_0^2]^{1/2}} + \frac{z+L/2}{[(z+L/2)^2+r_0^2]^{1/2}}\right\}. \tag{10.20}$$

For a very long solenoid ($L \to \infty$), for points $z \ll L/2$, we obtain from formula (10.20)

$$\lim_{L\to\infty} B_z = \mu_0 n I/L. \tag{10.21}$$

The magnetic field of a very long solenoid is not only constant along its axis but is also uniform over its cross section [see (8.38)].

Sec. 11. Field Transformation

The law of field transformation is derived proceeding from the invariance of the equation of motion of a charge in an electromagnetic field.

Invariance of the expression for force in an electromagnetic field. Expression (9.19) for the Lorentz force acting on a point charge in an electromagnetic field has been obtained from the requirement of the invariance of the relativistic

www.EngineeringBooksPDF.com

equation of motion. Consequently, *this expression must also be relativistically invariant, i.e. must have the same form in all systems of coordinates.* Thus, the expressions for forces in the coordinate systems K and K' have the form

$$\mathbf{F} = q\,(\mathbf{E} + \mathbf{u} \times \mathbf{B}), \tag{11.1}$$

$$\mathbf{F}' = q\,(\mathbf{E}' + \mathbf{u}' \times \mathbf{B}'). \tag{11.2}$$

Using the relativistic invariance of these expressions and taking into account (9.9), (9.11) and (9.12), we can obtain a relation between electric and magnetic field vectors in various coordinate systems.

A particular case of transformation of field vectors was considered earlier, when it was shown that while an electric field only exists in the coordinate system K, a magnetic field also appears in the system K. Similarly, it can be shown that if, a magnetic field only is present in one system of coordinates, the electric field also appears in another system. Let us consider the relation between electric and magnetic fields in the general case.

Transformation of fields. Substituting the expressions (11.1) and (11.2) for F_y and F'_y into formula (9.11), we obtain

$$E_y + (u_z B_x - u_x B_z) = \frac{1 - vu_x/c^2}{\sqrt{1-\beta^2}}\,[E'_y + (u'_z B'_x - u'_x B'_z)]. \tag{11.3}$$

Eliminating u'_x and u'_z from this formula by using the formulas for addition of velocities,

$$u'_x = \frac{u_x - v}{1 - vu_x/c^2}, \quad u'_z = \frac{u_z\sqrt{1-\beta^2}}{1 - vu_x/c^2}, \tag{11.4}$$

and transposing all the terms in (11.3) to the left-hand side, we get

$$\left(E_y - \frac{E'_y}{\sqrt{1-\beta^2}} - \frac{vB'_z}{\sqrt{1-\beta^2}}\right) + \left(-B_z + \frac{vE'_y}{c^2\sqrt{1-\beta^2}} + \frac{B'_z}{\sqrt{1-\beta^2}}\right) u_x + (B_x - B'_x)\,u_z = 0. \tag{11.5}$$

This equality is valid for any values of u_x and u_z. Consequently, the expressions in the parentheses in Eq. (11.5) are separately equal to zero. Equating these expressions to zero, we obtain the transformation formulas for field vectors:

$$E_y = \frac{E'_y + vB'_z}{\sqrt{1-\beta^2}}, \tag{11.6}$$

$$B_x = B'_x, \tag{11.7}$$

$$B_z = \frac{B'_z + (v/c^2)\,E'_y}{\sqrt{1-\beta^2}}. \tag{11.8}$$

The formulas for transforming other components can be obtained in a similar manner by proceeding from (9.12):

$$E_z = \frac{E'_z - vB'_y}{\sqrt{1-\beta^2}}, \tag{11.9}$$

$$B_x = B'_x, \tag{11.10}$$

$$B_y = \frac{B'_y - (v/c^2)\,E'_z}{\sqrt{1-\beta^2}}. \tag{11.11}$$

www.EngineeringBooksPDF.com

It is convenient to derive the transformation of the x-projection of a force through formula (9.4) which can be written in the form

$$F_x = \frac{1}{1+vu'_x/c^2}\left[F'_x + \frac{v}{c^2}(\mathbf{F}'\cdot\mathbf{u}')\right]. \tag{11.12}$$

Proceeding in the same way as before, we can reduce Eq. (11.12) to the form

$$\left(1+\frac{vu'_x}{c^2}\right)[E_x+(u_yB_z-u_zB_y)]-[E'_x+(u'_yB'_z-u'_zB'_y)]=\frac{v}{c^2}(\mathbf{E}'\cdot\mathbf{u}'), \tag{11.13}$$

where $\mathbf{F}' \cdot \mathbf{u}' = q\mathbf{E}' \cdot \mathbf{u}'$. With the help of formulas (11.8) and (11.11), we find that

$$E_x = E'_x. \tag{11.14}$$

Thus, the transformation formulas for the electric and magnetic field vectors have the form

$$\boxed{\begin{aligned} &E_x = E'_x, && B_x = B'_x, \\ &E_y = \frac{E'_y+vB'_z}{\sqrt{1-\beta^2}}, && B_y = \frac{B'_y-(v/c^2)E'_z}{\sqrt{1-\beta^2}}, \\ &E_z = \frac{E'_z-vB'_y}{\sqrt{1-\beta^2}}, && B_z = \frac{B'_z+(v/c^2)E'_y}{\sqrt{1-\beta^2}}. \end{aligned}} \tag{11.15}$$

The inverse transformation formulas for the field vector in accordance with the relativity principle can be obtained by replacing v by $-v$, the primed quantities by unprimed ones, and vice versa.

Application of formulas (11.15). Formulas (11.15) can be used to find the electromagnetic field vectors in any inertial system of coordinates if their value is known in any such system.

As an example, let us consider the field of a very long charged filament. The filament is fixed and lies along the X'-axis in the coordinate system K'. Consequently, this coordinate system only contains an electric field whose strength is given by formula (8.5) by taking into consideration the definition of the field strength. Hence, instead of (8.5), we obtain the following expression for the electric field strength:

$$E'_x = 0, \quad E'_y = \rho' S_0/(2\pi\varepsilon_0 y_0), \quad E'_z = 0. \tag{11.16}$$

The Y-axis may have any direction perpendicular to the filament. From formula (11.16) we find that the electric field strength of a very long charged filament is perpendicular to the filament and decreases in inverse proportion to the distance from it. Since the charges are fixed, there is no magnetic field in the coordinate system K'.

In the coordinate system K, the filament moves along its length in the positive direction of the X-axis with a velocity v. On the basis of (11.15), the electric field strength is

$$E_x = 0, \quad E_y = E'_y/\sqrt{1-\beta^2} = \rho' S'_0/(2\pi\varepsilon_0 y'_0\sqrt{1-\beta^2}), \quad E_z = 0, \tag{11.17}$$

www.EngineeringBooksPDF.com

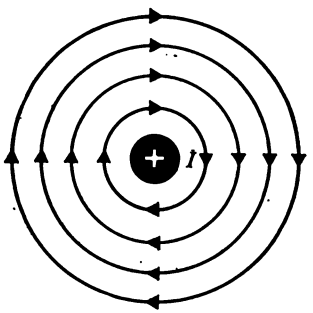


Fig. 29. Magnetic field lines of a charged filament moving along its length.

which is equivalent to (8.8), since the field strength is the ratio of force to charge.

Formulas (11.15) show that in addition to the electric field, the moving charged filament also generates a magnetic field in the space surrounding it. The induction of this field is

$$B_x = 0,\ B_y = 0,\ B_z = \frac{(v/c^2)\,E'_y}{\sqrt{1-\beta^2}} = \frac{v\rho' S'_0}{2\pi\varepsilon_0 c^2 y'_0 \sqrt{1-\beta^2}}. \qquad (11.18)$$

which is equivalent to formula (8.15) if we take into account (8.9), and go over from force f_y to the magnetic field induction in accordance with formulas (9.18) and (9.16), i.e. if we divide f_y in (8.15) by qv. Obviously, the magnetic field lines are concentric circles in planes perpendicular to the filament (Fig. 29), the centre of the circles lying on the filament.

While solving specific problems, we must choose the coordinate system in such a way that the electromagnetic field in it has the simplest form. This considerably simplifies the solution of the problem. *It should not be thought that there always exists a coordinate system in which the field can be reduced either to an electric field or to a magnetic one. There are configurations of electromagnetic fields for which electric and magnetic fields exist simultaneously in any coordinate system.* A general analysis of this question is carried out by considering the invariance of an electromagnetic field with respect to the Lorentz transformation (see Sec. 62).

Field of a point charge moving uniformly in a straight line. To begin with, we fix the origin of the coordinate system K' at the point charge q. In this system, the electric field strength is described by Coulomb's law, and the magnetic field is absent:

$$\mathbf{E}' = \frac{q}{4\pi\varepsilon_0}\frac{\mathbf{r}'}{r'^3},\quad \mathbf{B}' = 0, \qquad (11.19)$$

where $r'^2 = x'^2 + y'^2 + z'^2$. In the coordinate system K the charge q moves with a velocity v in the positive direction of the X-axis. The coordinate axes of the system K' are oriented so that they coincide with the corresponding axes of the system K at the instant $t' = t = 0$. Substituting (11.19) into (11.15) and using the Lorentz transformation, we obtain

$$E_x = E'_x = \frac{q}{4\pi\varepsilon_0}\frac{x'}{r'^3} = \frac{q\gamma\,(x-vt)}{4\pi\varepsilon_0\,[\gamma^2(x-vt)^2+y^2+z^2]^{3/2}}, \qquad (11.20)$$

where

$$\gamma = \frac{1}{(1-v^2/c^2)^{1/2}} \qquad (11.21)$$

Denoting by x_q the coordinate of the charge q in the system K at the instant t when the field is determined at the point (x, y, z), we can rewrite (11.20) in

www.EngineeringBooksPDF.com

the form

$$E_x = \frac{q}{4\pi\varepsilon_0} \frac{\gamma (x - x_q)}{[\gamma^2 (x - x_q)^2 + y^2 + z^2]^{3/2}}, \tag{11.22}$$

since $x_q = vt$ is the law of motion of the charge in the system K.

The other two components of the electric field strength are obtained in the same way:

$$E_y = \frac{q}{4\pi\varepsilon_0} \frac{\gamma y}{[\gamma^2 (x - x_q)^2 + y^2 + z^2]^{3/2}}, \tag{11.23}$$

$$E_z = \frac{q}{4\pi\varepsilon_0} \frac{\gamma z}{[\gamma^2 (x - x_q)^2 + y^2 + z^2]^{3/2}}. \tag{11.24}$$

The magnetic induction is calculated with the help of formulas (11.15). The result can be written in the vector form as follows:

$$\mathbf{B} = (1/c^2)\, \mathbf{v} \times \mathbf{E}, \tag{11.25}$$

where $\mathbf{E}$ is defined by formulas (11.22)-(11.24). It can be seen that the field lines of $\mathbf{B}$ are concentric circles with their centre on the X-axis along which the charge q moves.

The field configuration of a charge moving uniformly and rectilinearly does not change with time. Only the position of this configuration changes relative to the fixed coordinate system K. In other words, the invariable field configuration moves along with the charge. Let us consider this configuration at the instant when the charge is at the origin of the system K, i.e. when $x_q = 0$. In this case, we have [see (11.22)-(11.24)]

$$\mathbf{E} = \frac{q}{4\pi\varepsilon_0} \frac{\gamma \mathbf{r}}{(\gamma x^2 + y^2 + z^2)^{3/2}}. \tag{11.26}$$

where $\mathbf{r}$ is the radius vector from the point of location of the charge q to the point, where $\mathbf{E}$ is being determined. Thus, the field strength is directed along the radius vector, although its value depends on the direction of the latter. We denote by θ the angle between the direction of the velocity $\mathbf{v}$ of the charge and the radius vector. In this case, $x = r \cos\theta$, $y^2 + z^2 = r^2 \sin^2\theta$, $\gamma x^2 + y^2 + z^2 = r^2\gamma^2 (1 - \beta^2 \sin^2\theta)$, $\beta = v/c$ and formula (11.26) assumes the form

$$\mathbf{E} = \frac{q}{4\pi\varepsilon_0} \frac{\mathbf{r}}{r^3} \frac{1 - \beta^2}{(1 - \beta^2 \sin^2\theta)^{3/2}}.$$

The difference between the field of a moving charge and that of a fixed charge is manifested in a strong dependence of the field strength on the direction of the moving charge. The field strength along the line of motion of the charge ($\theta = 0$, $\theta = \pi$) and in a direction perpendicular to this line ($\theta = \pm\pi/2$) is given by

$$E_{\parallel} = \frac{q}{4\pi\varepsilon_0 r^2} (1 - \beta^2), \tag{11.27}$$

$$E_{\perp} = \frac{q}{4\pi\varepsilon_0 r^2} \frac{1}{\sqrt{1 - \beta^2}}. \tag{11.28}$$

www.EngineeringBooksPDF.com

For relativistic velocities ($\beta \simeq 1$), *the field of a moving charge at a given distance is weak along the line of its motion and strong in the direction perpendicular to it. In other words, the field is as if concentrated in the vicinity of the plane passing through the charge perpendicularly to its direction of motion.*

If an electric field exists in a certain coordinate system, a magnetic field also appears in another system, and vice versa. Through an appropriate choice of a reference system, we can strive to obtain simple configurations of electric and magnetic fields, or even eliminate one of them. However, it is not always possible to find a reference system in which the field can be reduced either to an electric field or to a magnetic one.

Proceeding from formulas for the transformation of quantities from the system K' to the system K, how can we obtain formulas for the transformation of the same quantities from the system K to the system K'? Verify by using formulas (11.5) as an example that the results obtained in both cases are the same.
Is the field of a rapidly moving point charge central? centrally symmetric?

Problems

1.1. Calculate div **r**.
1.2. Calculate grad (**r**·**A**), where **A** is a constant vector.
1.3. Calculate div ($\omega \times$ **r**), where ω is a constant vector.
1.4. Calculate div ($\mathbf{r}/r$).
1.5. Calculate div [**A** $\times$ (**r** $\times$ **B**)], where **A** and **B** are constant vectors.
1.6. What is the value of the magnetic induction at the centre of a square loop with side a, if a current I flows in it?
1.7. A wire is wound to form a spiral around a cylindrical insulator of radius a and makes n complete turns. The helical angle of the spiral is equal to α. Find the magnetic induction at the centre of the cylindrical insulator if a current I flows through the winding.

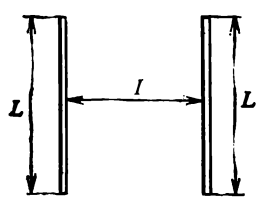


Fig. 30. Two finite-length conductors.

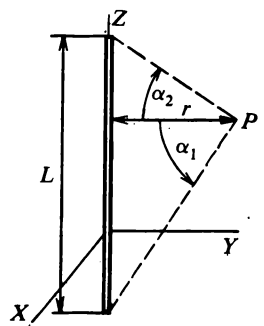


Fig. 31. Notation of angles in a chosen coordinate system.

1.8. Two point charges q and $-q$ are located at the points (a, 0, 0) and ($-a$, 0, 0) respectively. Find the electric field strength at the point (x, y, z).
1.9. A charge is distributed with a linear density τ over the length L along a radius vector drawn from the point where a point charge q is located. The distance between q and the nearest point on the linear charge is R. Find the force acting on the linear charge.
1.10. Two charges are distributed with the same linear density τ over a length L and are located parallel to each other at a distance l (Fig. 30). Find the force of interaction between them.

www.EngineeringBooksPDF.com

1.11. A disc has a surface charge density $\sigma = \alpha r^2$, where r is the distance from the centre of the disc. The radius of the disc is equal to r_0. Find the field strength at a height h along the normal passing through the centre of the disc.

1.12. Two uniformly charged surfaces are parallel to the plane X, Y and intersect the Z-axis at the points $z_1 = a_1$ and $z_2 = a_2 > a_1$. The surface charge densities are equal in magnitude but opposite in sign ($\sigma_1 = -\sigma_2$). Find the electric field strength at all points in space.

1.13. Find the strength of the electric field created at a point P by a charged filament of length L (Fig. 31). The linear charge density is τ. The point P lies in the (Y, Z) plane; this, however, does not involve any loss of generality of the solution since the field is axisymmetric.

1.14. A very long uniformly charged circular cylinder has a surface density σ. A very long filament charged to a linear density τ is arranged along the cylinder axis. Find the condition under which the electric field strength outside the cylinder is equal to zero.

1.15. A charge is uniformly distributed with the volume density $\rho = \alpha\sqrt{r}$ inside a spnere of radius a. Find the electric field strength.

1.16. A proton beam of circular cross section of radius 1 mm is accelerated by the potential difference of 10 kV. Assuming that the density of protons is constant over the cross section, find the volume charge density in the beam for the current of 5×10^{-6} A.

Answers

1.1 3. 1.2. $\mathbf{r}\cdot\mathbf{A}/r$. 1.3. 0. 1.4. $2/r$. 1.5. $2(\mathbf{A}\cdot\mathbf{B})$. 1.6. $2\sqrt{2}\,\mu_0 I/(\pi l)$. 1.7. $\frac{\mu_0 In}{2a} \times \frac{1}{\sqrt{1+\pi^2 n^2 \tan^2\alpha}}$ 1.8. $\mathbf{E} = \frac{q}{4\pi\varepsilon_0}\left\{\frac{(x-a)\,\mathbf{i}_x + y\mathbf{i}_y}{[(x-a)^2+y^2]^{3/2}} - \frac{(x+a)\,\mathbf{i}_x + y\mathbf{i}_y}{[(x+a)^2+y^2]^{3/2}}\right\}$ 1.9. $F = \frac{q\tau L}{4\pi\varepsilon_0 R\,(R+L)}$. 1.10. $F = \frac{\tau^2}{2\pi\varepsilon_0}\left[\left(1+\frac{L^2}{l^2}\right)^{1/2} - 1\right]$. 1.11. $E_h = \frac{\alpha h}{2\varepsilon_0}\left[\frac{r_0^2+2h^2}{(r_0^2+h^2)^{1/2}} - 2h\right]$. 1.12. $E_z = 0$ for $z < a_1$ and $z > a_2$; $E_z = \sigma_1/\varepsilon_0$ for $a_1 < z < a_2$. 1.13. $\mathbf{E} = \frac{\tau}{4\pi\varepsilon_0 r}[(\sin\alpha_1 + \sin\alpha_2)\,\mathbf{i}_y - (\cos\alpha_1 - \cos\alpha_2)\,\mathbf{i}_z]$. 1.14. $\tau = -2\pi r\sigma$. 1.15. $\mathbf{E} = \frac{2\alpha}{7\varepsilon_0}\sqrt{r}\,\mathbf{r}$, $0 < r < a$; $\mathbf{E} = \frac{2\alpha a^{7/2}}{7\varepsilon_0}\frac{\mathbf{r}}{r^3}$ for $r > a$. 1.16. $\rho = 1.15 \times 10^{-6}$ C/m^3.

www.EngineeringBooksPDF.com

CHAPTER 2

Constant Electric Field

Constant electric fields do not exist in nature since there are no fixed elementary charges. However, if the sum of elementary charges of each sign is nearly constant in an infinitely small physical volume, and if their average velocity is close to zero, the field generated by them at a sufficiently large distance from this volume element is nearly constant. It is called a constant electric field. A fixed point charge serves as the model of the charge generating this field. An aggregate of point charges may form space charge, surface charge or linear charges. As we go over to the continuous charge distribution model these aggregates are characterized by volume, surface, and linear charge densities.

Sec. 12. Constant Electric Field

The ideal model of a constant electric field and the limits of its applicability are discussed.

Fixed charge. Electrostatics studies electric fields of fixed charges. It is assumed that charges are held at various points in space by the forces of nonelectrostatic origin, whose nature is not specified in the framework of electrostatics. For example, electrostatics studies charge distribution over the surface of a conductor, the electric field created by these charges, the forces acting in the field, but does not explain why these charges remain on the surface of the conductor. The nature of forces holding the charges on the surface of conductors is not studied in electrostatics. The expression "a charge q is at the point (x, y, z) in vacuum" has a similar meaning. It is assumed that the charge q is as if fixed at the point (x, y, z) in space, and that there are no material particles in the vicinity of this charge (vacuum). Such a representation is obviously an idealization.
The essence of the model. There are no fixed elementary charges in nature, and hence *constant electric fields also do not exist.* However, in most of the phenomena studied in the classical theory of electricity, a superposition of fields created by many charges is observed rather than the field of an individual elementary charge. The contribution of the field of an individual elementary charge to the superposition of fields is rather small. Moreover, the electric field strength is defined as the quantity averaged over a certain physically small volume and a physically short interval of time. *The fluctuations from the mean value of the field strength are quite small. It is these mean values that are studied*

www.EngineeringBooksPDF.com

in the classical theory of electricity and magnetism. Hence, strictly speaking, *the constancy of electric field in time rather than the immobility of charges is essential for electrostatics.* In other words, in the model of constant fields the idealization is associated not with the constancy of the field but with the immobility of charges creating this field.

Limits of applicability of the model. Since the model is based on the existence of fields with very small fluctuations of mean values and not on the existence of fixed charges, *the limits of its applicability are determined by the requirement that the contribution of individual elementary charges to the observed field should be small.* Hence it follows, for example, that electrodynamics is inapplicable to the description of motion of individual electrons in an atom. Their motion is described by the quantum theory.

Sec. 13. Differential Form of Coulomb's Law

The physical factors responsible for the validity of the Gauss theorem are analysed. The differential form of the Coulomb law is given and its corollaries are discussed.

Gauss' theorem. The Gauss' electrostatic theorem establishes a mathematical relation between the electric flux through a closed surface and the charge located in the volume bounded by this surface.

Suppose that a point charge q is inside a volume V bounded by a closed surface S (Fig. 32). Let us consider the flux N of the field $\mathbf{E}$ across this surface:

$$N = \oint_S \mathbf{E} \cdot d\mathbf{S} \tag{13.1}$$

It should be recalled that the outward normal is always taken as the positive direction for closed surfaces. This means that the surface area element $d\mathbf{S}$ in (13.1) is directed outwards from the volume (Fig. 32). According to Coulomb's law, we have

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0} \frac{q}{r^2} \frac{\mathbf{r}}{r}. \tag{13.2}$$

Consequently, the integral in (13.1) can be expressed in the form

$$N = \frac{q}{4\pi\varepsilon_0} \oint_S \frac{1}{r^2} \left(\frac{\mathbf{r}}{r} \cdot d\mathbf{S} \right). \tag{13.3}$$

Let us consider the relation

$$\frac{\mathbf{r}}{r} \cdot d\mathbf{S} \left| \frac{\mathbf{r}}{r} \right| dS \cos (\widehat{\mathbf{r}, d\mathbf{S}}) = dS', \tag{13.4}$$

www.EngineeringBooksPDF.com

where dS' is the projection of the surface area element dS onto a plane perpendicular to the radius vector r. It is well known from geometry that

$$d\Omega = dS'/r^2, \tag{13.5}$$

where $d\Omega$ is the solid angle at which the surface area element dS' is seen from the origin of radius vectors, which coincides in the present case with the point where charge q is located. Taking (13.4) and (13.5) into consideration, we can write Eq. (13.3) in the form

$$N = \frac{q}{4\pi\varepsilon_0} \oint_S d\Omega. \tag{13.6}$$

The total solid angle at which a closed surface is seen from points inside the volume bounded by this surface is equal to 4π, i.e.

$$\oint_S d\Omega = 4\pi. \tag{13.7}$$

Consequently, we get from (13.6)

$$\boxed{N = q/\varepsilon_0.} \tag{13.8}$$

The flux of E through a closed surface when a point charge is located outside the volume bounded by it is calculated in a similar manner (Fig. 33) and is

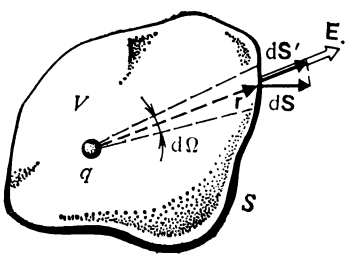


Fig. 32. Calculation of the electric flux through a closed surface when a point charge is inside the volume bounded by the surface.

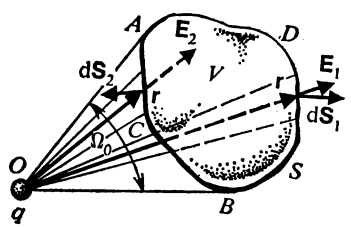


Fig. 33. Calculation of the electric flux through a closed surface when a point charge is outside the volume bounded by the surface.

defined by formula (13.3). The integrand, however, can now assume positive as well as negative values: if the angle $(\widehat{\mathbf{r}, d\mathbf{S}})$ is acute at some points on the surface, the integrand will be positive at these points. For points where the values of this angle are more than $\pi/2$, the integrand will be negative. This means that the integrand is positive on the surface ADB and negative on ACB. Hence the elements of the solid angle (13.5) on the surface ADB are positive, while on ACB, these elements are negative. We denote by Ω_0 the solid angle at the vertex

www.EngineeringBooksPDF.com

of a cone formed by the tangents from the point O to the surface under consideration (Fig. 33). In this case,

$$\oint_S \frac{1}{r^2}\left(\frac{\mathbf{r}}{r}\,d\mathbf{S}\right)=\int_{ADB} d\Omega-\int_{ACB} d\Omega=\Omega_0-\Omega_0=0, \tag{13.9}$$

since the surfaces ACB and ADB are seen from point O at the same solid angle Ω_0, but appear in the integral with different signs. When a point charge is outside the volume, the flux of field $\mathbf{E}$ through a closed surface is

$$N=0. \tag{13.10}$$

Combining results (13.8) and (13.10), we can write (13.1) in the following final form:

$$\oint_S \mathbf{E}\cdot d\mathbf{S}=\begin{cases} q/\varepsilon_0, & \text{when } q \text{ is inside the volume bounded by the surface } S; \\ 0, & \text{when } q \text{ is outside the volume bounded by the surface } S. \end{cases} \tag{13.11}$$

The statement expressed in (13.11) forms the content of the **Gauss electrostatic theorem for a point charge.**

The generalization of this theorem to a system of point charges is made with the help of the principle of superposition. For point charges q_i, the field strength $\mathbf{E}$ at each point is the sum of fields $\mathbf{E}_i$ created by each of the point charges:

$$\mathbf{E}=\sum \mathbf{E}_i. \tag{13.12}$$

Consequently,

$$\oint_S \mathbf{E}\,d\mathbf{S}=\sum_i \oint_S \mathbf{E}_i\,d\mathbf{S}. \tag{13.13}$$

While calculating each integral in the sum on the right-hand side of (13.13), we must take into account Eq. (13.11): the integral is equal to q_i/ε_0 for a point charge contained inside the volume, while it is equal to zero for a charge outside the volume. Consequently, Eq. (13.13) assumes the form

$$\oint_S \mathbf{E}\cdot d\mathbf{S}=\frac{1}{\varepsilon_0}\sum_V q_i=\frac{1}{\varepsilon_0}Q, \tag{13.14}$$

where V under the sign Σ means that the summation is carried out only over the charges located inside the volume V. The total charge inside this volume is denoted by Q in (13.14):

$$Q=\sum_V q_i. \tag{13.15}$$

www.EngineeringBooksPDF.com

Taking into account definition (4.1) for the volume density ρ for a continuous charge distribution, we can write formula (13.14) in the form

$$\boxed{\oint_S \mathbf{E}\cdot d\mathbf{S} = \frac{1}{\varepsilon_0}\int_V \rho\, dV = \frac{1}{\varepsilon_0} Q,} \tag{13.16}$$

where

$$Q = \int_V \rho\, dV \tag{13.17}$$

is the total charge in the volume bounded by the closed surface S. The statement expressed in formula (13.16) forms the content of Gauss' electrostatic theorem for a continuous charge distribution. Obviously this formula also contains expressions (13.14) and (13.11) as particular cases.

Measurement of charge. Gauss' theorem allows us to find the total charge contained in a volume by measuring the field flux through the surface bounding this volume. *Other methods of determining the charge do not give satisfactory results.* For example, it is impossible to find this charge by measuring the force exerted by it on a test charge outside this volume, since the force depends not only on the total charge, but also on its distribution over the volume which is, generally speaking, unknown. The charge can be determined by measuring the force acting on it in a known uniform external electric field. In this case, the field must be uniform. Clearly, this method is applicable only when the external uniform field does not significantly change the charge distribution inside the volume.

Physical foundation of the validity of Gauss' theorem. It can be seen from the derivation of Gauss' theorem that its validity stems from the possibility of reducing the integrand in (13.3) to the differential $d\Omega$ of the solid angle with the help of (13.4) and (13.5). This is possible only when $\mathbf{E}(r)$ decreases in inverse proportion to the square of the distance from the point charge. For any other dependence of $\mathbf{E}$ on r, the integrand in (13.6) must contain, besides the differential of the solid angle, a certain function of r which makes it impossible to express the electric flux through a surface as a function of charge. This is a violation of the Gauss theorem. Hence, *Coulomb's law forms the physical foundation of the Gauss theorem or, in other words, the Gauss theorem is the integral form of Coulomb's law.*

Differential form of Coulomb's law. Maxwell's equation for div E. With the help of the divergence formula (5.21), the flux of $\mathbf{E}$ through a closed surface can be transformed into the volume integral of div $\mathbf{E}$:

$$\oint_S \mathbf{E}\cdot d\mathbf{S} = \int_V \operatorname{div}\mathbf{E}\, dV, \tag{13.18}$$

Consequently, formula (13.16) assumes the form

$$\int_V (\operatorname{div}\mathbf{E} - \rho/\varepsilon_0)\, dV = 0. \tag{13.19}$$

www.EngineeringBooksPDF.com

This equality is satisfied for any volume V. Consequently, the integrand is identically equal to zero, i.e.

$$\boxed{\operatorname{div} \mathbf{E} = \rho/\varepsilon_0.} \tag{13.20}$$

The validity of (13.20) and of the Gauss theorem depends on the validity of Coulomb's law. Consequently, (13.20) is the differential form of Coulomb's law. The linearity of Eq. (13.20) reflects the validity of the principle of superposition for field strength. It has been derived here for fixed charges, although it is assumed that *this equation is valid for an arbitrary motion of charges.*

Lines of force. *An electric field line* (*line of force*) *is a line the tangent to which at any point coincides with the field strength* **E**. The lines of force are a convenient

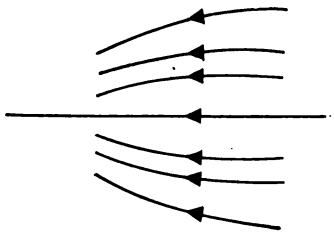


Fig. 34. Lines of force for a field whose strength increases from right to left

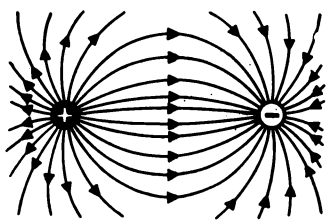


Fig. 35. Lines of force for two unlike charges

way in which a field can be graphically represented. The field strength is conventionally characterized by the number of field lines intersecting 1 m^2 of the surface perpendicular to the direction of the field lines at the corresponding point: the higher the density of the lines of force, the stronger the field. Figure 34 shows an electric field whose strength increases from right to left.

Sources and sinks of field E. It can be seen from Eq. (13.20) that the *field lines start where* $\operatorname{div} \mathbf{E} > 0$ *and terminate where* $\operatorname{div} \mathbf{E} < 0$. *In other words, the field lines originate at positive charges and terminate at negative ones.* Consequently, the positive and negative charges are respectively called the sources and sinks of field **E**. Such a distinction between charges is, of course, arbitrary and stems from the definition of the direction of a field. Positive and negative charges play an exactly identical role in the creation of an electric field. Figure 35 shows the lines of force for two unlike charges.

Charge invariance. Let us find the flux of **E** through a closed surface surrounding a point charge q which is moving uniformly and rectilinearly. The field created by this charge is given by formula (11.26). The electric flux is

$$N = \oint \mathbf{E} \cdot d\mathbf{S} = \oint Er^2 \, d\Omega = \oint Er^2 \sin\theta \, d\theta \, d\varphi, \tag{13.21}$$

where the surface of integration is taken in the form of a sphere having the moving point charge as its centre at a certain instant of time. It is considered

www.EngineeringBooksPDF.com

that **E** and d**S** are collinear with the radius vector **r**; θ and φ are the polar and the axial angle respectively of a polar coordinate system whose polar axis coincides with the X-axis of the fixed coordinate system. Substituting (11.26) into (13.21), we get

$$N = \frac{q(1-\beta^2)}{2\varepsilon_0} \int_0^{\pi} \frac{\sin\theta\, d\theta}{(1-\beta^2\sin^2\theta)^{3/2}}, \tag{13.22}$$

where the integration has been carried out over the angle $d\varphi$ on which the integrand in (13.21) does not depend. Since $\sin^2\theta = 1 - \cos^2\theta$, and $\sin\theta\, d\theta = -d\cos\theta$, we get

$$\int_{\theta}^{\pi} \frac{\sin\theta\, d\theta}{(1-\beta^2\sin^2\theta)^{3/2}} = 2\int_{\theta}^{1} \frac{dx}{(1-\beta^2+\beta^2x^2)^{3/2}} = \frac{2}{\beta^3}\left[\frac{x}{a^2\sqrt{a^2+x^2}}\right]_0^1 = \frac{2}{1-\beta^2},$$

where $a^2 = (1 - \beta^2)/\beta^2$. In this case, relation (13.22) assumes the form

$$N = q/\varepsilon_0, \tag{13.23}$$

which is identical with (13.8). This proves that the Gauss theorem is also valid for a point charge moving uniformly and rectilinearly. If the charge in a volume is determined by means of the flux of **E** through a closed surface bounding this volume, Eq. (13.23) expresses the invariance of charge.

The Gauss theorem expresses a relation between the electric field flux through a closed surface and the charge in the volume bounded by this surface. Coulomb's law serves as the physical foundation of the Gauss theorem. In other words, the Gauss theorem is the integral form of Coulomb's law.

A line the tangent to which at each point coincides with the electric field vector is called the line of force of this field.
Positive charges are the sources and negative charges are the sinks of the electric field. This distinction between charges, however, is purely arbitrary. They play identical roles in generating an electric field.

Sec. 14. Potential Nature of Electrostatic Field

The integral and differential forms of the definition of the potential nature of field are discussed. The scalar potential is introduced and its properties are analysed. The potential of charges distributed in a finite region of space is calculated. The Earnshaw theorem is proved.

Work in an electric field. Since the force acting on a point charge q in an electric field is $\mathbf{F} = q\mathbf{E}$, the work done to displace the charge by d**l** is

$$dA = \mathbf{F}\cdot d\mathbf{l} = q\mathbf{E}\cdot d\mathbf{l}. \tag{14.1}$$

www.EngineeringBooksPDF.com

The work performed per unit charge in displacing the charge is defined as the ratio of the work to the charge:

$$dA' = dA/q = \mathbf{E}\cdot d\mathbf{l} \tag{14.2}$$

and is measured in joules per coulomb. It follows from (14.2) that *the work performed by the field is assumed to be positive while the work done by the forces external relative to the field is assumed to be negative.* This sign rule is similar to that used in thermodynamics for the work of a system.

When a charge is displaced from point *1* to point *2* along the trajectory L (Fig. 36), the work per unit charge is

$$A' = \int_{(1)}^{(2)}_{L} \mathbf{E}\cdot d\mathbf{l}. \tag{14.3}$$

Potential nature of a Coulomb field. A force field is called a potential field if the work done upon a displacement in this field depends only on the initial and final points of the path and does not depend on the trajectory. *An equivalent definition of the potential nature is the requirements that the work must be equal to zero upon a displacement along any closed contour.*

It is well known that the force of gravity of a mass, which is inversely proportional to the square of the distance, is a potential force, its potential nature being determined precisely by this dependence on distance. Since the Coulomb force of a point charge decreases with distance according to the same law, it is a potential force. The entire mathematical part of the concept of potential was developed in the theory of gravitation. This concept first appeared in 1777 in the works of Lagrange (1736-1813), although he did not apply this term to the function representing the potential. The term "potential" was introduced in science in 1828 by Green and independently by Gauss (1777-1855). Laplace (1749-1827) and Poisson (1781-1840) also made significant contributions to the theory of potential.

According to the principle of superposition, the potential nature of an arbitrary electrostatic field follows from the potential nature of the field of a point charge. Mathematically, this statement is proved as follows:

$$\oint \mathbf{E}\cdot d\mathbf{l} = \oint \left(\sum_i \mathbf{E}_i\right)\cdot d\mathbf{l} = \sum_i \oint \mathbf{E}_i\cdot d\mathbf{l} = \sum_i 0 = 0, \tag{14.4}$$

where

$$\mathbf{E} = \sum \mathbf{E}_i, \quad \oint \mathbf{E}_i \cdot d\mathbf{l} = 0. \tag{14.5}$$

Curl of a vector. The criterion of the potential nature of the field used by us so far does not have a differential form and for this reason cannot always be applied easily and effectively. Its application boils down to the verification of the statement that the work over any closed contour is equal to zero. This means that

www.EngineeringBooksPDF.com

we must investigate an infinite number of closed paths, which is generally impossible. This criterion can be applied only when the general expression for the work over any path is given in an analytic form. Such a formula can be obtained only in rare cases. Hence it is desirable to obtain another criterion for the potential nature, which could be easily and conveniently applied in practice. The differential form given with the help of the curl of a vector serves as such a criterion.

First of all, let us consider the definition of the curl of vector **A** (denoted by curl **A**) in vector form. A vector is specified by three components which do not

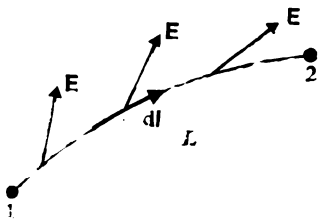


Fig. 36. Work performed in an electric field during the displacement of a point charge

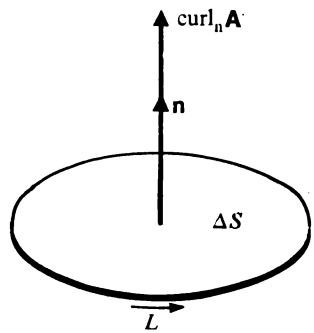


Fig. 37. To the vector definition of curl

lie in the same plane. We choose a certain direction characterized by a unit vector **n**. In the plane perpendicular to **n** we bound a surface area element ΔS by a very small closed contour L (Fig. 37). As usual, the direction of the positive circumvention of the contour L is connected with the direction of **n** by the right-hand screw rule. The **curl** is a vector whose projection onto the direction of **n** is defined by

$$\operatorname{curl}_n \mathbf{A} = \lim_{\Delta S \to 0} \frac{\oint \mathbf{A} \cdot d\mathbf{l}}{\Delta S}. \qquad (14.6)$$

The curl characterizes the vorticity of a vector, which is reflected in the name of the operation. Suppose, for example, that the vector **A** is equal to the velocity **v** of the points of a rigid body rotating at an angular velocity $\boldsymbol{\omega}$ about an axis collinear with **n**. We shall find $\operatorname{curl}_n \mathbf{v}$ for points lying on the axis of rotation. For the contour L, we choose a circle of radius r having its centre on the axis and lying in a plane perpendicular to this axis. Obviously, we have $v = \omega r$, $\Delta S = \pi r^2$, and $\mathbf{A} \cdot d\mathbf{l} = v\, dl$, where dl is the scalar value of an element of the circle. Hence, using (14.6), we obtain

$$\operatorname{curl}_n \mathbf{v} = \lim_{r \to 0} \frac{\omega r \oint dl}{\pi r^2} = \lim_{r \to 0} \frac{\omega r 2\pi r}{\pi r^2} = 2\omega, \qquad (14.7)$$

www.EngineeringBooksPDF.com

where $\oint dl = 2\pi r$ is the circumference of the circle. Thus, the curl of the linear velocity of the points of a perfectly rigid rotating body is equal to twice the angular velocity of its rotation. It can be shown that this statement is valid not only for the points lying on the axis of rotation but for all points.

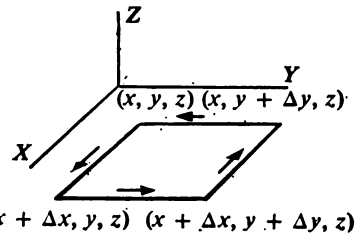


Fig. 38. To the definition of curl in terms of coordinates

For a practical computation of a curl, it is more convenient to use formulas written in terms of coordinates instead of (14.6). Let us find the projections of curl A in a rectilinear Cartesian system of coordinates. We take, for instance, the Z-axis (Fig. 38). The contour L is a rectangle with the sides Δx and Δy. The direction of positive circumvention is shown in the figure. In this case, we obtain

$$\oint_L \mathbf{A}\cdot d\mathbf{l} = \int_{(x, y, z)}^{(x+\Delta x, y, z)} A_x(x, y, z)\,dx + \int_{(x+\Delta x, y, z)}^{(x+\Delta x, y+\Delta y, z)} A_y(x+\Delta x, y, z)\,dy + \int_{(x+\Delta x, y+\Delta y, z)}^{(x, y+\Delta y, z)} A_x(x, y+\Delta y, z)\,dx + \int_{(x, y+\Delta y, z)}^{(x, y, z)} A_y(x, y, z)\,dy, \quad (14.8)$$

where the integration is performed along the sides of the rectangle between its vertices whose coordinates are used in (14.8) as limits of integration. Considering that Δx and Δy can be as small as desired, in the integrands of the second and third integrals we can expand A_x and A_y into a series in Δx and Δy and confine ourselves to the linear terms:

$$A_x(x, y+\Delta y, z) = A_x(x, y, z) + \Delta y \frac{\partial A(x, y, z)}{\partial y} + \ldots \quad (14.9a)$$

$$A_y(x+\Delta x, y, z) = A_y(x, y, z) + \Delta x \frac{\partial A_y(x, y, z)}{\partial x} + \ldots \quad (14.9b)$$

Let us calculate the sum of the first and third integrals:

$$I_1 = \int_{(x, y, z)}^{(x+\Delta x, y, z)} A_x(x, y, z)\,dx + \int_{(x+\Delta x, y+\Delta y, z)}^{(x, y+\Delta y, z)} A_x(x, y+\Delta y, z)\,dx$$

$$= \int_{(x, y, z)}^{(x+\Delta x, y, z)} A_x(x, y, z)\,dx - \int_{(x, y, z)}^{(x+\Delta x, y, z)} \left[A_x(x, y, z) + \Delta y \frac{\partial A_x(x, y, z)}{\partial y}\right] dx, \quad (14.10)$$

where in calculating the second integral in (14.10) we use formula (14.9a), and the minus sign appears as a result of the reversal of the order of integration. In this formula, the terms containing $A_x(x, y, z)$ in the integrands are cancel-

www.EngineeringBooksPDF.com

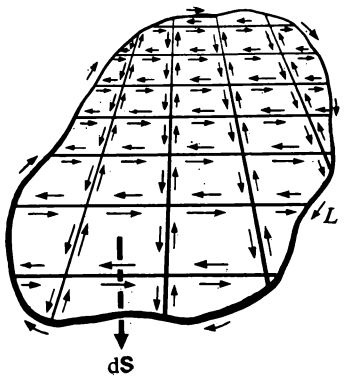


Fig. 39. To the proof of Stokes' theorem

led out, and hence

$$I_1 = -\frac{\partial A_x(x, y, z)}{\partial y}\Delta y\,\Delta x. \qquad (14.11)$$

Similarly, we calculate the sum of the second and fourth integrals in (14.8):

$$I_2 = \frac{\partial A_y(x, y, z)}{\partial x}\Delta x\,\Delta y. \qquad (14.12)$$

Using formula (14.6), we find

$$(\operatorname{curl}\mathbf{A})_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}. \qquad (14.13)$$

The projections onto the other axes of coordinates are calculated in a similar way:

$$(\operatorname{curl}\mathbf{A})_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z},\quad (\operatorname{curl}\mathbf{A})_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}. \qquad (14.14)$$

Denoting, as usual, the unit vectors of the coordinate axes by $\mathbf{i}_x$, $\mathbf{i}_y$, and $\mathbf{i}_z$, we can write the vector curl A in the form

$$\operatorname{curl}\mathbf{A} = \mathbf{i}_x\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + \mathbf{i}_y\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + \mathbf{i}_z\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right). \qquad (14.15)$$

Stokes' integral theorem. This theorem relates the circulation of a vector around the contour bounding a surface to the flux of its curl through this surface. Its derivation is based on definition (14.6). Let us calculate the flux of the vector curl A through the surface S bounded by a contour L (Fig. 39), dividing the surface into elements ΔS_i:

$$\int_S \operatorname{curl}\mathbf{A}\cdot d\mathbf{S} = \sum_i \int_{\Delta S_i} \operatorname{curl}\mathbf{A}\cdot d\mathbf{S}. \qquad (14.16)$$

Since ΔS_i are very small, we obtain, according to (14.6), the following expression for each element

$$\int_{\Delta S_i} \operatorname{curl}\mathbf{A}\cdot d\mathbf{S} = \int_{\Delta S_i} (\operatorname{curl}\mathbf{A})_n\, dS \approx (\operatorname{curl}\mathbf{A})_n\,\Delta S \approx \oint_{L_i} \mathbf{A}\cdot d\mathbf{l}, \qquad (14.17)$$

where L_i is the contour bounding ΔS_i. Hence Eq. (14.6) can be represented in the form

$$\int_S \operatorname{curl}\mathbf{A}\cdot d\mathbf{S} \approx \sum_i \oint_{L_i} \mathbf{A}\cdot d\mathbf{l}. \qquad (14.18)$$

The parts of the contours L_i, which are the boundaries between ΔS_i's, appear in two terms of the sum (14.18): once in the integration along the contour of a given area element ΔS_i, and for the second time, in the integration along the contour of the neighbouring area element. These integrals are equal in magnitude but opposite in sign, since the paths of integration along the boundary have op-

www.EngineeringBooksPDF.com

posite directions. Thus, in (14.18) all parts of the integrals over the boundaries between ΔS_i cancel out, and we are left only with the sum of integrals over those parts of the contours L_i which do not form a boundary between ΔS_i. In other words, we are left with the integral over the contour L bounding the surface S. As $\Delta S_i \to 0$, the approximate equality (14.18) becomes exact:

$$\boxed{\int_S \operatorname{curl} \mathbf{A} \cdot d\mathbf{S} = \oint_L \mathbf{A} \cdot d\mathbf{l}.} \tag{14.19}$$

This relation is known as **Stokes' integral theorem.**

Differential form of the potential nature of the field. The fact that the work performed during the displacement of a charge in an electrostatic field is independent of the path is expressed by the equality

$$\int_{A\,L_1}^{B} \mathbf{E} \cdot d\mathbf{l} = \int_{A\,L_2}^{B} \mathbf{E} \cdot d\mathbf{l}, \tag{14.20}$$

where L_1 and L_2 are different paths between points A and B. Considering that $\int_{A\,L_2}^{B} \mathbf{E} \cdot d\mathbf{l} = -\int_{B\,L_2}^{A} \mathbf{E} \cdot d\mathbf{l}$, we can represent (14.20) in the form

$$\int_{A\,L_1}^{B} \mathbf{E} \cdot d\mathbf{l} + \int_{B\,L_2}^{A} \mathbf{E} \cdot d\mathbf{l} = \oint_L \mathbf{E} \cdot d\mathbf{l} = 0, \tag{14.21}$$

where $L = L_1 + L_2$. This formula is a mathematical expression of the statement that the work done in displacing a charge over any closed contour in an electrostatic field is equal to zero.

Using (14.19) and (14.21), we get

$$\boxed{\int_S \operatorname{curl} \mathbf{E} \cdot d\mathbf{S} = 0,} \tag{14.22}$$

where S is the surface bounded by the contour L. Since S is arbitrary, it follows from (14.22) that

$$\operatorname{curl} \mathbf{E} = 0. \tag{14.23}$$

This equation is the **differential form of the statement that the electrostatic field is a potential field.**

Gradient. Let $\varphi(x, y, z)$ be a scalar position function. The gradient of φ is defined as the vector

$$\boxed{\operatorname{grad} \varphi = \mathbf{i}_x \frac{\partial \varphi}{\partial x} + \mathbf{i}_y \frac{\partial \varphi}{\partial y} + \mathbf{i}_z \frac{\partial \varphi}{\partial z}.} \tag{14.24}$$

www.EngineeringBooksPDF.com

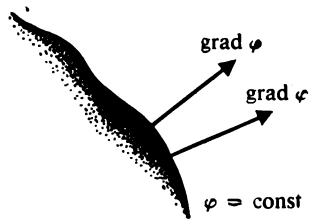


Fig. 40. The direction of grad φ

In order to elucidate the meaning of this vector, let us calculate the total differential of the function φ upon a displacement by $\mathbf{dr} = \mathbf{i}_x\, dx + \mathbf{i}_y\, dy + \mathbf{i}_z\, dz$:

$$d\varphi = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} dz = \operatorname{grad} \varphi \cdot \mathbf{dr}. \tag{14.25}$$

Thus, an infinitely small increment dφ due to a displacement in a certain direction is equal to the component of grad φ in this direction, multiplied by the magnitude of the displacement. We construct a family of surfaces φ = const (Fig. 40). As we move over the surface φ = const, dφ=0. Hence [see (14.25)], grad φ ⊥ dr, i.e. *the vector* grad φ *is normal to the surface* φ=const. The magnitude of this vector is equal to the derivative of φ with respect to the direction perpendicular to the surface φ = const.

Scalar potential. Since the work done in displacing a charge in a potential field does not depend on the path and depends only on the initial and final points of the trajectory, it can be expressed in terms of the coordinates of these points. This can be done with the help of potential.

It can be directly verified that the following identity always holds:

$$\operatorname{curl} \operatorname{grad} \varphi = 0. \tag{14.26}$$

Consequently, Eq. (14.23) will be satisfied if **E** is represented in the form

$$\boxed{\mathbf{E} = -\operatorname{grad} \varphi.} \tag{14.27}$$

The sign is chosen so that the field strength **E** is directed towards decreasing values of φ. The scalar function φ related to the field strength **E** through formula (14.27) is called the scalar potential of electric field.

Field strength can be measured experimentally. *The potential* φ *does not have any definite numerical value, and it is meaningless to speak about an experimental measurement of its value.*

Ambiguity of scalar potential. Formula (14.27) shows that if a certain constant is added to φ, the field described by this potential does not change since the derivatives of a constant quantity with respect to coordinates are equal to zero. Consequently, *the potential* φ *of a given electric field is defined only to within an additive constant.*

Normalization. In view of the ambiguity of scalar potential, we can ascribe to it any preset value at any preset point. After this, the potential at all other points has a quite definite value, i.e. it will be single-valued. *This procedure of making the potential single-valued by ascribing to it a certain value at one of the points is called potential* **normalization**. When electric fields are investigated near the Earth surface, the potential of the Earth is usually taken as the zero potential. When general questions are analysed, and the charges are located in a finite region of space, it is more convenient to assume that the po-

www.EngineeringBooksPDF.com

tential is zero at an infinite distance from the charges. Such a normalization will be frequently used in this book.

Expression of work in terms of potential. If a charge moves between points *1* and *2*, the work per unit charge is

$$A' = \int_{(1)}^{(2)} \mathbf{E}\cdot d\mathbf{l} = -\int_{(1)}^{(2)} \operatorname{grad}\varphi\cdot d\mathbf{r} = -\int_{(1)}^{(2)} d\varphi = \varphi(1) - \varphi(2), \qquad (14.28)$$

where we have used formula (14.25) and the fact that $d\mathbf{l} = d\mathbf{r}$. It follows from this equation that the work indeed depends on the initial and final points of the path and does not depend on the shape of the trajectory. It also shows that the potential difference between two points has a clear physical meaning and can be measured experimentally. Thus, *it is not the potential itself but the potential difference between different points that has a physical meaning.*

Field potential of a point charge. We shall normalize the potential to zero at infinity. Assuming that point *2* in formula (14.28) is at infinity, we put $\varphi(2) = \varphi(\infty) = 0$ and obtain the following expression for the potential at point *1*:

$$\varphi(1) = \int_{(1)}^{\infty} \mathbf{E}\cdot d\mathbf{l}. \qquad (14.29)$$

The path from point *1* to infinity can be arbitrary. However, we must choose it so as to simplify the calculations as much as possible.

The field of a point charge is spherically symmetric. In accordance with formula (14.29), the potential at a distance r from the point charge q is given by

$$\varphi(r) = \frac{q}{4\pi\varepsilon_0}\int_{r}^{\infty} \frac{1}{r^2}\left(\frac{\mathbf{r}}{r}\cdot d\mathbf{l}\right). \qquad (14.30)$$

Since for any $d\mathbf{l}$ the equality $(\mathbf{r}\, d\mathbf{l}/r) = dr$ is observed, it follows from (14.30) that

$$\boxed{\varphi(r) = \frac{q}{4\pi\varepsilon_0}\int_{r}^{\infty} \frac{dr}{r^2} = \frac{q}{4\pi\varepsilon_0}\left[-\frac{1}{r}\right]_r^{\infty} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r}.} \qquad (14.31)$$

We recommend to the reader to verify that Coulomb's law

$$\mathbf{E} = -\operatorname{grad}\varphi = -\frac{q}{4\pi\varepsilon_0}\operatorname{grad}\frac{1}{r} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}\frac{\mathbf{r}}{r} \qquad (14.32)$$

can be obtained from this formula.

Field potential of a system of point charges. According to the principle of superposition, *the field potential of a system of point charges is equal to the sum of the potentials created at a given point by each of the charges.* This is obvious since

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = -\operatorname{grad}\varphi_1 - \operatorname{grad}\varphi_2 = -\operatorname{grad}(\varphi_1 + \varphi_2).$$

www.EngineeringBooksPDF.com

Consequently, using formula (14.31) we can write the following expression for the potential created by a system of point charges q_i:

$$\boxed{\varphi(x, y, z) = \frac{1}{4\pi\varepsilon_0} \sum \frac{q_i}{r_i},} \tag{14.33}$$

where $r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$ is the distance from the point charge q_i located at the point (x_i, y_i, z_i) to the point (x, y, z) at which the potential is calculated.

Field potential of continuously distributed charges. As before, *we assume that all charges are located in a finite region of space and that the potential is normalized to zero at infinity.* Denoting the volume charge density by $\rho(x', y', z')$, we obtain the following expression for the potential instead of (14.33):

$$\varphi(x, y, z) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho(x', y', z')\, dx'\, dy'\, dz'}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}}. \tag{14.34}$$

This formula can be written in a more compact form (without detailed specification of variables):

$$\boxed{\varphi = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho\, dV}{r},} \tag{14.35}$$

where dV is the volume element over which the integration is performed. This brief form will be often used in further analysis.

Field potential of surface charges. If a charge is located on the surface, the charge distribution is characterized by the surface charge density σ. On the area element dS (which is a scalar and not the vector of area element), there is a charge $\sigma\, dS$, and hence the potential at a certain point is given by a formula similar to (14.35):

$$\boxed{\varphi = \frac{1}{4\pi\varepsilon_0} \int_S \frac{\sigma\, dS}{r},} \tag{14.36}$$

where r is the distance between the area element dS and the point at which the potential is calculated. Integral (14.36) is valid for all surfaces carrying surface charges.

Infinite value of the field potential of a point charge. It follows from (14.31) that as $r \to 0$, the potential $\varphi\ (r \to 0) \to \infty$. This is due to the fact that the volume density of a point charge is formally equal to infinity since its volume is equal to zero. *It is the infinite volume charge density that is responsible for the infinite value of the potential.*

Finite value of the potential for a continuous charge distribution with a finite density. If a charge is distributed continuously with a finite density, the potential does not assume infinite value anywhere. This can be verified by calculating the potential with the help of formula (14.34). We take the point $(x,$

www.EngineeringBooksPDF.com

y, z) as the origin ($x = y = z = 0$) and carry out calculations in a spherical system of coordinates. In this system, the volume element is expressed by $dx'\,dy'\,dz' = r'^2 \sin\theta'\,d\theta'\,d\alpha'\,dr'$, where $r' = \sqrt{x'^2 + y'^2 + z'^2}$. Then [see (14.34)]

$$\varphi(0, 0, 0) = \frac{1}{4\pi\varepsilon_0} \int \rho(r', \alpha', \theta')\, r' \sin\theta'\, d\theta'\, d\alpha'\, dr'. \qquad (14.37)$$

Consequently, *if* ρ *is finite, the potential* φ *is also finite,* Q.E.D.

Continuity of potential. The derivative of the potential with respect to a Cartesian coordinate gives the corresponding component of the electric field strength. Obviously, the intensity cannot be infinite. Therefore, the derivatives of the potential with respect to coordinates must be finite. This means that the potential is a continuous function. Thus, *the potential* φ *is a continuous and finite function with finite coordinate derivatives.* These conditions are important for solving differential equations for potential.

Earnshaw's theorem. This theorem states that *there exists no configuration of fixed charges, which would be stable in the absence of forces other than the forces of Coulomb's interaction between the charges of the system.*

The proof of the Earnshaw theorem follows from the Gauss theorem. Suppose that the equilibrium is stable. Then the displacement of any charge from the equilibrium position in any direction will give rise to a force tending to return the charge to the initial position. And this means that the field created in the vicinity of each fixed charge by all other charges is directed along the radii emerging from the point of location of this charge. The flux of this field through a closed surface around the charge differs from zero, since the field has the same direction along the radii (in the vicinity of a positive charge it is directed to the charge, while near a negative charge, it is directed away from it). In accordance with the Gauss theorem, the flux through a closed surface is created by the charge located in the volume bounded by this surface. This is in contradiction with the initial assumption according to which the flux is created by the charges located outside the volume. Thus, the assumption about the equilibrium configuration of fixed charges is rejected and the Earnshaw theorem is proved.

Stable configurations of fixed charges may exist only when, in addition to the forces of interaction between them, there are some extraneous forces holding the charges in equilibrium positions. Stable states of moving charges are possible, for example, in the form of the motion of two unlike charges in ellipses around the centre of mass (naturally, if we ignore radiation).

The sign rule: the work done by the field is assumed to be positive while the work done by the forces external to the field is assumed to be negative.

The differential form of the statement that the electrostatic field is of a potential nature: curl $\mathbf{E} = 0$.

The minus sign in the expression $\mathbf{E} = -\text{grad}\,\varphi$ is chosen conventionally to show that E is directed towards decreasing φ.

The application of Poisson's equation for solving a problem is not based on the assumption that the potential is normalized in a certain way and that there are no charges at infinity. The potential is a continuous and finite function having finite derivatives with respect to coordinates.

www.EngineeringBooksPDF.com

What methods do you know for determining the field strength for a given charge distribution? What determines the choice of the method of solving the problem in each specific case?
What are the advantages of determining the field strength by solving the Laplace and Poisson equations in comparison with other methods?
What are the properties of the potential as a solution of the corresponding differential equations?
Which forms of the statement of the potential nature of an electrostatic field do you know? What are the advantages of the differential form?
What physical factors determine the possibility of the normalization of the scalar potential? What normalization conditions are used more frequently and when are they most expedient?

Example 14.1. *Calculate* grad $\varphi(r)$.

We have

$$\operatorname{grad}\varphi = \mathbf{i}_x \frac{\partial\varphi}{\partial x} + \mathbf{i}_y \frac{\partial\varphi}{\partial y} + \mathbf{i}_z \frac{\partial\varphi}{\partial z},$$

$$\frac{\partial\varphi}{\partial x} = \frac{\partial\varphi}{\partial r}\frac{\partial r}{\partial x} = \varphi' \frac{\partial r}{\partial x}, \quad r = \sqrt{x^2+y^2+z^2}.$$

Similarly, we calculate $\partial\varphi/\partial y$ and $\partial\varphi/\partial z$. The prime denotes the derivative with respect to the argument r. Considering that $\frac{\partial r}{\partial x} = \frac{2x}{2\sqrt{x^2+y^2+z^2}} = \frac{x}{r}$, we obtain

$$\operatorname{grad}\varphi(r) = \frac{d\varphi}{dr}(\mathbf{i}_x x + \mathbf{i}_y y + \mathbf{i}_z z) = \frac{d\varphi}{dr}\frac{\mathbf{r}}{r}.$$

In particular, grad $r = \mathbf{r}/r$, for $\varphi(r) = r$, while grad $(1/r) = -\mathbf{r}/r^3$ for $\varphi(r) = 1/r$.

Example 14.2. *Calculate the circulation of the vector* $\boldsymbol{\omega} \times \mathbf{r}$ *around the circle* L *of radius* r_0, *lying in the plane perpendicular to the constant vector* $\boldsymbol{\omega}$, *both directly and with the help of Stokes' theorem. The centre of the circle coincides with the origin of coordinates.*

The vector $\boldsymbol{\omega} \times \mathbf{r}_0$ at each point is directed along the tangent to the circle. Consequently,

$$\oint_L \boldsymbol{\omega}\times\mathbf{r}\cdot d\mathbf{l} = \omega r_0 \int_L dl = 2\pi\omega r_0^2. \tag{14.38}$$

The direction of circumvention is chosen so that the vectors $\boldsymbol{\omega} \times \mathbf{r}$ and $d\mathbf{l}$ are collinear at each point. If the direction of circumvention is reversed, the integral will have the opposite sign.

With the help of Stokes' theorem, the problem is solved in a different way.

$$\oint_L \boldsymbol{\omega}\times\mathbf{r}\cdot d\mathbf{l} = \int_S \operatorname{rot}(\boldsymbol{\omega}\times\mathbf{r})\cdot d\mathbf{S},$$

where S is the surface bounded by the circle L. For $\boldsymbol{\omega}$ = const, curl $(\boldsymbol{\omega} \times \mathbf{r}) = 2\boldsymbol{\omega}$ and

$$\int_S \operatorname{curl}(\boldsymbol{\omega}\times\mathbf{r})\cdot d\mathbf{S} = 2\int_S \boldsymbol{\omega}\cdot d\mathbf{S} = 2\omega\int_S dS = 2\pi\omega r_0^2, \tag{14.39}$$

which, as expected, coincides with (14.38).

In can be easily seen that the surface S can be any surface stretched over the circle and not only a plane surface. We have

$$\int_{S_1} \operatorname{curl}(\boldsymbol{\omega}\times\mathbf{r})\cdot d\mathbf{S} = 2\int_{S_1} \boldsymbol{\omega}\cdot d\mathbf{S} = 2\boldsymbol{\omega}\cdot\int_{S_1} d\mathbf{S}. \tag{14.40}$$

www.EngineeringBooksPDF.com

We take into account that

$$\oint_{S'} dS = 0, \tag{14.41}$$

where S' is the closed surface consisting of the surface S_1 from (14.40) and the surface S of the circle from (14.39), i.e. $S' = S_1 + S$. Form (14.41) we obtain

$$\int_{S_1} dS = -n\pi r_0^2, \tag{14.42}$$

where n is a unit vector perpendicular to the plane of the circle. In (14.42) we took into account the fact that the element dS in (14.41) is directed along the outward normal to the closed surface. Substituting (14.42) into (14.40), we obtain a formula identical to (14.39).

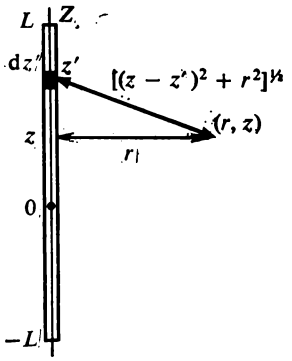


Fig. 41. A linear charge of a finite length

Example 14.3. *Find the potential and the strength of the field created in the space surrounding a uniformly charged filament of a finite length 2L. The linear charge density of the filament is* τ.

We place the origin of the Cartesian system of coordinates at the middle of the filament (point O) and direct the Z-axis along the filament (Fig. 41). In view of axial symmetry, the potential depends only on r and the coordinate z.

Figure 41 shows the plane passing through the point (r, z) and the Z-axis. The charge $\tau\, dz'$ located on the element of length dz' creates at the point (r, z) the potential

$$d\varphi = \frac{1}{4\pi\varepsilon_0} \frac{\tau\, dz'}{\sqrt{r^2 + (z - z')^2}}.$$

Consequently, the potential created by the entire charged filament is

$$\varphi = \frac{1}{4\pi\varepsilon_0} \int_{-L}^{L} \frac{\tau\, dz'}{\sqrt{r^2 + (z - z')^2}} = -\frac{\tau}{4\pi\varepsilon_0} \ln\left(\frac{z - L + \sqrt{r^2 + (z - L)^2}}{z + L + \sqrt{r^2 + (z + L)^2}}\right). \tag{14.43}$$

The components of electric field are given by the following formulas:

$$E_z = -\frac{\partial\varphi}{\partial z} = \frac{\tau}{4\pi\varepsilon_0}\left(\frac{1}{\sqrt{r^2 + (z - L)^2}} - \frac{1}{\sqrt{r^2 + (z + L)^2}}\right), \tag{14.44}$$

$$E_r = -\frac{\partial\varphi}{\partial r} = -\frac{\tau}{4\pi\varepsilon_0 r}\left(\frac{z - L}{\sqrt{r^2 + (z - L)^2}} - \frac{z + L}{\sqrt{r^2 + (z + L)^2}}\right). \tag{14.45}$$

For $L \to \infty$, we obtain

$$E_z = 0, \quad E_r = \tau/(2\pi\varepsilon_0 r).$$

As $L \to \infty$, the potential also tends to infinity:

$$\varphi = -\frac{\tau}{2\pi\varepsilon_0}[\ln r - \ln(2L)] \to \infty.$$

This is a consequence of the fact that the charge is not concentrated in a finite region of space, and hence formula (14.43) cannot be used for calculating the potential when $L \to \infty$.

www.EngineeringBooksPDF.com

For very large distances from the middle of the filament [($R = \sqrt{r^2 + z^2} \gg L$), formula (14.43) gives

$$\varphi = \frac{\tau 2L}{4\pi\varepsilon_0 R} = \frac{1}{4\pi\varepsilon_0}\frac{Q}{R},$$

where $Q = \tau 2L$ is the total charge of the filament. Thus, at distances that are large in comparison with the linear dimensions of the filament, the field is close to the Coulomb field.

Sec. 15. Electrostatic Field in Vacuum

The basic methods for calculating the potential and strength of an electrostatic field are described and examples of calculations are analysed.

Formulation of the problem. Let us solve the following problem of electrostatics:

Determine the electric field created by a given charge distribution.

This problem can be solved in several ways. All these methods are equivalent in principle, but different in practice, depending on the circumstances, since they involve different amounts of computational work. It is expedient to choose the method which leads to the required result in the simplest manner.

Direct application of Coulomb's law. In this case, the field strength at a point is determined as the sum of the fields created by all the elements $\rho\, dV$ and $\sigma\, dS$ of the volume and surface charges. This is the most natural method, though not the simplest one since it involves the addition of vectors which considerably complicates the calculations. An example of using this method was considered in Sec. 8 while calculating the forces of interaction between a point charge and a very long charged filament.

Calculation of potential. Formulas (14.35) and (14.36) can be used only for charge distribution in a finite region of space and when the potential is normalized to zero at infinity.

By way of an example, let us consider the field at the points on the perpendicular passing through the centre of a uniformly charged disc of radius a (Fig. 42). The total charge on the disc is equal to Q. For the potential at a distance h from the surface of the disc, we have [see (14.36)]

$$\varphi(h) = \frac{1}{4\pi\varepsilon_0}\int_S \frac{\sigma\, dx\, dy}{\sqrt{x^2+y^2+h^2}}, \tag{15.1}$$

where $\sigma = Q/(\pi a^2)$ is the surface charge density of the disc. It is convenient to calculate this integral in polar coordinates by putting $x^2 + y^2 = r^2$ and $dx\, dy = dS = r\, dr\, d\alpha$. This gives [see (15.1)]

$$\varphi(h) = \frac{\sigma}{4\pi\varepsilon_0}\int_0^{2\pi} d\alpha \int_0^a \frac{r\, dr}{\sqrt{r^2+h^2}} = \frac{1}{2\pi\varepsilon_0}\frac{Q}{a^2}\left(\sqrt{a^2+h^2} - h\right). \tag{15.2}$$

www.EngineeringBooksPDF.com

From the axial symmetry of charge distribution it follows that the electric field vector is directed along the axis of the disc and is equal to

$$E_h = -\frac{\partial \varphi}{\partial h} = \frac{1}{2\pi\varepsilon_0}\frac{Q}{a^2}\left(1 - \frac{h}{\sqrt{a^2+h^2}}\right). \tag{15.3}$$

For $h \gg a$, it can be assumed that

$$\frac{h}{\sqrt{a^2+h^2}} = \frac{1}{\sqrt{1+a^2/h^2}} \approx 1 - \frac{1}{2}\frac{a^2}{h^2} + \ldots. \tag{15.4}$$

Consequently,

$$E_h \approx \frac{1}{4\pi\varepsilon_0}\frac{Q}{h^2}, \tag{15.5}$$

as could be expected even without calculations, since at large distances the field due to a charged body is equal to the field due to a point charge.

Application of Gauss' theorem. In the presence of symmetry, Gauss' theorem is sometimes found to be the most effective means for determining the field strength.

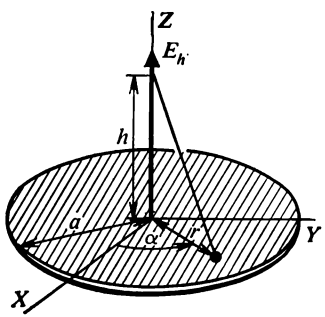


Fig. 42. Field along the axis of a uniformly charged disc

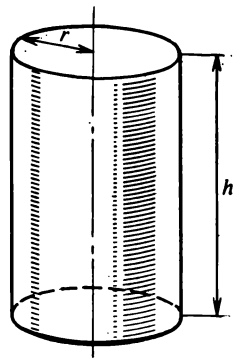


Fig. 43. To the calculation of the electric field of a very long charged filament with the help of the Gauss theorem

For example, suppose that it is required to find the field strength due to a very long straight charged filament with linear density τ. We construct a right circular cylinder of radius r, whose axis coincides with the filament (Fig. 43). We denote the height of the cylinder by h. Applying the Gauss theorem to the volume of the cylinder, we get

$$\int_S \mathbf{E}\, d\mathbf{S} = Q/\varepsilon_0 \tag{15.6}$$

where Q is the charge in the volume of the cylinder and S is the cylinder surface. Obviously, $Q = \tau h$. The flux of $\mathbf{E}$ through the bases of the cylinder is equal to zero, since the vector $\mathbf{E}$ is parallel to the bases. The flux of $\mathbf{E}$ through the lateral surface can be easily calculated, since in this case the vector $\mathbf{E}$ coincides with

www.EngineeringBooksPDF.com

the normal to the surface, and its absolute value is constant. This gives

$$\int_S \mathbf{E}\cdot d\mathbf{S} = \int_{S_{lat}} \mathbf{E}\cdot d\mathbf{S} = E\cdot 2\pi rh. \tag{15.7}$$

Thus, Gauss' theorem leads to the equality

$$E\cdot 2\pi rh = \tau h/\varepsilon_0, \tag{15.8}$$

whence

$$E = \frac{1}{2\pi\varepsilon_0}\frac{\tau}{r}. \tag{15.9}$$

In a field of this strength, the force acting on a point charge is equal to (8.5), which is obtained directly from the Coulomb law.

Laplace's equation and Poisson's equation. In order to find the field strength it is preferable in many cases to reduce the problem to the solution of the differential equation for potential. In order to obtain this equation, we substitute into

$$\operatorname{div}\mathbf{E} = \rho/\varepsilon_0 \tag{15.10}$$

the expression

$$\mathbf{E} = -\operatorname{grad}\varphi. \tag{15.11}$$

This gives

$$\operatorname{div}\operatorname{grad}\varphi = -\rho/\varepsilon_0. \tag{15.12}$$

We consider that

$$\operatorname{div}\operatorname{grad}\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} = \nabla^2\varphi, \tag{15.13}$$

where ∇^2 is the Laplace operator, equal to the sum of the second derivatives with respect to coordinates. Sometimes, this operator is denoted by $\Delta \equiv \nabla^2$. With the help of (15.13), we can write Eq. (15.12) in the form

$$\nabla^2\varphi = -\rho/\varepsilon_0. \tag{15.14}$$

This equation is called **Poisson's equation.** In those regions of space where charges do not exist ($\rho = 0$), this equation is reduced to the form

$$\nabla^2\varphi = 0 \tag{15.15}$$

called **Laplace's equation.**

After the potential φ has been determined as the solution of (15.14), we can calculate the electric field strength by using formula (15.11). The solution must satisfy the requirements formulated for the potential (see Sec. 14): *the potential φ is a continuous and finite function with finite derivatives with respect to coordinates.*

If all the charges are concentrated in a finite region of space, it follows from the uniqueness of the solution of problems of electromagnetism that (14.35) will be the solution of (15.14) (see Sec. 58).

www.EngineeringBooksPDF.com

The biggest advantage in finding the field strength with the help of Poisson's differential equation for potential is that this method is quite general and it can be widely applied. Formulas (14.35) and (14.36) assume that all charges are located in a finite region of space, and hence it is reasonable to normalize the potential to zero at infinity. *Poisson's equation, however, does not assume any definite normalization of the potential and the absence of charges at infinity.*

A very long uniformly charged circular cylinder. Using Poisson's equation, let us determine the potential created by a very long circular cylinder of radius a and the volume charge density $\rho = \text{const}$.

We direct the Z-axis along the axis of the cylinder. In view of the axial symmetry of the charge distribution, the potential φ is also axially symmetric, i.e. $\varphi = \varphi(r)$. It is therefore convenient to use a cylindrical system of coordinates, whose axial angle is denoted by α. In this system of coordinates, the Laplace operator has the form

$$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial r^2} + \frac{1}{r}\frac{\partial\varphi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial\alpha^2} + \frac{\partial^2\varphi}{\partial z^2}. \tag{15.16}$$

Since in this case the potential φ depends only on r, Eq. (15.16) can be simplified as follows:

$$\nabla^2\varphi = \frac{d^2\varphi}{dr^2} + \frac{1}{r}\frac{d\varphi}{dr} = \frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi}{dr}\right), \tag{15.17}$$

while Poisson's equation (15.14) can be written in the form:

$$\begin{aligned} &\frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi_1}{dr}\right) = -\rho/\varepsilon_0 \quad (0<r<a), \\ &\frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi_2}{dr}\right) = 0 \qquad (r>a). \end{aligned} \tag{15.18}$$

The general solutions of this equation are obtained by integration:

$$\begin{aligned} \varphi_1 &= -\frac{1}{4}\frac{\rho}{\varepsilon_0}r^2 + A_1\ln r + B_1, \\ \varphi_2 &= A_2\ln r + B_2, \end{aligned} \tag{15.19}$$

where A_1, A_2, B_1 and B_2 are integration constants. Since the potential must be finite at all points in space, and $\ln r \to \infty$ as $r \to 0$, we must put $A_1 = 0$ in the solution (15.19). It is convenient to normalize the potential by the condition $\varphi_1(0) = 0$, which gives $B_1 = 0$.

Since there are no surface charges, the electric field strength on the surface of a sphere is continuous. In other words, the derivative of the potential is continuous. The continuity conditions for the potential and its derivative at $r = a$ give two algebraic equations for determining the two remaining unknown constants A_2 and B_2:

$$A_2\ln a + B_2 = -\frac{1}{4}\frac{\rho}{\varepsilon_0}a^2, \quad \frac{A_2}{a} = -\frac{1}{2}\frac{\rho}{\varepsilon_0}a \tag{15.20}$$

www.EngineeringBooksPDF.com

Consequently,

$$\varphi_1(r) = -\frac{1}{4}\frac{\rho}{\varepsilon_0} r^2 \qquad (0 < r \leqslant a),$$
$$\varphi_2(r) = \frac{1}{2}\frac{\rho}{\varepsilon_0} a^2 \ln\frac{a}{r} - \frac{1}{4}\frac{\rho}{\varepsilon_0} a^2 \quad (r \geqslant a). \tag{15.21}$$

This gives

$$E_r = -\frac{\partial\varphi_1}{\partial r} = \frac{1}{2}\frac{\rho}{\varepsilon_0} r \qquad (0 < r \leqslant a),$$
$$E_r = -\frac{\partial\varphi_2}{\partial r} = \frac{1}{2}\frac{\rho}{\varepsilon_0}\frac{a^2}{r} \quad (r \geqslant a). \tag{15.22}$$

Considering that $\rho\pi a^2 = \tau$ is the charge per meter length of the cylinder, the second equation in (15.22) can be rewritten in the form

$$E_r = \frac{1}{2\pi\varepsilon_0}\frac{\tau}{r}. \tag{15.23}$$

A comparison of (15.23) and (15.9) shows that the field outside a uniformly charged cylinder is the same as if the entire charge were concentrated on the axis.

Direct application of Coulomb's law for calculating the electric field strength from a given charge distribution is the most natural, although not the simplest approach.
For a symmetric charge distribution, it is usually expedient to determine the electric field strength with the help of the Gauss theorem.

What is the physical meaning of the potential within the framework of electrostatics?
What is the physical meaning of the potential difference?

Example 15.1. *Find the field strength for a uniformly charged finite straight filament with linear charge density* τ (Fig. 44). *The parameters have the following values*: $\tau = 10^{-10}$ C/m, $l = 1$ m, $d = 0.5$ m *and* $a = 0.5$ m.

According to Coulomb's law, we have

$$dE_x = \frac{\tau\, dy \cos\alpha}{4\pi\varepsilon_0 (y^2+d^2)} = \frac{d\tau\, dy}{4\pi\varepsilon_0 (y^2+d^2)^{3/2}},$$
$$dE_y = \frac{\tau\, dy \sin\alpha}{4\pi\varepsilon_0 (y^2+d^2)} = -\frac{\tau y\, dy}{4\pi\varepsilon_0 (y^2+d^2)^{3/2}},$$

whence

$$E_x = \frac{\tau d}{4\pi\varepsilon_0}\int_{-(l-a)}^{a} \frac{dy}{(y^2+d^2)^{3/2}}, \qquad E_y = -\frac{\tau}{4\pi\varepsilon_0}\int_{-(l-a)}^{a} \frac{y\, dy}{(y^2+d^2)^{3/2}}.$$

Carrying out the substitution of variables $y = d\tan\alpha$, $dy = d\, d\alpha/\cos^2\alpha$, $1 + \tan^2\alpha = 1/\cos^2\alpha$, and evaluating the integrals, we get

$$E_x = \frac{\tau}{4\pi\varepsilon_0 d}(\sin\alpha_2 + \sin\alpha_1) = 1.27 \text{ V/m},$$
$$E_y = \frac{\tau}{4\pi\varepsilon_0 d}(\cos\alpha_2 - \cos\alpha_1) = 0. \tag{15.24}$$

www.EngineeringBooksPDF.com

For a very long filament ($l \to \infty$), $\alpha_1 = \alpha_2 = \pi/2$, and hence $E_y = 0$ and $E_x = \tau/(2\pi\varepsilon_0 d)$.

Example 15.2. *With the help of the potential, find the field strength at points on the perpendicular to the plane of a disc on which a charge Q is uniformly distributed. The radius of the disc is equal*

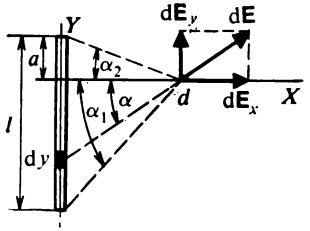


Fig. 44. To the calculation of the electric field of a linear charge of a finite length

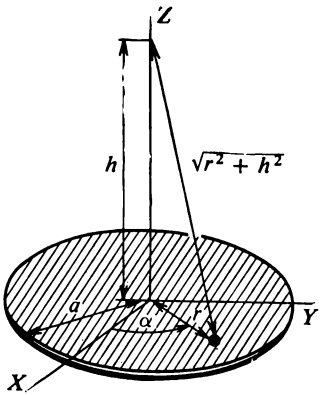


Fig. 45. To the calculation of the electric field of a charged disc

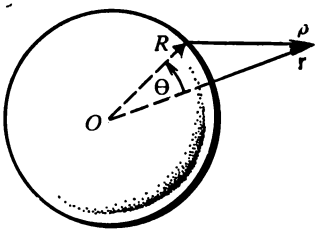


Fig. 46. To the calculation of the electric field of the surface charge of a sphere

to a (Fig. 45), *and it is assumed that* $Q = 10^{-10}$ C, $a = 10$ cm, $h = 20$ cm (*the distance between a point and the plane of the disc*).

In accordance with formula (14.36), we have

$$\varphi(h) = \frac{1}{4\pi\varepsilon_0} \int\limits_S \frac{\sigma\, dx\, dy}{\sqrt{x^2+y^2+h^2}}, \qquad \sigma = \frac{Q}{\pi a^2}.$$

In order to calculate the integral, we go over to polar coordinates in the plane of the disc: $x^2 + y^2 = r^2$, $dx\, dy = r\, dr\, d\alpha$. This gives

$$\varphi(h) = \frac{\sigma}{4\pi\varepsilon_0} \int\limits_0^{2\pi} d\alpha \int\limits_0^{a} \frac{r\, dr}{\sqrt{r^2+h^2}} = \frac{1}{2\pi\varepsilon_0}\frac{Q}{a^2}\left(\sqrt{a^2+h^2}-h\right), \tag{15.25}$$

whence

$$E_h = -\frac{\partial\varphi}{\partial h} = \frac{1}{2\pi\varepsilon_0}\frac{Q}{a^2}\left(1-\frac{h}{\sqrt{a^2+h^2}}\right) = 18\ \text{V/m}. \tag{15.26}$$

This formula is similar to (15.3).

Example 15.3. *Find the strength of the electric field created by the surface charge of a sphere of radius R. The total charge on the sphere is Q, and the surface charge density is* $\sigma = Q/(4\pi R^2)$.

The potential created by a charged surface element at a point characterized by r (Fig. 46) is

$$d\varphi = \frac{1}{4\pi\varepsilon_0}\frac{\sigma R^2 \sin\theta\, d\theta\, d\alpha}{\rho}, \tag{15.27}$$

where $R^2 \sin\theta\, d\theta\, d\alpha$ is a surface element of the sphere in spherical coordinates whose polar axis coincides with the vector r, and α is the axial angle. It can be seen from the figure that $\rho = R - r$. Squaring both sides of this equation, we get $\rho^2 = R^2 + r^2 - 2Rr\cos\theta$. Differentiating both sides, we obtain

$$2\rho\, d\rho = 2Rr \sin\theta\, d\theta,$$

www.EngineeringBooksPDF.com

whence $R^2 \sin\theta\, d\theta = (\rho R/r)\, d\rho$. In this case [see (15.27)]

$$d\varphi = \frac{1}{4\pi\varepsilon_0} \frac{\sigma R}{r}\, d\rho\, d\alpha. \tag{15.28}$$

Integrating this equation over the entire surface of the sphere, we get

$$\varphi = \frac{1}{4\pi\varepsilon_0} \frac{\sigma R}{r} \int_0^{2\pi} d\alpha \int_{|r-R|}^{r+R} d\rho = \frac{1}{2} \frac{\sigma R}{r} [\rho]_{|r-R|}^{r+R} = \begin{cases} \dfrac{\sigma R^2}{\varepsilon_0 r} = \dfrac{1}{4\pi\varepsilon_0} \dfrac{Q}{r} & (r > R), \\ \dfrac{\sigma R}{\varepsilon_0} = \dfrac{1}{4\pi\varepsilon_0} \dfrac{Q}{R} & (r < R). \end{cases} \tag{15.29}$$

This leads to the following expression for the electric field strength:

$$E_r = -\frac{\partial\varphi}{\partial r} = \begin{cases} \dfrac{1}{4\pi\varepsilon_0} \dfrac{Q}{r^2} & (r > R), \\ 0 & (r < R), \end{cases}$$

i.e. the field strength outside a uniformly charged sphere is the same as *if the entire charge were concentrated at its centre. There is no field inside the sphere.*

Sec. 16. Electrostatic Field in the Presence of Conductors

The effect of conductors on an electric field is considered. Basic physical phenomena due to charge distribution over the surface of a conductor (such as charge leakage from a point) are described. Quantitative characteristics of electrical properties of solitary conductors and systems of conductors are analysed. The essence of the image method is discussed.

Differential form of Ohm's law. Conductors are material bodies in which the motion of charges, i.e. electric current, appears due to an electric field. The law connecting the current created in a conductor when a potential difference is applied across its ends was discovered experimentally in 1827 by Ohm (1787-1854). This law has the form

$$I = U/R, \tag{16.1}$$

where R is called the resistance of the conductor. *In differential form, Ohm's law is obtained if we write this relation for the current density.* Let us consider a very small element of a conductor (Fig. 47; Δl is the length and ΔS is the cross-sectional area of the conductor and $\Delta\varphi$ is the potential difference across its ends). Let γ be the electric conductivity of the substance, viz. the reciprocal of the electric resistivity. The electric resistance of the conductor element and the current in it are respectively given by

$$R = \frac{1}{\gamma} \frac{\Delta l}{\Delta S}, \quad \text{(16.2a)} \qquad I_\tau = j_\tau \Delta S, \tag{16.2b}$$

www.EngineeringBooksPDF.com

where the subscript τ indicates that the current density component is taken along the conductor element. For this element, Ohm's law is written in the form

$$\Delta\varphi = j_\tau \Delta S \frac{1}{\gamma} \frac{\Delta l}{\Delta S} \tag{16.3}$$

Considering that $(\Delta\varphi/\Delta l) = E_\tau$ is the electric field component along the element under investigation, we obtain from (16.3)

$$j_\tau = \gamma E_\tau. \tag{16.4}$$

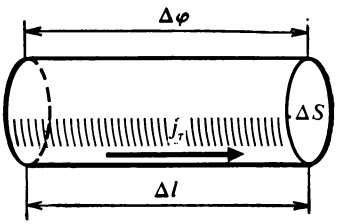


Fig. 47. To the derivation of Ohm's law in differential form

This relation is valid for any orientation of the conductor element, and hence can be written in vector form:

$$\boxed{\mathbf{j} = \gamma \mathbf{E}.} \tag{16.5}$$

This equation is called the differential form of Ohm's law.

Classification of materials according to conductivity. Electric conductivity γ depends on the properties of a material. Depending on conductivity, the materials are divided into three classes: dielectrics, semiconductors, and conductors. There is no sharp boundary between them. According to their conductivity, these materials are specified as follows.

(a) Dielectrics are substances with low electric conductivity. An ideal dielectric is characterized by the absence of conduction. This, however, can be observed only at 0 K. At temperatures other than 0 K, all materials have a certain conductivity, and hence ideal dielectrics do not exist. Dielectrics are the materials whose electric conductivity $\gamma < 10^{-5}$ S/m.

(b) Semiconductors have electric conductivity between 10^{-5} and 10^3 S/m.

(c) Conductors are characterized by an electric conductivity higher than 10^3 S/m. Mainly, these are metals. The best conductors among them are copper and silver which have electric conductivity of the order of 10^7 S/m.

Absence of electric field inside a conductor. In electrostatics, we consider the case when charges are fixed, i.e. $\mathbf{j} = 0$. Equation (16.5) gives for this case

$$\mathbf{E} = 0, \tag{16.6}$$

i.e. *there is no electric field inside a conductor in electrostatic equilibrium.*

Absence of volume charges inside a conductor. The equation

$$\operatorname{div} \mathbf{E} = \rho/\varepsilon_0 \tag{16.7}$$

for $\mathbf{E} = 0$ gives

$$\rho = 0, \tag{16.8}$$

i.e. *there are no volume charges inside a conductor.* This means that the charge of the conductor is concentrated on its surface in a layer of atomic thickness. Of course, *both positive and negative charges exist inside the conductor, but they compensate each other, and the interior of the conductor is neutral on the whole* [see (16.8)].

www.EngineeringBooksPDF.com

The neutrality is established very quickly. Let us suppose that in a certain volume inside a conductor the density of free charges differs from zero ($\rho(0) \neq 0$) at the instant $t = 0$. The continuity equation (5.24) combined with (16.5) assumes the form

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\gamma \mathbf{E}) = \frac{\partial \rho}{\partial t} + \gamma \operatorname{div} \mathbf{E} = 0,$$

where $\gamma = \text{const}$ (for a homogeneous conductor). Taking into account (16.7), we obtain from this expression the equation for time variation of ρ:

$$\frac{\partial \rho}{\partial t} = -\frac{\gamma}{\varepsilon_0} \rho.$$

The solution of this equation has the form

$$\rho(t) = \rho(0)\, e^{-(\gamma/\varepsilon_0)t},$$

i.e. the charge density decreases exponentially. In accordance with the general rule, we can assume that the *space charge in the conductor "is assimilated" during the time* $\tau = \varepsilon_0/\gamma$, which is called the **relaxation time**. For metals, this time is extremely short. For example, for copper ($\gamma = 6 \times 10^7$ S/m), $\tau \simeq 10^{-19}$ s. This interval of time is extremely short even on the scale of intra-atomic processes. Consequently, in nonstationary cases, when fields vary with time, we can assume with a high degree of accuracy that for moderate frequencies *free charges in a conductor are distributed over its surface and space charges are absent.* This conclusion remains valid even if we take into account the dependence of the conductivity γ on frequency, although in this case the relaxation time increases by several orders of magnitude.

The establishment of neutrality is associated with currents which, however, do not create a charge in the regions through which they flow. Let us illustrate this by a simple example. Suppose we have a sphere of radius a_2, made of a material with the permittivity ε and electric conductivity γ. At the initial moment $t = 0$, the spherical region of radius $a_1 < a_2$ is uniformly charged with the charge density ρ_0. The spherical layer between radii a_1 and a_2 is neutral. Let us consider the process of charge neutralization in the volume of the sphere.

The time variation of charge density at different points of the sphere is given by the formula

$$\rho(r, t) = \begin{cases} \rho_0 e^{-t/\tau} & (r < a_1), \\ 0 & (r > a_1), \end{cases}$$

where $\tau = \varepsilon/\gamma$. The total charge of the sphere $Q_0 = (4/3)\pi a_1^3 \rho_0$ remains constant, while the charge of the spherical region of radius a_1 decreases in accordance with the law

$$Q_1(t) = \frac{4}{3}\pi a_1^3 \rho_0 e^{-t/\tau} = Q_0 e^{-t/\tau}.$$

Conduction current carries this charge through a spherical layer between radii a_1 and a_2 to the surface of the sphere, where it is concentrated in the form of a surface charge.

www.EngineeringBooksPDF.com

At any instant of time, the charge distribution is spherically symmetric, and hence the Gauss theorem gives the following expression for the electric field strength:

$$E_r = \begin{cases} \frac{Q_0 e^{-t/\tau} r}{4\pi\varepsilon a_1^3} & (0 < r < a_1) \\ \frac{Q_0 e^{-t/\tau}}{4\pi\varepsilon r^2} & (a_1 < r < a_2), \\ \frac{Q_0}{4\pi\varepsilon_0 r^2} & (r > a_2). \end{cases}$$

The surface charge of the sphere increases. It can be calculated with the help of the law of charge conservation or from the boundary conditions. In the former case, we obtain

$$\sigma = \frac{1}{4\pi a_2^2}[Q_0 - Q_1(t)] = \frac{Q_0}{4\pi a_2^2}(1 - e^{-t/\tau}).$$

In the latter case,

$$\sigma|_{r=a_2} = D_r|_{r=a_2+0} - D_r|_{r=a_2-0} = \varepsilon_0 E_r|_{r=a_2+0} - \varepsilon E_r|_{r=a_2-0} = \frac{Q_0}{4\pi a_2^2}(1 - e^{-t/\tau}),$$

where the values of the function with the arguments $r = a_2 + 0$ and $r = a_2 - 0$ are taken from the inside and outside of the surface of the sphere.

The conduction current density is given by

$$j_r = \gamma E_r = \begin{cases} \frac{\gamma Q_0 e^{-t/\tau} r}{4\pi\varepsilon a_1^3} & (0 < r < a_1), \\ \frac{\gamma Q_0 e^{-t/\tau}}{4\pi\varepsilon r^2} & (a_1 < r < a_2), \\ 0 & (a_2 < r < \infty). \end{cases}$$

The conduction current through a spherical surface of radius r is determined from the formula

$$I_r = j_r 4\pi r^2 = \begin{cases} \frac{\gamma Q_0 e^{-t/\tau}}{\varepsilon} \frac{r^3}{a_1^3} & (0 < r < a_1), \\ \frac{\gamma Q_0 e^{-t/\tau}}{\varepsilon} & (a_1 < r < a_2), \\ 0 & (a_2 < r < \infty). \end{cases}$$

Thus, the total current in the region $0 < r < a_1$ increases with the radius. This is due to the fact that each point in this volume is a source of conduction current. In the region $a_1 < r < a_2$ there are no sources of conduction current, and hence the total current through the spherical surface does not depend on the radius.

Electrostatic induction. If a neutral conductor is placed in an external electric field, the surface charges are redistributed over its surface in such a way that

www.EngineeringBooksPDF.com

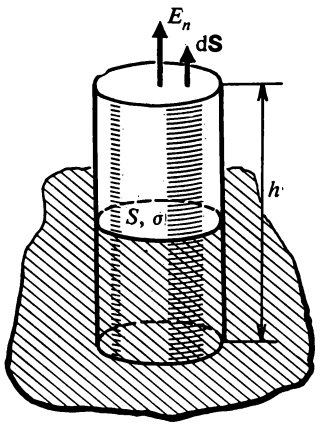


Fig. 48. To the derivation of the formula for the normal component of the electric field vector near the surface of a conductor

the field created by them inside the conductor completely compensates the external field. As a result, the total field strength inside the conductor is equal to zero. *The redistribution of surface charges on a conductor placed in an external electric field is called the electrostatic induction.* If the conductor is charged, its charges are also redistributed under the action of the external field.

The field near the surface of a conductor. Let us isolate an area element ΔS on the surface of a conductor and construct a right cylinder of height h so that it crosses the surface (Fig. 48). We apply the Gauss theorem to this cylinder:

$$\int_S \mathbf{E} \cdot d\mathbf{S} = Q/\varepsilon_{0\varepsilon} \tag{16.9}$$

where S is the surface of the cylinder and Q is the charge in the cylinder volume.

Inside the cylinder, the charge exists only on the surface of the conductor and is characterized by the surface charge density σ. Hence $Q = \sigma S$. The field inside the conductor is equal to zero, and hence the flux of $\mathbf{E}$ through the part of the cylinder surface located in the volume of the conductor is equal to zero. The flux through the part of the cylinder lying outside the conductor is the sum of the fluxes through the cylinder base and the lateral surface. In the limit, we take the height h of the cylinder as small as desired ($h \to 0$), and hence the area of the lateral surface of the cylinder and the flux of $\mathbf{E}$ through it will be as small as desired. Therefore, in the limit $h \to 0$, only the flux through the cylinder base is left:

$$\int_{\Delta S} \mathbf{E} \cdot d\mathbf{S} = E_n \Delta S, \tag{16.10}$$

where E_n is the normal component of $\mathbf{E}$. It should be recalled that in the Gauss theorem, the positive direction of the normal is that of the outward normal to the closed surface. In the case under consideration, this means that the positive normal is directed outwards from the conductor surface. As $h \to 0$, expression (16.9) combined with (16.10) becomes

$$E_n \Delta S = \sigma \Delta S/\varepsilon_0, \tag{16.11}$$

whence

$$E_n = \sigma/\varepsilon_0. \tag{16.12}$$

Thus, *the normal component of the field strength at the surface of a conductor is uniquely determined by the surface charge density.*

Let us consider now the tangential component of the field strength. We shall show that it should be equal to zero proceeding from the fact that a perpetual

www.EngineeringBooksPDF.com

motion machine cannot exist. We consider a closed loop L crossing the surface of a conductor so that its upper part is parallel to the surface of the conductor outside it, while the lower part is inside the conductor (Fig. 49). The electric field strength **E** is equal to zero inside the conductor, and hence the tangential field component is absent. Suppose that outside the conductor the tangential component is not equal to zero. We take a positive charge and move it along the closed loop in the direction shown in Fig. 49 by arrows. On the section AB, the field performs a positive work. The section BC can be made in the limit as small as desired since the sections AB and CD can be arranged as close to the conductor surface as desired. Consequently, the motion on the section BC is associated with the work which can be as small as desired. The motion of the

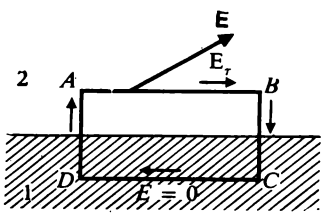


Fig. 49. To the proof of the absence of the tangential component of the electric field outside a conductor

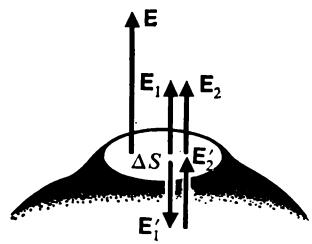


Fig. 50. Emergence of an electric field near the surface of a conductor

charge in the section CD does not involve any work since the field inside the conductor is absent. The work associated with the motion of the charge along the section DA, as in the case of the section BC, can be as small as desired. Thus, as a result of motion of the charge along the closed loop, the electric field performs a positive work, and no changes occur in the system. We can repeat this cycle and obtain the same work, and so on. Thus, we have realized a perpetual motion machine of the first kind, which is impossible. This perpetual motion machine performs work at the expense of the tangential component of the electric field near the surface of the conductor. Hence, this component must be equal to zero. In other words, *the fact that the tangential component of the electric field near the surface of a conductor is equal to zero is a consequence of the potential nature of electrostatic field and of the absence of the field inside the conductor.*

The equality

$$E_\tau = 0 \tag{16.13}$$

indicates that *the electric field strength near the surface of a conductor is perpendicular to the surface and is equal to* σ/ε_0 [see (16.12)].

Mechanism of creation of the field near the surface of a conductor. Electric charges are the only source of electric field in electrostatics. Hence the field in the vicinity of the surface of a conductor is created by all surface charges of this

www.EngineeringBooksPDF.com

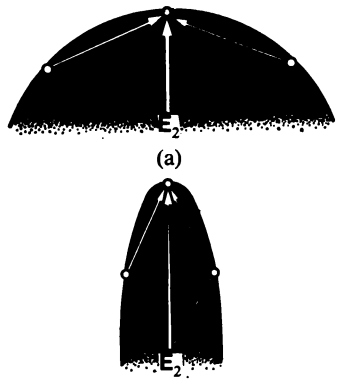


Fig. 51. Dependence of the surface charge density on the curvature of the surface

conductor and all charges outside it. Let us isolate a very small area element ΔS of the conductor surface (Fig. 50). The field strength $\mathbf{E}$ near the surface of the conductor is the sum of two components, viz. the strength of the field $\mathbf{E}_1$ created by the charges contained in the element ΔS and the field $\mathbf{E}_2$ created by the remaining charges outside the element ΔS. Clearly, the charges of the area element ΔS create the field on both sides of the element. Since both sides of the element ΔS are equivalent, we may conclude that the vectors $\mathbf{E}_1$ and $\mathbf{E}_1'$ are oppositely directed and equal in magnitude: $|\mathbf{E}_1| = |\mathbf{E}_1'|$. The field $\mathbf{E}_2$ is created by all charges located outside the element ΔS. Obviously, these charges create not only the field $\mathbf{E}_2$ outside the conductor, but also a field $\mathbf{E}_2'$ inside the conductor. Since this field is in the space outside the charges and is created by them, it must be continuous, and hence $\mathbf{E}_2 = \mathbf{E}_2'$. The total field strength inside the conductor is equal to zero, i.e. $\mathbf{E}' = \mathbf{E}_1' + \mathbf{E}_2' = 0$. Hence it follows that $\mathbf{E}_1' = -\mathbf{E}_2'$. Taking into account the equality $|\mathbf{E}_1| = |\mathbf{E}_1'|$, we conclude that

$$|\mathbf{E}_1| = |\mathbf{E}_2|.$$

Hence it follows that

$$\mathbf{E}_1 = \mathbf{E}_2 = \frac{1}{2}\mathbf{E}, \tag{16.14}$$

i.e. *the field near the surface of a conductor is the sum of two equal components, one of which is created by the surface charges of the adjoining surface element, while the other is created by all the remaining charges located outside this surface element.*

Dependence of the surface charge density on the curvature of the surface. The charge on the surface of a conductor is distributed nonuniformly, and the surface charge density depends on the curvature of the surface. In order to verify this, let us analyse the distribution of the field strength near a certain element of the surface (Fig. 51). If the surface is slightly curved (Fig. 51*a*), the charges lying outside dS create a small normal component of the field $\mathbf{E}_2'$ near this element. Therefore, to compensate the normal component, the charges located on this area element must create a comparatively weak field $\mathbf{E}_1' = -\mathbf{E}_2'$. In accordance with formulas (16.14) and (16.12), we conclude that the surface charge density $\sigma = 2\varepsilon_0 E_1'$ on this element must be comparatively low. On the other hand, if the curvature of the surface near the element under consideration is large, the field $\mathbf{E}_2'$ created by the charges located outside the area element dS is strong, and accordingly the field created by the charges located on the area element must be considerably stronger. This means that the surface charge density on this element must be higher. Thus, we can conclude that the *surface charge*

www.EngineeringBooksPDF.com

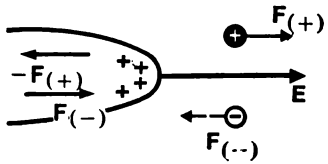


Fig. 52. Charge leakage from a tip

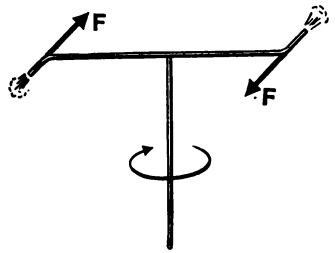


Fig. 53. Electrical "Segner" wheel

density increases with the curvature of the surface, i.e. increases with decreasing radius of curvature.

Similar arguments can be used to show that *the surface charge density on the concave surface of a conductor is lower in comparison with the flat surface.*

An increase in the surface charge density on convex surfaces is manifested most visually in the leakage of charge from a tip.

Charge leakage from a tip. Let us analyse the phenomenon occurring near the tip of a charged conductor (Fig. 52). The field E near this point is very strong. The surrounding air contains charges (ions and electrons) which are acted upon by a force in the field E. In accordance with Newton's third law, an equal and opposite force acts on the charges on the tip. Consequently, as a result of interaction, the charges in the air in the vicinity of the tip and the tip itself receive equal and opposite momenta. The charges in the air, which move towards the point under the action of the force, transfer their momentum and charge to the point upon impingement. This momentum is equal in magnitude to the momentum received by the point as a result of interaction with the corresponding charge, but is opposite in direction. Consequently, as the point is hit by the charges, these momenta compensate each other, and the net result of interaction is equal to zero.

Thus, *the interaction between the charges on the tip with unlike charges in the surrounding air does not lead to a force acting on the point.*

A different situation arises for like charges: the force acting on the charges at the tip is always directed inside the conductor (in Fig. 52, this force is denoted by $-\mathbf{F}_{(+)}$). If the tip is positively charged, negative charges hitting the tip (as shown in Fig. 52) neutralize the corresponding positive charges. This looks as if positive charges leave the tip or, as is usually said, leak from the tip. The force $-\mathbf{F}_{(+)}$ acting in this case on the tip is equivalent to the reactive force due to the leakage of charges from the tip. If the tip is negatively charged, the electrons in fact leave it, i.e. actually leak from the tip. The mechanism of appearance of "reactive force" in this case is completely identical to that described above.

This means that the "reactive force" appears not only at the moment the electrons start leaking from the surface of the conductor, but at all subsequent

www.EngineeringBooksPDF.com

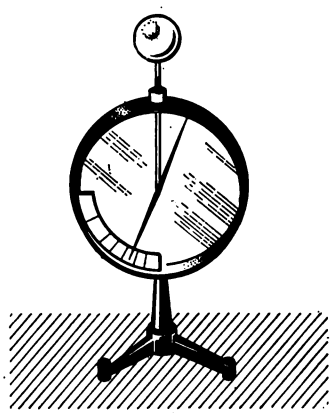


Fig. 54. Schematic diagram of an electrometer

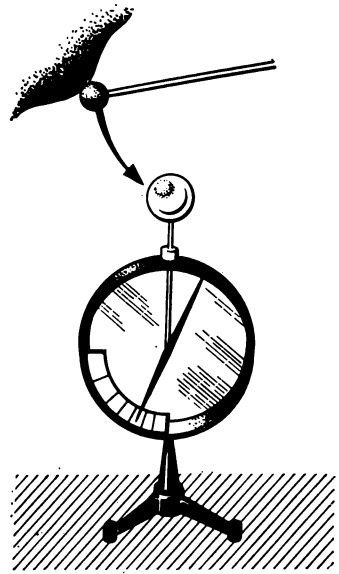


Fig. 55. Demonstration of the surface charge density dependence on the curvature of the surface with the help of an electrometer

instants of time, when an electron is accelerated by the field of charges left on the point.

An effective demonstration of the appearance of "reactive force" as a result of leakage of charges from a tip is the rotation of the electrical Segner wheel (Fig. 53). The dashed arrows show the direction of charge leakage, as a result of which a "reactive force" appears and the horizontal segment of the conductor is set into rapid rotation around the vertical axis.

Electroscopes and electrometers. The simplest device for detecting electric charges is a vertical metallic rod or plate with a light conducting foil or pointer attached to it at one end (Fig. 54). If there is no charge on the metallic rod and foil (pointer), the latter hangs vertically, parallel to the rod. In the presence of charge, the repulsive forces acting between like charges on the rod and on the foil (pointer) deflect the foil from the vertical position by a certain angle. Thus, the device may serve as an indicator of the presence of charge, i.e. an electroscope. The angle of deflection of the pointer from the vertical is the larger, the larger the charge on the rod. This makes it possible to graduate the electroscope and to determine the amount of electricity on it by the angle of deflection. Such an electroscope, adapted for quantitative measurement of charge is called an electrometer. The charge depends on the potential of the rod and of the

www.EngineeringBooksPDF.com

pointer. Consequently, an electrometer can be used for measuring the potential difference. The device is enclosed in a case (Fig. 54).

The dependence of the surface charge density on the curvature of the conductor surface can be illustrated with the help of an electrometer as follows. A small conducting ball fixed on a nonconducting handle touches a part of the surface of a conductor (Fig. 55). The charge on the ball is the larger, the higher the surface charge density on the part of the surface touched by the ball. After this, the ball is separated from the surface of the conductor and brought in contact with the rod of an electrometer. The charge transferred to the electrometer depends on the charge on the ball. Consequently, the deflection of the

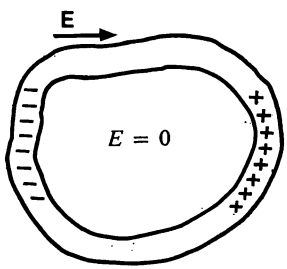


Fig. 56. A metallic screen for external fields

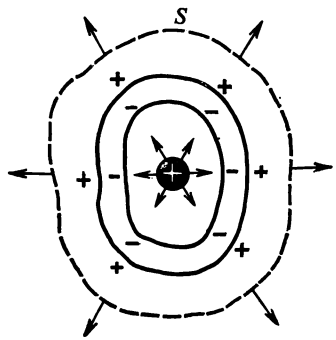


Fig. 57. A charge surrounded by a closed conducting shell

pointer indicates the surface charge density of the part of the surface from which the charge was transferred to the electrometer. The ratio of the surface charge densities of the corresponding parts of the surface of the conductor can be judged by the ratio of the angles of deflection of the pointer. Depending on the curvature of the surface, the surface charge density may vary significantly.

Metallic screen. The annihilation of the field inside a conductor by the charge distribution on its surface indicates that the inner parts of the conductor have nothing to do with the field and hence can be done away with. As a result, we are left with a closed conducting shell (Fig. 56). *In the space surrounded by the shell, the electric field is equal to zero. The closed shell is called a* **screen**. *It shields the internal space from the external electric field.* The screens (shields) are used for protecting technical devices from the influence of external electric fields. They are usually made from a mesh with small cells rather than from a solid conducting material. Experiments and calculations show that the screening effect of such a mesh is slightly inferior to that of a solid screen, but the material expenditures are much smaller and the construction is much simpler.

Does a closed conducting shell screen the external space from the charges located inside a cavity? In other words, does the field of charges in the volume surrounded by a closed conducting shell penetrate the surrounding space? Yes, it does. In order to verify this, we have to analyse the situation in greater detail.

www.EngineeringBooksPDF.com

Suppose that the charge

$$Q = \int_V \rho \, dV \tag{16.15}$$

is distributed in the volume V inside a cavity. In accordance with the law of electrostatic induction, an opposite charge appears on the inner surface of the shell (Fig. 57). In order to find its magnitude, we shall apply the Gauss theorem to the volume inside the closed shell:

$$\int_{S_{in}} \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\varepsilon_0} \int_V \rho \, dV, \tag{16.16}$$

where S_{in} is the inner surface of the shell.

Denoting by σ the surface charge density on the inner surface, we obtain the following expression for the field $\mathbf{E}$ near the surface [see (16.12)]:

$$\mathbf{E} = \frac{\sigma}{\varepsilon_0} \mathbf{n}, \tag{16.17}$$

where $\mathbf{n}$ is the normal to the inner surface of the shell, directed inside the volume bounded by the shell. We consider that $d\mathbf{S}$ in (16.16) is directed along the outward normal to the volume V, i.e. oppositely to $\mathbf{n}$, and hence

$$\mathbf{n} \cdot d\mathbf{S} = dS \cos(\widehat{\mathbf{n}, d\mathbf{S}}) = dS \cos \pi = -dS. \tag{16.18}$$

Using (16.17) and (16.18), we can write the integral on the left-hand side of (16.16) in the form

$$\int_{S_{in}} \mathbf{E} \cdot d\mathbf{S} = -\frac{1}{\varepsilon_0} \int_{S_{in}} \sigma \, dS. \tag{16.19}$$

Then the Gauss theorem (16.16) assumes the form

$$-\int_{S_{in}} \sigma \, dS = \int_V \rho \, dV = Q. \tag{16.20}$$

Consequently, *the charge formed on the inner surface of the shell is equal and opposite to the charge inside the cavity.*

The field strength inside the shell is equal to zero since the shell is a conductor. The charge on the outer surface has a sign opposite to that of the charge on the inner surface, its magnitude being equal, in accordance with the law of conservation of charge, to the magnitude of the charge on the inner surface.

In order to prove the existence of an electric field in the surrounding space, we shall use the Gauss theorem. The dashed line in Fig. 57 shows the closed surface surrounding the shell. The total charge in the volume bounded by this closed surface is equal to the charge inside the cavity bounded by the shell

www.EngineeringBooksPDF.com

since the charge of the shell is equal to zero. Consequently, the Gauss theorem has the form

$$\int_S \mathbf{E}\cdot d\mathbf{S} = \frac{1}{\varepsilon_0}\int_V \rho\, dV = Q/\varepsilon_0 \neq 0, \quad (16.21)$$

i.e. the field strength **E** in the space surrounding the shell differs from zero.

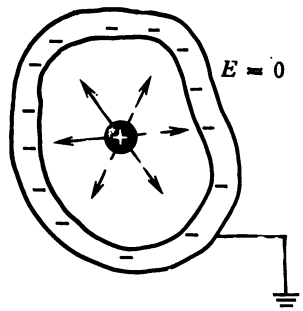


Fig. 58. An earthed closed shell shields the external space from the charges inside a volume

Let us "earth" the shell, i.e. connect it by a conductor with a very large remote conducting body. Usually, the Earth serves as such a body (Fig. 58). In order to simplify the analysis, we represent this body in the form of an infinite conducting medium that fills the entire space outside the shell and is in contact with it. All the charges will go from the outer surface of the shell to infinity, and only the charges inside the cavity and on the inner surface of the shell will remain. The field strength in the conducting medium surrounding the shell is equal to zero. In this case, the medium ensures the removal of the charge from the outer surface of the shell to infinity. Hence, at a finite distance from the shell a thin wire conductor will ensure the charge exchange between the shell and sufficiently remote regions of the medium. Obviously, after the removal of the conducting medium from the region surrounding the shell, the field strength at the points of this region is, as before, equal to zero. Thus, *the earthed closed shell shields the external space from the charges located in the volume surrounded by this shell. An unearthed shell does not provide such a screening.*

Potential of a conductor. The fact that the field strength **E** inside a conductor is equal to zero means that the potential at all points of the conductor has the same value, i.e. the potential difference between points *1* and *2* of the conductor [see (14.28)] is

$$\varphi(2) - \varphi(1) = \int_{(1)}^{(2)} \mathbf{E}\cdot d\mathbf{l} = 0. \quad (16.22)$$

The potential, whose value is the same at all points of a conductor, is called **the potential of the conductor.**

Suppose that we have an isolated charged conductor. In the space surrounding the conductor, there is an electric field created by the charge of the conductor. We shall normalize the potential to zero at infinity. Then [see (14.29)] the potential of the conductor can be expressed by the formula

$$\varphi = \int_{\substack{\text{(conductor}\\ \text{surface)}}}^{\infty} \mathbf{E}\cdot d\mathbf{l}. \quad (16.23)$$

In this formula, the integration path starts at any point of the conductor and terminates at infinity.

www.EngineeringBooksPDF.com

Capacitance of an isolated conductor. What determines the potential of an isolated conductor? It follows from formula (16.23) that in accordance with the principle of superposition, the potential must be proportional to the charge since **E** in the integrand of (16.23) is proportional to the charge. Further, it is clear that the potential depends on the size and shape of the conductor, which are taken into account by its capacitance.

The **capacitance** of a conductor is defined as the ratio of the charge Q of an isolated conductor to its potential φ:

$$\boxed{C = Q/\varphi.} \tag{16.24}$$

The capacitance of a conductor is measured in farads (F). Formula (16.24) gives

$$1 \text{ F} = 1 \text{ C/V}. \tag{16.25}$$

In the CGS system, the capacitance is expressed in centimetres and the formula for capacitance coincides with (16.24). Since 1 V = (1/300) CGS units, 1 C = 3×10^9 CGS units, it follows from (16.24) that

$$1 \text{ F} = 9 \times 10^{11} \text{ cm}. \tag{16.26}$$

A farad is a very large unit. Let us calculate, for example, the capacitance of a sphere of radius R carrying a charge Q. Since the strength of the electric field created by such a sphere in the surrounding space is

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0} \frac{Q}{r^2} \frac{\mathbf{r}}{r}, \tag{16.27}$$

the potential and capacitance are expressed by the formulas

$$\varphi = \int_R^\infty E \, dr = \frac{1}{4\pi\varepsilon_0} \frac{Q}{R}, \tag{16.28}$$

$$C = Q/\varphi = 4\pi\varepsilon_0 R. \tag{16.29}$$

For the radius of the sphere equal to 1 cm, we obtain

$$C = 10^{-2}/(9 \times 10^9) \simeq 10^{-12} \text{ F}. \tag{16.30}$$

For this reason, the capacitance is usually expressed in fractional units.

A system of conductors. If we have several conductors, the potential of each of them depends not only on the charge of the conductor but also on the strength of the fields created by other conductors or, in other words, on the charges of other conductors. In accordance with the principle of superposition, the potential is proportional to these charges.

Let us consider, for the sake of definiteness, two conductors (Fig. 59). It follows from what has been said above that

$$\varphi_1 = \alpha_{11}Q_1 + \alpha_{12}Q_2, \quad \varphi_2 = \alpha_{21}Q_1 + \alpha_{22}Q_2, \tag{16.31}$$

www.EngineeringBooksPDF.com

where α_{ij} are the **potential coefficients** which depend on the shape and size of the conductors and on their mutual arrangement. The theoretical calculation of these coefficients is a complicated mathematical problem. Usually, they are determined experimentally.



Fig. 59. A system of conductors

The potential coefficients are not independent of each other. This can be shown as follows. Let σ_1 and σ_2 be surface charge densities, r_{11} the distance from the surface element dS_1 of the first conductor to a certain fixed point inside it, and r_{12} the distance from the area element dS_2 of the second conductor to the same point. Then the potentials of the first and second conductors are given by

$$\varphi_1 = \frac{1}{4\pi\varepsilon_0}\int_{S_1}\frac{\sigma_1\,dS_1}{r_{11}} + \frac{1}{4\pi\varepsilon_0}\int_{S_2}\frac{\sigma_2\,dS_2}{r_{12}}, \tag{16.32}$$

$$\varphi_2 = \frac{1}{4\pi\varepsilon_0}\int_{S_2}\frac{\sigma_2\,dS_2}{r_{22}} + \frac{1}{4\pi\varepsilon_0}\int_{S_1}\frac{\sigma_1\,dS_1}{r_{21}}. \tag{16.33}$$

(the meaning of r_{22} and r_{21} is similar to that of r_{11} and r_{12}). The charges of the conductors are

$$Q_1 = \int_{S_1}\sigma_1\,dS_1, \quad Q_2 = \int_{S_2}\sigma_2\,dS_2. \tag{16.34}$$

Let us suppose that the charges of the conductors have changed:

$$Q_1' = \int_{S_1}\sigma_1'\,dS_1, \quad Q_2' = \int_{S_2}\sigma_2'\,dS_2. \tag{16.35}$$

We multiply both sides of (16.32) by Q_1' and (16.33) by Q_2' and add the obtained equalities termwise:

$$\begin{aligned}
Q_1'\varphi_1 + Q_2'\varphi_2 &= \frac{1}{4\pi\varepsilon_0}\int_{S_1}\sigma_1'\,dS_1\int_{S_1}\frac{\sigma_1\,dS_1}{r_{11}} + \frac{1}{4\pi\varepsilon_0}\int_{S_1}\sigma_1'\,dS_1\int_{S_2}\frac{\sigma_2\,dS_2}{r_{12}} \\
&+ \frac{1}{4\pi\varepsilon_0}\int_{S_2}\sigma_2'\,dS_2\int_{S_2}\frac{\sigma_1\,dS_2}{r_{22}} + \frac{1}{4\pi\varepsilon_0}\int_{S_2}\sigma_2'\,dS_2\int_{S_1}\frac{\sigma_1\,dS_1}{r_{21}} \\
&= \frac{1}{4\pi\varepsilon_0}\int_{S_1}\sigma_1\,dS_1\int_{S_1}\frac{\sigma_1'\,dS_1}{r_{11}} + \frac{1}{4\pi\varepsilon_0}\int_{S_2}\sigma_2\,dS_2\int_{S_1}\frac{\sigma_1'\,dS_1}{r_{12}} \\
&+ \frac{1}{4\pi\varepsilon_0}\int_{S_2}\sigma_2\,dS_2\int_{S_2}\frac{\sigma_2'\,dS_2}{r_{22}} + \frac{1}{4\pi\varepsilon_0}\int_{S_1}\sigma_1\,dS_1\int_{S_2}\frac{\sigma_2'\,dS_2}{r_{21}} \\
&= Q_1\varphi_1' + Q_2\varphi_2',
\end{aligned} \tag{16.36}$$

where the order of integration has been changed since the integration is carried out with respect to different independent variables. The quantities φ_1' and φ_2' are the potentials of the conductors when their charges are equal to Q_1' and Q_2'.

www.EngineeringBooksPDF.com

The relation

$$\boxed{Q_1'\varphi_1 + Q_2'\varphi_2 = Q_1\varphi_1' + Q_2\varphi_2'} \tag{16.37}$$

obtained in (16.36) is called the **reciprocity theorem**. From this theorem, we can obtain the condition that must be satisfied by the potential coefficients α_{ij}.

If the charge of the second conductor is equal to zero ($Q_2 = 0$, $Q_1 \neq 0$), then [see (16.31)]

$$\varphi_1 = \alpha_{11}Q_1, \quad \varphi_2 = \alpha_{21}Q_1, \tag{16.38}$$

If the charge of the first conductor is equal to zero ($Q_1' = 0$, $Q_j' \neq 0$), then [see (16.31)]

$$\varphi_1' = \alpha_{12}Q_2', \quad \varphi_2' = \alpha_{22}Q_2'. \tag{16.39}$$

The reciprocity theorem (16.37) for these two cases assumes the form

$$Q_2'\varphi_2 = Q_1\varphi_1' \tag{16.40}$$

Substituting into this expression the expressions for φ_2 and φ_1' [see (16.38) and ([16.39)] and cancelling the common factor $Q_2'Q_1$ on both sides of the obtained equality, we find

$$\alpha_{12} = \alpha_{21}, \tag{16.41}$$

i.e. *the potential coefficients are symmetric relative to their indices.*

These calculations can be easily extended for any number of conductors by writing the initial relations (16.31) for n conductors in the form

$$\varphi_i = \sum_{j=1}^{n} \alpha_{ij}Q_j \tag{16.42}$$

All further calculations are similar to (16.32)-(16.37) and lead to the following formula instead of (16.37), which expresses the reciprocity theorem in the general case:

$$\sum_{i=1}^{n} Q_i'\varphi_i = \sum_{i=1}^{n} Q_i\varphi_i'. \tag{16.43}$$

Instead of (16.41), we obain from this equation the general condition for the symmetry of potential coefficients:

$$\alpha_{ij} = \alpha_{ji}. \tag{16.44}$$

The system of equations (16.42) can be solved for Q_i:

$$Q_i = \sum_{j=1}^{n} C_{ij}\varphi_j. \tag{16.45}$$

Here $C_{ij} = A_{ij}/D$, where D is the determinant of the coefficients of the system of equations (16.42), A_{ij} being the complement of the element α_{ij} in this determinant. On the basis of (16.44) we conclude that the coefficients C_{ij} satisfy

www.EngineeringBooksPDF.com

the condition

$$C_{ij} = C_{ji}, \tag{16.46}$$

where C_{ij} are the capacitance coefficients, C_{ii} is the capacitance coefficient of the i-th conductor and C_{ij} the capacitance coefficient between the i-th and j-th conductors. The capacitance coefficient of an isolated conductor is called the capacitance of the conductor.

Since a positive charge creates in an isolated conductor a positive potential, we may conclude that *all capacitance coefficients with similar indices* (C_{11}, C_{22}, . . .) *are positive.* In order to verify this, let us earth all the conductors with the

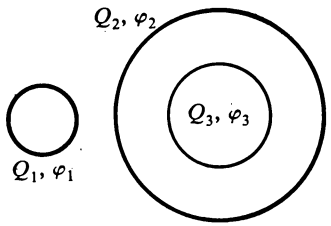


Fig. 60. To the determination of capacitance coefficients for the case of two spheres

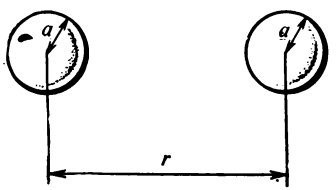


Fig. 61. To the calculation of capacitance coefficients for two conducting spheres

exception of the i-th conductor, retaining on it a positive charge. In other words, we shall assume that $Q_i > 0$. Obviously, in this case $\varphi_i > 0$ and $\varphi_j = 0$ for $j \neq i$. Consequently, Eq. (16.45) for Q assumes the form

$$Q_i = C_{ii}\varphi_i. \tag{16.47}$$

Since $\varphi_i > 0$ and $Q_i > 0$, $C_{ii} > 0$ as well, Q.E.D.

Similarly, we can prove that the *capacitance coefficients with unlike indices cannot be positive. They are either negative or equal to zero.* Let us consider, for example, two conductors one of which is earthed while the other is isolated and positively charged. On account of electrostatic induction, this positive charge will induce a negative charge on the earthed conductor. Formula (16.45) assumes the following form for the charge on the second conductor:

$$Q_2 = C_{21}\varphi_1. \tag{16.48}$$

Since $Q_2 < 0$ and $\varphi_1 > 0$, $C_{21} < 0$. This result does not exclude that the coefficient can be equal to zero, but by no means can this coefficient be positive.

Let us consider three conducting spheres (Fig. 60). We denote their potentials and charges by φ_1, φ_2, φ_3 and Q_1, Q_2, Q_3 respectively. For determining C_{ij}, we have Eqs. (16.45) which in this case have the form

$$\begin{aligned} Q_1 &= C_{11}\varphi_1 + C_{12}\varphi_2 + C_{13}\varphi_3, \\ Q_2 &= C_{21}\varphi_1 + C_{22}\varphi_2 + C_{23}\varphi_3, \\ Q_3 &= C_{31}\varphi_1 + C_{32}\varphi_2 + C_{33}\varphi_3. \end{aligned} \tag{16.49}$$

www.EngineeringBooksPDF.com

In order to calculate the coefficients C_{ij}, it is necessary to have a sufficient number of equations (16.49) with known Q_i's and φ_i' s, from which C_{ij} can be found.

Let us assume that $Q_3 = 0$ and that the second sphere is earthed. Then $\varphi_3 = \varphi_2 = 0$, and equations (16.49) become

$$Q_1 = C_{11}\varphi_1, \quad Q_2 = C_{21}\varphi_1, \quad 0 = C_{31}\varphi_1. \tag{16.50}$$

This gives $C_{31} = C_{13} = 0$, i.e. the capacitance coefficient (mutual capacitance) for screened conductors is equal to zero.

Suppose now that the first and second spheres are earthed, i.e. $\varphi_1 = 0$ and $\varphi_2 = 0$, while the charge $Q_3 \neq 0$. In this case, Eqs. (16.49) assume the form

$$Q_1 = 0, \quad Q_2 = C_{23}\varphi_3, \quad Q_3 = C_{33}\varphi_3. \tag{16.51}$$

It was shown above that the charge induced on the inner surface of an earthed conducting shell is equal in magnitude to the charge in the cavity bounded by the shell but has the opposite sign, i.e. $Q_2 = -Q_3$. Equations (16.51) give

$$C_{23} = -C_{33}. \tag{16.52}$$

Thus, the capacitance coefficient for a conductor which completely envelopes another conductor is equal and opposite to the capacitance coefficient of the inner conductor. This fact is very important for the theory of capacitors.

Suppose that we have two spheres at a distance r from each other, this distance being large in comparison with their radius a (Fig. 61) (we denote by r the distance between the centres of the spheres). Since $a \ll r$, we can ignore the redistribution of charges on the spheres due to mutual electrostatic induction while calculating the field strength at large distances from the spheres. Then the formulas for the potentials of the spheres become

$$\varphi_1 = \frac{1}{4\pi\varepsilon_0}\left(\frac{Q_1}{a} + \frac{Q_2}{r}\right), \quad \varphi_2 = \frac{1}{4\pi\varepsilon_0}\left(\frac{Q_1}{r} + \frac{Q_2}{a}\right), \tag{16.53}$$

where Q_1 and Q_2 are the charges on the first and second spheres. These equations can be solved for Q_1 and Q_2:

$$\begin{aligned} Q_1 &= 4\pi\varepsilon_0 \frac{ar^2}{r^2-a^2}\varphi_1 - 4\pi\varepsilon_0 \frac{a^2 r}{r^2-a^2}\varphi_2, \\ Q_2 &= -4\pi\varepsilon_0 \frac{ra^2}{r^2-a^2}\varphi_1 + 4\pi\varepsilon_0 \frac{r^2 a}{r^2-a^2}\varphi_2. \end{aligned} \tag{16.54}$$

Then

$$C_{11} = C_{22} = 4\pi\varepsilon_0 \frac{ar^2}{r^2-a^2} = C > 0, \tag{16.55}$$

$$C_{12} = C_{21} = -4\pi\varepsilon_0 \frac{ra^2}{r^2-a^2} = \gamma < 0. \tag{16.56}$$

Taking into account (16.55) and (16.56), let us represent (16.54) in the form

$$Q_1 = C\varphi_1 + \gamma\varphi_2, \quad Q_2 = \gamma\varphi_1 + C\varphi_2. \tag{16.57}$$

www.EngineeringBooksPDF.com

For $r \to \infty$, we obtain $C_{11} = C_{22} = 4\pi\varepsilon_0 a$, $C_{12} = C_{21} = 0$, i.e. the electric coupling between the spheres vanishes, and each of them behaves as an isolated conductor. The capacitance coefficient for each of the spheres simply becomes the capacitance of an isolated sphere.

Let us now consider a typical problem.

It should be recalled that the capacitance coefficients are constant for an invariable configuration of conductors and their mutual arrangement, regardless of the change in their charges and potentials. Therefore, we must analyse different situations equal in number to the unknown capacitance coefficients and then solve the system of equations.

Suppose that the spheres receive certain charges as a result of which their potentials become φ_1 and φ_2. After this, the second sphere is earthed. What will be the charges and potentials of the spheres after earthing?

Before we earthed the second sphere, the charges and potentials of the spheres were related through Eqs. (16.57). Since the potentials are known, the charges can be calculated with the help of these formulas. After earthing, the potential of the second sphere becomes equal to zero ($\varphi_2' = 0$), and the charge Q_2' is unknown. The charge of the first sphere, as before, is $Q_1' = Q_1$ since the sphere is isolated. Its potential φ_1' is unknown. Let us write Eqs. (16.57) for the case when the second sphere is earthed:

$$Q_1' = C\varphi_1', \quad Q_2' = \gamma\varphi_1', \quad Q_1' = Q_1. \tag{16.58}$$

The solution of these equations has the form

$$\varphi_1' = \frac{Q_1}{C} = \frac{C\varphi_1 + \gamma\varphi_2}{C} = \varphi_1 + \frac{\gamma}{C}\varphi_2, \quad Q_2' = \gamma\frac{Q_1}{C}. \tag{16.59}$$

It follows from (16.55) and (16.56) that

$$\gamma/C = -a/r. \tag{16.60}$$

Consequently, expressions (16.59) become

$$\varphi_1' = \varphi_1 - (a/r)\varphi_2, \quad Q_2' = -(a/r)Q_1, \tag{16.61}$$

i.e. after we earthed the second sphere, the potential of the first sphere changed by a fraction a/r of the potential which primarily was on the second sphere. The induced charge remaining on the second sphere is equal to the fraction a/r of the charge on the first sphere and has a sign opposite to that of the first sphere.

Let us remove the earthing of the second sphere and then earth the first sphere. Now we determine the potential of the second sphere and the charge of the first sphere.

Obviously, after earthing, the potential of the first sphere will be equal to zero ($\varphi_1'' = 0$), while the charge Q_1'' is unknown. Since the second sphere is now isolated, its charge remains unchanged upon the earthing of the first sphere ($Q_2'' = Q_2'$). Equations (16.57) assume the following form in this case:

$$Q_1'' = \gamma\varphi_2'', \quad Q_2'' = C\varphi_2'', \quad Q_2'' = Q_2', \tag{16.62}$$

www.EngineeringBooksPDF.com

whence

$$\varphi_2'' = \frac{Q_1'}{C} = -\frac{a}{rC} Q_1 = -\frac{a}{r}\varphi_1 + \left(\frac{a}{r}\right)^2 \varphi_2,$$
$$Q_1'' = \frac{r}{C} Q_2' = -\left(\frac{a}{r}\right)^2 Q_1. \tag{16.63}$$

These examples illustrate the methods of calculating capacitance coefficients, charges, and potentials for a system of several conductors in an electrostatic field.

Capacitors. A **capacitor** is a system of any two conductors carrying charges equal in magnitude and opposite in sign. *These conductors are called the capacitor plates.* Putting $Q_1 = Q$ and $Q_2 = -Q$ in (16.31) we obtain $\varphi_1 = Q(\alpha_{11} - \alpha_{12})$ and $\varphi_2 = Q(\alpha_{21} - \alpha_{22})$. Then the potential difference between the plates is

$$\Delta\varphi = \varphi_1 - \varphi_2 = Q(\alpha_{11} + \alpha_{22} - \alpha_{12} - \alpha_{21}). \tag{16.64a}$$

This means that the potential difference between the capacitor plates is proportional to the charge on a plate and, hence, the capacitor is characterized by a single parameter called the **capacitance. The capacitance** of a capacitor is defined by

$$\boxed{C = \frac{Q}{\Delta\varphi}} \tag{16.64b}$$

and is assumed to be positive by definition, i.e. Q and $\Delta\varphi$ in (16.64b) must have the same sign. A comparison of (16.64b) with (16.64a) shows that the capacitance of a capacitor is expressed in terms of the potential coefficients through the formula

$$C = (\alpha_{11} + \alpha_{22} - 2\alpha_{12})^{-1}, \tag{16.64c}$$

where $\alpha_{12} = \alpha_{21}$. Since α_{12} and α_{21} are negative, the *capacitance C in* (16.64c) *is always positive* [see (16.64b)]. Taking into account the meaning of the potential coefficients in (16.64c), we conclude that the *capacitance of a capacitor depends only on the geometrical characteristics of the capacitor plates and their mutual arrangement.*

Proceeding from (16.45) and using definition (16.64b), we obtain the following expression for the capacitance in terms of the capacitance coefficients:

$$C = \frac{C_{11}C_{22} - C_{12}^2}{C_{11} + C_{22} + 2C_{12}}. \tag{16.64d}$$

In most cases, the shape of the capacitor plates and their mutual arrangement is chosen in such a way that the external fields do not significantly affect the electric field between the plates and the field lines emerging from one plate necessarily terminate on the other. Owing to this, the equality of magnitudes of the charges on the plates is always ensured.

A capacitor can be represented in the form of a conductor placed into the cavity surrounded by a closed shell (Fig. 62*a*). If the inner conductor is a ball

www.EngineeringBooksPDF.com

or a sphere and the closed shell is the sphere concentric to it, we get a **spherical** capacitor (Fig. 62*b*). If the inner conductor is a straight solid cylinder and the shell is a hollow straight cylinder coaxial with it, we obtain a **cylindrical** capacitor (Fig. 62*c*). A system of two parallel plane conducting plates forms a **parallel-plate** capacitor (Fig. 62*d*).

The calculation of the capacitance of a capacitor boils down to determining the potential difference between the capacitor plates for a known value of charge on the plates. If, for example, the inner plate of a spherical capacitor

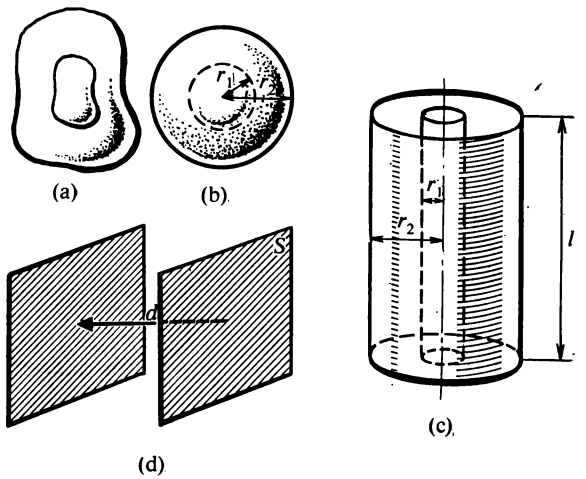


Fig. 62. Capacitors: general (a), spherical (b), cylindrical (c), and parallel-plate (d)

has the charge Q, the field strength in the gap between the inner and outer plates is equal to $E = Q/(4\pi\varepsilon_0 r^2)$ and is directed along the radius. Hence, the potential difference between the plates is

$$\varphi_2 - \varphi_1 = \int_{r_1}^{r_2} E\,dr = \frac{Q}{4\pi\varepsilon_0}\int_{r_1}^{r_2}\frac{dr}{r^2} = \frac{Q}{4\pi\varepsilon_0}\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \qquad (16.65)$$

Using formula (16.64b), we obtain the following expression for the capacitance of a spherical capacitor:

$$C = 4\pi\varepsilon_0 r_1 r_2/(r_2 - r_1). \qquad (16.66)$$

Similarly, we can find the capacitances of a cylindrical and a parallel-plate capacitor:

$$C = 2\pi\varepsilon_0 l/\ln(r_2/r_1), \quad C = \varepsilon_0 S/d.$$

Let us calculate the capacitance of a parallel-plate capacitor with the area of the plates equal to 1 $cm^2 = 10^{-4}\ m^2$ and the distance between the plates

www.EngineeringBooksPDF.com

$d = 1$ mm $= 10^{-3}$ m:

$$C = \frac{1}{4\pi \cdot 9 \cdot 10^9} \frac{10^{-4}}{10^{-3}} \text{ F} \approx 10^{-12} \text{ F} = 1 \text{ pF}. \tag{16.67}$$

Capacitors can be connected in series (Fig. 63*a*) or in parallel (Fig. 63*b*). *In the case of a series connection, the potential differences are added, while for a parallel connection, the charges on the plates are added.*

For a series connection, we have

$$U = U_1 + U_2, \quad U = Q/C, \quad U_1 = Q/C_1, \quad U_2 = Q/C_2, \tag{16.68}$$

where U is the potential difference between the outer plates of the capacitors, U_1 and U_2 are the potential differences between the plates of each capacitor,

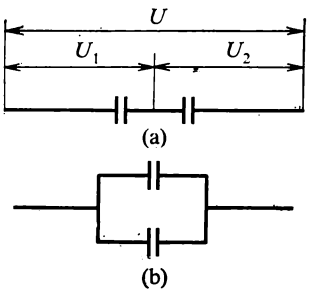


Fig. 63. Series (a) and parallel (b) connection of capacitors

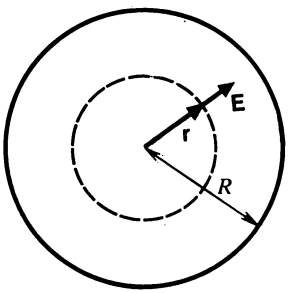


Fig. 64. Field inside a uniformly charged sphere

Q is the magnitude of charge on each capacitor plate (the charges on all the plates are modulo equal), C is the capacitance of the two capacitors, and C_1 and C_2 are the capacitances of each capacitor. It follows from (16.68) that

$$\boxed{\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}.} \tag{16.69}$$

Thus, with a series connection, the reciprocal values of capacitances are added.

For a parallel connection, we have

$$Q = Q_1 + Q_2, \quad Q = UC, \quad Q_1 = UC_1, \quad Q_2 = UC_2. \tag{16.70}$$

In this case,

$$\boxed{C = C_1 + C_2,} \tag{16.71}$$

i.e. the capacitances of the capacitors are added.

A conducting sphere in a uniform field. The field which appears when a conducting sphere is introduced into a uniform external electric field can be determined with the help of elementary methods.

www.EngineeringBooksPDF.com

First of all, let us find the field strength inside a uniformly charged sphere of radius R (Fig. 64), which, of course, is not a conductor. Suppose that the volume charge density inside the sphere is ρ. Then the charge contained in the spherical volume of radius $r < R$ is equal to $Q_r = (4/3)\,\pi r^3 \rho$. Using the Gauss theorem for this spherical volume, we obtain (ε_0 is the permittivity of the material from which the sphere is made)

$$E(r)\,4\pi r^2 = Q_r/E_0 = 4\pi r^3 \rho/(3\varepsilon_0) \tag{16.72}$$

and hence the field inside a uniformly charged sphere at a point characterized by the radius vector $\mathbf{r}$ is

$$\mathbf{E}(\mathbf{r}) = [(\rho/(3\varepsilon_0)]\,\mathbf{r}, \tag{16.73}$$

the origin of the radius vector coinciding with the centre of the sphere.

Now suppose that we have two spheres of the same radius and with the same volume density of unlike charges (Fig. 65). Let the negatively charged sphere

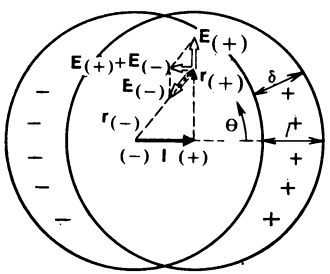


Fig. 65. To the calculation of the electric field of two spheres displaced relative to each other

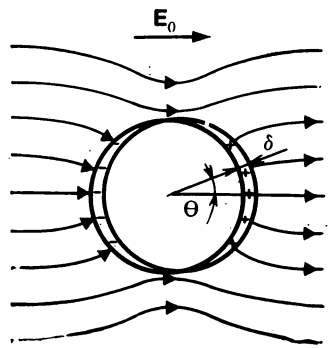


Fig. 66. A conducting sphere in a uniform electric field

be shifted to the left. The vector drawn from its centre to the centre of the other sphere is denoted by $\mathbf{l}$. We shall find the electric field strength at the inner points of these spheres. The fields created by the charges of each sphere are

$$\mathbf{E}_{(+)} = [|\rho|/(3\varepsilon_0)]\,\mathbf{r}_{(+)}, \quad \mathbf{E}_{(-)} = -[|\rho|/(3\varepsilon_0)]\,\mathbf{r}_{(-)}, \tag{16.74}$$

where $\mathbf{E}_{(+)}$ and $\mathbf{E}_{(-)}$ represent the fields created by the charges of the corresponding sign, and $\mathbf{r}_{(+)}$ and $\mathbf{r}_{(-)}$ are the radius vectors drawn to the point under consideration from the centres of the spheres with charges of the corresponding sign. The total electric field strength is given by

$$\mathbf{E} = \mathbf{E}_{(+)} + \mathbf{E}_{(-)} = [|\rho|/(3\varepsilon_0)]\,(\mathbf{r}_{(+)} - \mathbf{r}_{(-)}) = -[|\rho|/(3\varepsilon_0)]\,\mathbf{l}, \tag{16.75}$$

where

$$\mathbf{r}_{(-)} = \mathbf{l} + \mathbf{r}_{(+)} \tag{16.76}$$

(see Fig. 65). Thus, the electric field inside the spheres is constant and directed along the line connecting their centres.

www.EngineeringBooksPDF.com

At the points where the volumes of the spheres intersect, the charge density is equal to zero, since the positive and negative charge densities compensate each other. Only the crescent-shaped nonintersecting parts of the spheres are charged (see Fig. 65). The maximum width of these regions, equal to l, can be as small as desired.

Suppose now that a conducting sphere is placed into a uniform external electric field of strength $\mathbf{E}_0$. The electrostatic induction will lead to the appearance of surface charges. The signs of these charges and the direction of the external field are shown in Fig. 66. Inside the sphere, the electric field must be equal to zero, i.e. the surface charge distribution will be the same as in Fig. 65, and the field appearing in this case compensates the external field. Then [see (16.75)]

$$(|\rho|/3\varepsilon_0)\,\mathbf{l} = \mathbf{E}_0. \tag{16.77}$$

Thus, the centres of the imaginary charged spheres are shifted relative to each other along the line of force of the external field. Since $\mathbf{l}$ in (16.77) coincides in direction with $\mathbf{E}_0$, for scalar quantities we can write

$$|\rho|\, l = 3\varepsilon_0 E_0.$$

Obviously, the shift $\mathbf{l}$ of the spheres can be as small as desired if $|\rho|$ is sufficiently large. Hence, the charges appearing in this case can actually be considered as surface charges with varying surface density.

Let us find the distribution of the surface charge density as a function of the angle θ. The distance between the surfaces of the spheres in the direction of the angle θ is $\delta = l \cos\theta$ (Fig. 65). If the volume charge between the surfaces of the spheres is treated as the surface charge and if its surface density is denoted by σ, we obtain

$$\sigma \Delta S = \rho \Delta S \delta, \tag{16.78}$$

where the left-hand side contains the expression for the charge contained in the area element ΔS in terms of the surface density while the right-hand side expresses the same quantity in terms of the volume density. Consequently [see (16.78)],

$$\sigma = \rho\delta = \rho l \cos\theta = 3\varepsilon_0 E_0 \cos\theta, \tag{16.79}$$

where $\delta = l \cos\theta$.

We can now find the field strength at the surface of the conducting sphere:

$$E_n = \sigma/\varepsilon_0 = 3E_0 \cos\theta, \tag{16.80}$$

from which it follows that it varies between zero and thrice the value of the uniform field strength. Naturally, at all points of the spherical surface the field is directed along the normal to the surface.

Outside the sphere, at a finite distance from its surface, the field strength is the sum of the strengths of the external field and the fields created by the charged spheres shifted relative to each other or, which is the same, by the corresponding surface charges. The field outside a uniformly charged sphere is the same as if the entire charge were concentrated at its centre. Thus, we must find the field created by two unlike point charges of the same magnitude, located at a small

www.EngineeringBooksPDF.com

distance from each other. Such a system of charges is called a **dipole** (Fig. 67). The vector **l** drawn from the negative charge to the positive one is called the dipole arm. The vector

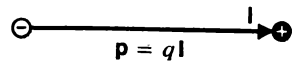


Fig. 67. A dipole

$$\mathbf{p} = q\mathbf{l} \qquad (16.81)$$

is called the **dipole moment.** In this formula, q indicates the magnitude of each of the dipole charges. In order to find the electric field outside the conducting sphere, we must find the field of a dipole whose charges are located at the centres of displaced spheres. It follows from (16.77) that the dipole moment is equal to

$$\mathbf{p} = (4/3)\pi R^3 \rho\mathbf{l} = 4\pi\varepsilon_0 R^3 \mathbf{E}_0, \qquad (16.82)$$

where R is the radius of the sphere.

The field of a dipole. The electric field of a dipole is the sum of the fields created by dipole charges. The dipole arm is as small as desired, and hence it can be assumed much smaller than the distance to the points at which the field is calculated. Let us find the potential of the dipole. At the point P (Fig. 68), the potential is given by

$$\varphi(P) = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{r_{(+)}} - \frac{1}{r_{(-)}}\right) = \frac{q}{4\pi\varepsilon_0}\left(\frac{r_{(-)} - r_{(+)}}{r_{(+)} r_{(-)}}\right). \qquad (16.83)$$

Since $l \ll r$, we can assume that $r_{(-)} - r_{(+)} \simeq l\cos\theta$ and $r_{(-)} r_{(+)} \simeq r^2$. We can characterize the position of the point P by the radius vector $\mathbf{r}$ with the origin at any point of the dipole, since the dipole has as small geometric dimensions as desired.

Then [see (16.83)] we can write

$$\boxed{\varphi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3},} \qquad (16.84)$$

where $ql\cos\theta = (\mathbf{p}\cdot\mathbf{r})/r$, whence

$$\mathbf{E} = -\operatorname{grad}\varphi = \frac{1}{4\pi\varepsilon_0}\left[\frac{3(\mathbf{p}\cdot\mathbf{r})\,\mathbf{r}}{r^5} - \frac{\mathbf{p}}{r^3}\right]. \qquad (16.85)$$

The electric field of a dipole decreases in inverse proportion to the third power of the distance, i.e. more rapidly than the Coulomb field of a charge. The lines of force of the dipole field are shown in Fig. 69.

Formula (16.85) allows us to construct the field lines when a conducting sphere is placed in a uniform external field. At each point, the field strength is equal to the sum of the strengths of the uniform external field $\mathbf{E}_0$ and the field $\mathbf{E}$ created by the charges induced on the surface of the conducting sphere. The field lines for this case are shown in Fig. 66.

Method of image charges. While solving the problem about a conducting sphere in a uniform external field, we made an assumption whose validity was not proved. Namely, we constructed a certain field satisfying all the conditions of the problem and assumed that there is no other field that would satisfy the same conditions. In other words, wa assumed that the solution of the problem is

www.EngineeringBooksPDF.com

unique. Otherwise, the obtained concrete solution would not necessarily be the solution which is realized in fact. It is proved in the theory of electricity and magnetism that the solution of a problem, which satisfies all necessary conditions, is unique. Later we shall consider the conditions mentioned here and give a rough proof of this statement. At the moment, we shall admit without proof that this statement is correct. This allows us to find the solution of the problem with the help of some conjectures or constructions and then conclude,

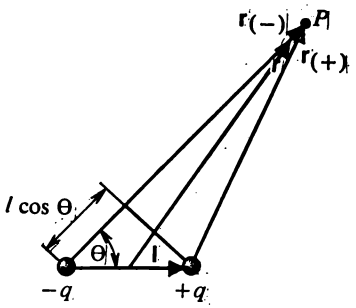


Fig. 68. To the calculation of a dipole field

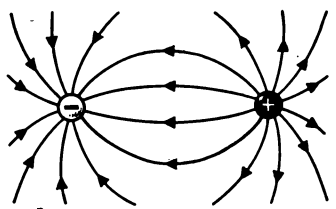


Fig. 69. Field lines in the vicinity of a dipole

on the basis of the uniqueness theorem, that the field found in this way is the solution of the problem. The solution of the above problem about a conducting sphere in a uniform external electric field may serve as an example of a successful conjecture.

There exists a visual method of constructing the field satisfying the conditions of the problem, which is called the method of image charges. The essense of this method consists in the following. The field of a point charge is well known. Hence, a system of charges is sought whose total field satisfies all the conditions of the problem. On the basis of the uniqueness theorem, we conclude that this field gives the required solution. Mathematically, this problem is reduced to determining the potential satisfying the conditions of the problem. The field vector $\mathbf{E}$ is normal to the equipotential surfaces and is calculated as the gradient of potential, taken with the opposite sign. The shape of equipotential surfaces of the system of point charges can, in principle, be easily obtained. Let us consider, for example, the field of two positive point charges q located at a distance $2d$ from each other (Fig. 70). Since the potential of a point charge at a distance r from it is $\varphi = q/(4\pi\varepsilon_0 r)$, the potential of the system of two identical point charges (see Fig. 70) at a point (x, y, z) is defined by

$$\varphi(x, y, z) = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{\sqrt{(x-d)^2+y^2+z^2}} + \frac{1}{\sqrt{(x+d^2)+y^2+z^2}}\right). \qquad (16.86)$$

From this equation, we can obtain the equation for equipotential surfaces:

$$\frac{1}{\sqrt{(x-d)^2+y^2+z^2}} + \frac{1}{\sqrt{(x+d)^2+y^2+z^2}} = \text{const}. \qquad (16.87)$$

www.EngineeringBooksPDF.com

Each of them is characterized by the corresponding potential $\varphi_1 = \text{const}$, $\varphi_2 = \text{const}$, etc.

Figure 70 shows the lines of intersection of the plane XY with equipotential surfaces. The equipotential surfaces proper can be obtained by rotating the pattern depicted in Fig. 70 around the X-axis.

Suppose that an isolated conducting surface coincides with one of the equipotential surfaces, whose potential is equal to φ_0. If we assume that the charge on this surface is $2q$ and the surface potential is equal to φ_0, the system of equipotential surfaces and the field corresponding to it completely satisfy the conditions of the problem about the field of a charged surface. At all points external relative

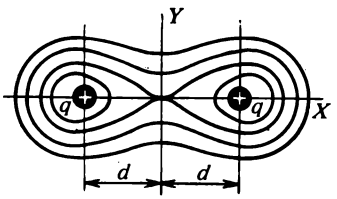


Fig. 70. Equipotential surfaces of two like point charges

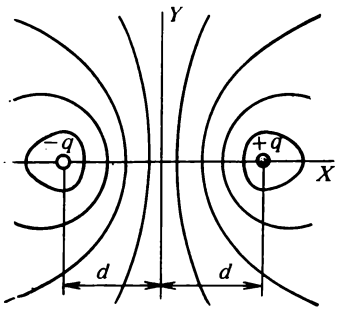


Fig. 71. Equipotential surfaces of two unlike and unequal point charges

to this surface, the potential is determined by formula (16.86). Thus, the determination of the characteristics of the field created by the charged conductor is reduced to the determination of the characteristics of the field created by two like and equal point charges. This is the essence of the method of image charges. The examples considered below will clarify the origin of the name given to this method.

The potential of two unlike point charges is determined in a way similar to (16.86):

$$\varphi = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{\sqrt{(x-d)^2+y^2+z^2}} - \frac{1}{\sqrt{(x+d)^2+y^2+z^2}}\right). \qquad (16.88)$$

The shape of equipotential surfaces in this case is shown in Fig. 71. The potential along the Y-axis is equal to zero, and hence it is equal to zero in the plane $X = 0$.

Suppose now that the entire infinite half-space $X < 0$ is filled by a conductor bounded by the plane YX and suppose that the charge $+q$ is located as shown in Fig. 71. Obviously, on account of electrostatic induction, this charge will induce on the surface of the conductor the charge $-q$. In this case, the potential φ of the conductor must be equal to zero, and the lines of force at each point of the surface must be normal to it. It is clear that the pattern of the lines of force in the half-space $X > 0$, shown in Fig. 71, completely satisfies these conditions.

www.EngineeringBooksPDF.com

Hence, the problem of determining the characteristics of the field of the point charge $+q$, located at the distance d from a plane surface of the conductor filling the half-space $X < 0$, is reduced to finding the characteristics of the fields of two point charges q and $-q$. The charge $-q$ is located at the point which would be the image of the point charge q if the plane $X = 0$ were a mirror. Hence the name of the method of image charges (method of images). Instead of the conducting body filling the half-space $X < 0$, we could take an earthed conducting plate parallel to the plane $X = 0$. The method of calculation and the field remain unchanged. If the plate is not earthed, at the surface of the plate facing the negative values of the X-axis positive surface charges are induced, which completely alter the nature of the field. In this case, the field is no longer the superposition of the fields created by the charge q and by its image.

Let us find the field due to a charge q located at a point $x = d$ in the presence of the earthed conducting plane $X = 0$. At all points $x > 0$, the field potential is given by formula (16.88). The electric field strength in the plane $Z = 0$ is

$$E_x = -\frac{\partial \varphi}{\partial x} = \frac{q}{4\pi\varepsilon_0}\left\{\frac{x-d}{[(x-d)^2+y^2]^{3/2}} - \frac{x+d}{[(x+d)^2+y^2]^{3/2}}\right\}, \tag{16.89}$$

$$E_y = -\frac{\partial \varphi}{\partial y} = \frac{q}{4\pi\varepsilon_0}\left\{\frac{y}{[(x-d)^2+y^2]^{3/2}} - \frac{y}{[(x+d)^2+y^2]^{3/2}}\right\}. \tag{16.90}$$

In the plane $X = 0$, the component E_y vanishes, while

$$E_x = -\frac{q}{2\pi\varepsilon_0}\frac{d}{(z^2+y^2+d^2)^{3/2}}. \tag{16.91}$$

In the plane $X = 0$, the surface charge density is given by [see (16.12)]

$$\sigma = -\frac{q}{2\pi}\frac{d}{(z^2+y^2+d^2)^{3/2}}. \tag{16.92}$$

The total surface charge on the plane $X = 0$ is

$$\int_{-\infty}^{\infty}\int \sigma \, dz \, dy = -\frac{qd}{2\pi}\int_{-\infty}^{\infty}\int \frac{dz\,dy}{(z^2+y^2+d^2)^{3/2}} = -q, \tag{16.93}$$

i.e. the charge induced on the conductor is equal to the inducing charge with the opposite sign [see (16.20)].

The force of interaction between the point charge q and the charge on the surface $x = 0$ is equal to the force of interaction between the charge q and its image:

$$F = -q^2/(16\pi\varepsilon_0 d^2). \tag{16.94}$$

The minus sign indicates that the point charge is attracted to the conducting earthed surface.

Of course, the method of images is not reduced in all cases to finding the mirror image of charges in the literal sense. Let us consider a pattern of equipotential surfaces created by two charges of different magnitude. For the sake of convenience, we introduce the polar system of coordinates with the origin

www.EngineeringBooksPDF.com

at the point O (Fig. 72). The polar axis passes through the point charges q_1 and q_2. The polar coordinates of q_1 and q_2 are $\theta_1 = 0$, $r_1 = d_1$ and $\theta_2 = 0$, $r_2 = d_2$ respectively. The potential at the point P is given by

$$\varphi(r, \theta) = \frac{1}{4\pi\varepsilon_0}\left(\frac{q_1}{\sqrt{r^2+d_1^2-2rd_1\cos\theta}} + \frac{q_2}{\sqrt{r^2+d_2^2-2rd_2\cos\theta}}\right). \quad (16.95)$$

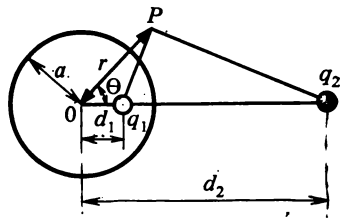


Fig. 72. To the determination of equipotential surfaces of two point charges of different magnitude

If $d_1 = a^2/d_2$ $(a < d_2)$ and $q_2 = -aq_2/d_2$, then $\varphi(a, \theta) = 0$, i.e. on the sphere of radius a, the potential is equal to zero. Consequently, this sphere is an equipotential surface with the zero value of the potential. If we replace it by a real conducting earthed surface, the field will remain unchanged. Thus, if we have a conducting earthed sphere of radius a and a point charge q_2 outside it at a distance d_2 from the centre of the sphere, the field outside the sphere is the same as that created by the charge q_2 and its "image", viz. the charge $q_1 = -aq_2/d_2$, placed at the point with the coordinates $d_1 = a^2/d_2$, $\theta = 0$ inside the sphere. The force of interaction between the charge q_2 and this sphere is given by

$$F = \frac{q_1 q_2}{4\pi\varepsilon_0 (d_2-d_1)^2} = -\frac{d_2 a q_2^2}{4\pi\varepsilon_0 (d_2^2-a^2)^2}. \quad (16.96)$$

In electrostatics, there is no field inside a conductor, and volume charges do not exist. Near the surface of a conductor, the electric field vector is normal to the surface and is proportional to the surface charge density.
On the convex surface of a conductor, the surface charge density and field strength increase with the curvature of the surface, i.e. with decreasing radius of curvature. On the concave surface of a conductor, the surface charge density decreases.
In differential form, Ohm's law is valid for the varying as well as for a constant electrical conductivity, regardless of the causes and nature of its variation.

The capacitance of an isolated conductor depends only on its shape and size. The potential and capacitance coefficients are determined only by geometrical characteristics of conductors and their mutual arrangement.
Capacitance coefficients with identical indices are always positive, while those with different indices are either equal to zero or negative.

Due to which property of the electrostatic field is the tangential component of the field near the surface of a conductor equal to zero?

Examle 16.1. *Find the force of interaction between a conducting sphere of radius a and a point charge q_2 located at a distance d_2 from the centre of the sphere, if the charge Q is distributed over the sphere.*

The mutual arrangement of the sphere and the charge is shown in Fig. 72. The charge q_2 induces on the sphere its image in the form of the charge $q_1 = -q_2 a/d_2$ at a distance $d_1 = a^2/d_2$ from the centre of the sphere. However, now the interaction is not reduced to the force of attraction between the charge q_2 and its image since, by assumption, the sphere has the charge Q and not q_1. Consequently, in order to describe the interaction, we must add one more "image" of the charge which creates a constant potential on the sphere and is equal to $Q - q_1$.

www.EngineeringBooksPDF.com

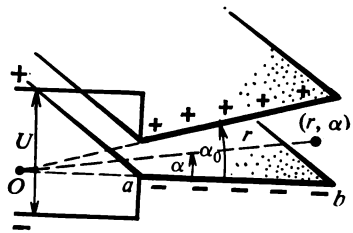


Fig. 73. To the calculation of the field of a capacitor with nonparallel plates

Hence, we must place at the centre of the sphere the charge $Q - q_1 = Q + q_2a/d_2$. The interaction of the point charge q_2 with the sphere carrying a charge Q is the sum of interactions of q_2 with the "images" q_1 and $Q + q_2a/d_2$. Thus, the force of interaction is

$$F = \frac{q_2}{4\pi\varepsilon_0}\left[\frac{Q+q_2a/d_2}{d_2^2} - \frac{q_2a}{d_2(d_2-d_1)^2}\right]. \quad (16.97)$$

Example 16.2. *Find the force of interaction between a conducting sphere of radius a, maintained at a constant potential φ_0 and a point charge q_2 at a distance d_2 from the centre of the sphere.*

The mutual arrangement of the sphere and the charge is shown in Fig. 72. The potential created by the charge q_2 and its image q_1 on the sphere is equal to zero. In order to make it equal to φ_0 it is necessary to place the "image" $Q = 4\pi\varepsilon_0 a\varphi_0$ at the centre of the sphere. The force of interaction between the point charge q_2 and the sphere maintained at the potential φ_0 is given by

$$F = \frac{q_2}{4\pi\varepsilon_0}\left[\frac{Q}{d_2^2} - \frac{q_2a}{d_2(d_2-d_1)^2}\right]. \quad (16.98)$$

Example 16.3. *Two plane conducting plates form an angle α_0* (Fig. 73). *These plates are perpendicular to the plane of the figure and are infinitely long. A constant potential difference U_0 is applied between the planes. Find the field strength between the plates and the capacitance over a length l. The width of the plane is b — a. We assume that the plates do not touch each other at the point O but are separated by a sufficiently small distance so that edge effects can be neglected.*

The field is axisymmetric. Consequently, it is convenient to use the cylindrical system of coordinates, where the Z-axis is normal to the plane of the figure. We denote the axial angle by α and the distance from the axis by r. Then the Laplace equation becomes

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\varphi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial\alpha^2} = 0, \quad (16.99)$$

where we took into account that $\partial^2\varphi/\partial z^2 = 0$ due to the cylindrical symmetry of the field. We seek its solution in the form

$$\varphi(r, \alpha) = R(r)\,\Phi(\alpha). \quad (16.100)$$

Substituting this equation into (16.99), we find

$$\frac{\Phi}{r}\frac{d}{dr}\left(r\frac{dR}{dr}\right) + \frac{R}{r^2}\frac{d^2\Phi}{d\alpha^2} = 0.$$

Multiplying both sides of this equation by $r^2/R\Phi$, we obtain

$$\frac{r}{R}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = -\frac{1}{\Phi}\frac{d^2\Phi}{d\alpha^2}. \quad (16.101)$$

The left- and right-hand sides of this equation contain different independent variables. Consequently, the equality is possible only if its left- and right-hand sides are separately equal to the same constant. Hence we put

$$\frac{r}{R}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = n^2, \quad \frac{1}{\Phi}\frac{d^2\Phi}{d\alpha^2} = -n^2, \quad (16.102)$$

where n^2 is a constant. The solution of the equation for Φ is obvious:

$$\Phi = \begin{cases} B_1\alpha + B_2 & \text{for } n = 0 \\ A_1\sin n\alpha + A_2\cos n\alpha & \text{for } n \neq 0. \end{cases} \quad (16.103)$$

www.EngineeringBooksPDF.com

We shall seek the solution of the equation for R in the form $R = Ar^{\beta}$ ($\beta \neq 0$). Substituting this expression into the first equation of (16.102) we obtain the equality

$$\beta^2 = n^2, \tag{16.104}$$

from which it follows that $\beta = \pm n$. For $n = 0$, the first of Eqs. (16.102) is simplified:

$$r\frac{dR}{dr} = \text{const}$$

and can be satisfied by the function

$$R = D_1 \ln r + D_2.$$

Consequently, the solution of (16.102) can be written in the final form as follows:

$$R = \begin{cases} D_1 \ln r + D_2 & \text{for } n = 0, \\ C_1 r^n + C_2 r^{-n} & \text{for } n \neq 0. \end{cases} \tag{16.105}$$

Let us try to find the solution of the problem which would be independent of r. In other words, for $n = 0$, $D_1 = 0$ we have [see (16.103)] $\varphi(\alpha) = B_1\alpha + B_2$. The boundary conditions for φ have the form $\varphi(0) = 0$, $\varphi(\alpha_0) = U_0$, i.e. $0 = B_2$, $U_0 = B_1\alpha_0$. Consequently,

$$\varphi(\alpha) = U_0\alpha/\alpha_0. \tag{16.106}$$

The electric field strength is given by

$$E_\alpha = -\frac{1}{r}\frac{\partial\varphi}{\partial\alpha} = -U_0/(r\alpha_0) \tag{16.107a}$$

The surface density of charges on the plates is

$$\sigma_1 = \varepsilon E_\alpha(\alpha = 0) = -\varepsilon U_0/(r\alpha_0), \quad \sigma_2 = -\varepsilon E_\alpha(\alpha = \alpha_0) = \varepsilon U_0/(r\alpha_0). \tag{16.107b}$$

The charge on the length l of each plate (in magnitude) is expressed by

$$Q = l\int_a^b \sigma\, dr = (l\varepsilon_0 U_0/\alpha_0)\ln(b/a) \tag{16.108}$$

The capacitance corresponding to the length l is equal to

$$C = \frac{Q}{U_0} = \frac{l\varepsilon_0 \ln(b/a)}{\alpha_0}. \tag{16.109}$$

Sec. 17. Electrostatic Field in the Presence of a Dielectric

The influence of a dielectric on an electrostatic field and various mechanisms of polarization are discussed. The relations between the volume and surface densities of bound charges and polarization are introduced. The phenomena occurring at the interface between dielectrics are discussed.

Dipole moment of a continuous charge distribution. The effect of matter on the electric and magnetic fields was experimentally discovered and investigated by Faraday. The results of these investigations led Faraday to put forth the idea

www.EngineeringBooksPDF.com

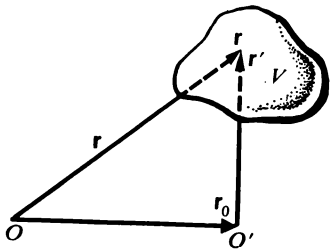


Fig. 74. To the determination of the dipole moment of a continuously distributed charges

of short-range interactions and the concept of a field. The electrostatic induction was discovered by him in 1837, when he also introduced the terms "dielectric" and "dielectric constant".

Suppose that in a certain volume V (Fig. 74) we have a continuously distributed charge with a volume density ρ and that the volume is electrically neutral as a whole. This, however, does not mean that the positive and negative charges compensate each other at each point inside this volume. If positive and negative charges are distributed in the volume according to different laws, the overall charge density ρ will be positive at some points in the volume and negative at some other points. Mathematically, the condition of neutrality of volume V can be expressed as follows:

$$\int_V \rho \, dV = 0. \tag{17.1}$$

If $\rho = 0$ at all points in the volume, the material system is electrically neutral inside the volume V: it is acted upon by an external electric field and it does not generate any electric field on its own. However, *if the charge density ρ is positive at some points of the volume V and negative at some other points, the system will have electric properties even though the total charge in the volume V is equal to zero: the system is acted upon by an external electric field and it itself generates an electric field.* To a first approximation, the electric properties of a neutral system are characterized by its dipole moment. Formula (16.81) serves as the definition of dipole moment for two point charges. For a continuous charge distribution, the dipole moment (Fig. 74) is defined by the formula

$$\boxed{\mathbf{p} = \int_V \rho \mathbf{r} \, dV.} \tag{17.2}$$

In this equation, the radius vector $\mathbf{r}$ is measured from any point O which can be taken as the reference point. Obviously, the form of the expression (17.2) is independent of the choice of this point. In order to prove this, let us take point O' as the reference point from which measurements are made, and suppose that the position of this point relative to O is characterized by the radius vector $\mathbf{r}_0$ (see Fig. 74). Formula (17.2) has the following form for point O':

$$\mathbf{p}' = \int_V \rho \mathbf{r}' \, dV. \tag{17.3}$$

This equation can be transformed as follows:

$$\mathbf{p}' = \int_V \rho (\mathbf{r} - \mathbf{r}_0) \, dV = \int_V \rho \mathbf{r} \, dV - \int_V \mathbf{r}_0 \rho \, dV = \int_V \rho \mathbf{r} \, dV = \mathbf{p}, \tag{17.4}$$

www.EngineeringBooksPDF.com

Q.E.D. Here, $\mathbf{r} = \mathbf{r}_0 + \mathbf{r}'$, and [see (17.1)]

$$\int_V \mathbf{r}_0 \rho \, dV = \mathbf{r}_0 \int_V \rho \, dV = 0. \tag{17.5}$$

Let us use formula (17.2) for calculating the dipole moment of two point charges which can be treated as charges lying in indefinitely small volumes ΔV_1 and ΔV_2 (Fig. 75):

$$\mathbf{p} = \int_V \rho \mathbf{r} \, dV = \int_{\Delta V_1} \rho \mathbf{r} \, dV + \int_{\Delta V_2} \rho \mathbf{r} \, dV$$
$$= \mathbf{r}_1 \int_{\Delta V_1} \rho \, dV + \mathbf{r}_2 \int_{\Delta V_2} \rho \, dV = \mathbf{r}_1 Q_1 + \mathbf{r}_2 Q_2, \tag{17.6}$$

where Q_1 and Q_2 are charges in volumes ΔV_1 and ΔV_2 respectively, and $\mathbf{r}_1$ and $\mathbf{r}_2$ are the radius vectors of these volumes. For example, suppose that a positive

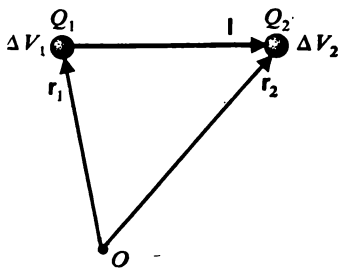


Fig. 75. To the calculation of the dipole moment of two point charges with the help of the formula for continuous distribution of charges

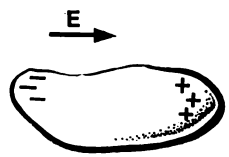


Fig. 76. Polarization of nonpolar dielectrics in an electric field

charge $Q_2 = Q$ is located in volume ΔV_2. In view of the electrical neutrality of the system, in this case $Q_1 = -Q$, and formula (17.6) assumes the form

$$\boxed{\mathbf{p} = Q(\mathbf{r}_2 - \mathbf{r}_1) = Q\mathbf{l},} \tag{17.7}$$

which is analogous to (16.81).

The field strength of a neutral system with a dipole moment $\mathbf{p}$ is given by formulas (16.84) and (16.85).

Polarization of dielectrics. Dielectrics are materials in which the application of an electric field does not lead to a displacement of charges as, for example, in conductors. *This, however, does not mean that charges in a dielectric do not move at all when an electric field is applied. The charges do move in such a case, but are not displaced by large distances.*

Let us consider an electrically neutral volume of a dielectric (Fig. 76). An external electric field tends to displace positive charges in the direction of the

www.EngineeringBooksPDF.com

field and the negative charges in the opposite direction. Hence an excess charge is accumulated in the direction of the field, while a deficiency of charge is created in the opposite direction. Consequently, the dielectric acquires a dipole moment. This process is called **polarization**.

The extent of polarization of a dielectric is characterized by the **dielectric polarization** which is defined as the ratio of the dipole moment $\Delta \mathbf{p}$ of a dielectric to its volume ΔV:

$$\boxed{\mathbf{P} = \frac{\Delta \mathbf{p}}{\Delta V}.} \tag{17.8}$$

Molecular pattern of polarization. A dielectric consists of atoms and molecules, and any infinitely small physical volume element of a dielectric is electrically neutral. The positive charge is concentrated at the atomic nuclei, while the negative charge is distributed over the electron shells of atoms and molecules. Positive and negative charges are located at different points in space and hence atoms and molecules can have electric dipole moments which vary with the frequency of electron oscillations in atoms, which is of the order of 10^{15} s^{-1}.

If in the absence of an external electric field the distribution of the *electron cloud is spherically symmetric with respect to the nucleus, the atom does not possess an electric dipole moment. Similarly, the positive and negative charges in a molecule may have such a symmetry of distribution that the molecule does not have a dipole moment.* Such molecules and atoms are called **nonpolar** and include, among others, the helium atom, diatomic molecules consisting of identical atoms (H_2, N_2, O_2, . . .), and symmetric polyatomic molecules like CO_2 and CH_4. In the absence of an external electric field, such a dielectric is not polarized.

Molecules and atoms which possess an electric dipole moment in the absence of an external electric field are called **polar**, and include CO, N_2O, SO_2 etc. The permanent dipole moment in such molecules is of the order of 10^{-29}-10^{-30} C·m. This corresponds to a dipole consisting of two elementary charges of 1.6×10^{-19} C separated by a distance 10^{-10} m, i.e. of the order of atomic dimensions.

In the absence of an external electric field, the permanent dipole moments of individual molecules are oriented at random and hence their sum in an infinitely small physical volume is equal to zero. In other words, the dielectric is not polarized.

When an electric field is applied to a dielectric, the positive charges tend to move along the field vector, while the negative charges tend to move in the opposite direction. *As a result, nonpolar molecules acquire a dipole moment and the dielectric is polarized.* Polar molecules also acquire an additional dipole moment induced by the external field and also get polarized, although this polarization is insignificant. The basic polarization mechanism for polar molecules is different: *in an external electric field, the permanent dipole moments of molecules are acted upon by the moments of force* [Fig. 77, see (19.7)] *which tend to orient the dipole moments along the field vector. Consequently, the molecules are reoriented in such a way that infinitely small physical volume elements of the dielectric acquire dipole moments, i.e. the dielectric is polarized.* The polarization

www.EngineeringBooksPDF.com

due to the reorientation of molecules is much stronger than due to the formation of additional dipole moments induced by the external field.

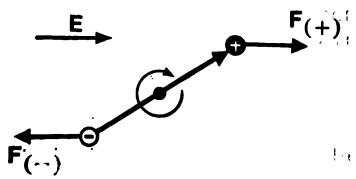


Fig. 77. Polarization of polar dielectrics in an electric field

Besides these mechanisms of polarization, there exists another. *In ionic crystals, the positive ions under the action of an external electric field are displaced along the field, while the negative ions are displaced in the opposite direction.* This results in a certain deformation of the crystal lattice or a relative displacement of sublattices, which leads to the emergence of dipole moments in the dielectric, i.e. to the polarization of the dielectric. Such a polarization is called **ionic lattice polarization.**

The quantitative measure of polarization in all cases is the dielectric polarization **P**. The polarization mechanism is revealed only as a result of the investigation of the dependence of **P** on the strength of the applied electric field and other factors (see Chap. 3). The formula which relates the electric field strength, the electric displacement and polarization remains unchanged [see (17.29)].

The polarization of nonpolar molecules is

$$\mathbf{P}=\frac{1}{\Delta V}\sum_{\Delta V}\mathbf{p}_i=N\mathbf{p}_0, \tag{17.9}$$

where ΔV under the symbol $\sum$ means that the summation is extended to all molecules in the volume ΔV, N is the concentration of molecules; $\mathbf{p}_0$ is the induced dipole moment (its value is the same for all molecules) whose direction coincides with that of the external electric field **E**. In the absence of an external field, $\mathbf{p}_0 = 0$, and hence $\mathbf{P} = 0$, i.e. there is no polarization.

The principal mechanism of polarization in polar molecules is the reorientation of the direction of permanent dipole moments under the action of an external field. The formula for polarization has the form

$$\mathbf{P}=\frac{1}{\Delta V}\sum_{\Delta V}\mathbf{p}_i=N\langle\mathbf{p}\rangle, \tag{17.10}$$

where $\langle\mathbf{p}\rangle$ is the average value of the dipole moments which are equal in magnitude but are oriented in different directions in space. In isotropic dielectrics, the direction of the average dipole moments coincides with that of the external electric field. In anisotropic dielectrics, which have different properties in different directions, such a coincidence is not observed. The relation between the polarization and the field strength is more complicated in such dielectrics (see Chap. 3). In polar dielectrics, the contribution from induced dipole moments to the polarization is much smaller than the contribution from the realignment of permanent dipole moments and is usually not taken into account. If necessary, this contribution can be taken into account by adding the right-hand side of Eq. (17.9) to the right-hand side of formula (17.10).

Ionic lattice polarization is defined by formula (17.10) in which $\langle\mathbf{p}\rangle$ is the average value of the dipole moments created in volume ΔV due to a displace-

www.EngineeringBooksPDF.com

ment of ions at the crystal lattice sites. In most cases, this polarization is anisotropic.

Dependence of polarization on the electric field strength. In electrets and ferroelectrics, the polarization may be nonzero in the absence of electric field ($E = 0$, $P \neq 0$). The polarization in other dielectrics in the absence of electric field is equal to zero. In the general case, the dependence of polarization on field strength can be expressed in the form

$$P_i = \varepsilon_0 \sum_j \varkappa_{ij} E_j + \varepsilon_0 \sum_{j,k} \varkappa_{ijk} E_j E_k + \ldots,$$

where the indices i, j, k . . . enumerate the components of the quantities along the Cartesian coordinate axes. ($i = x, y, z$; $j = x, y, z, \ldots$). Hence, in the general case, polarization depends not only on the first power of the electric field strength, but also on its higher powers. *If the dependence on the higher powers is significant, the dielectric is called nonlinear.* Such a nonlinearity is usually manifested only in very strong electric fields, although there are some special materials in which nonlinearity is observed in comparatively weak fields.

If the nonlinearity is insignificant, the polarization is expressed in terms of first powers of the field components:

$$P_i = \varepsilon_0 \sum_j \varkappa_{ij} E_j.$$

Such a dielectric is called **linear.** If the properties of such a dielectric are different in different directions, the dielectric is called **anisotropic.** The set of nine quantities $\varkappa_{ij}$ constitutes the **dielectric susceptibility tensor** which completely characterizes the electric properties of a dielectric. If the properties of a dielectric are identical in all directions, it is called a **linear isotropic dielectric.** Its dielectric properties are then characterized by a scalar quantity called the dielectric susceptibility.

For a linear isotropic dielectric, we have

$$\mathbf{P} = \varkappa \varepsilon_0 \mathbf{E}, \tag{17.11}$$

where $\varkappa$ is the dielectric susceptibility. In the Gaussian absolute system of units the dielectric susceptibility $\varkappa'$ is a quantity 4π times smaller than $\varkappa$ in formula (17.11):

$$\varkappa' = \varkappa/(4\pi). \tag{17.12}$$

For most solid and liquid dielectrics, the dielectric susceptibility is expressed by numbers of the order of several units. The dielectric susceptibility of most gases is a fraction of a thousandth part of unity and in most cases need not be taken into consideration. However, there are dielectrics whose susceptibility reaches very high values. For example, $\varkappa = 80$ for water, 25-30 for alcohol, while in ferroelectrics (Rochelle salt, barium titanate, etc.), the dielectric susceptibility attains values of several thousands.

www.EngineeringBooksPDF.com

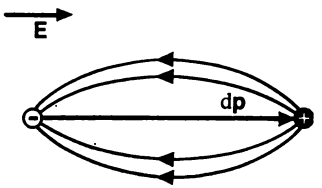


Fig. 78. Mechanism of field weakening during polarization

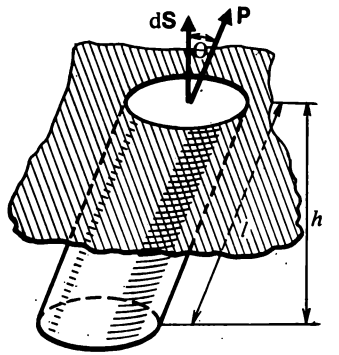


Fig. 79. Calculation of the charge crossing a surface area element upon polarization

The effect of polarization on electric field. In accordance with formula (17.8), the dipole moment of a volume element dV is

$$d\mathbf{p} = \mathbf{P}\, dV = \varkappa\varepsilon_0 \mathbf{E}\, dV, \tag{17.13}$$

i.e. coincides in direction with the electric field $\mathbf{E}$, since $\varkappa > 0$. Hence, the field created by the dipole moment is directed against the external field and weakens it (Fig. 78). Thus, *the field inside a dielectric is weakened as a result of polarization.* The role of polarization is reduced just to a separation of positive and negative charges, leading to the appearance of charges in the volume and on the surface of the dielectric. These charges are called **polarization charges** or **bound charges**, since they are as if attached to different places in the dielectric and cannot move freely in its volume or on its surface. *Bound charges give rise to an electric field in the same way as free charges, and are in no way different from them in this respect.* Thus, the presence of a dielectric is taken into account by considering the electric field created by bound charges induced as a result of polarization. Hence it is necessary to find an expression for bound charges.

Volume and surface density of bound charges. Let us consider a surface element dS (Fig. 79) inside a nonpolarized dielectric. As a result of polarization the charges move across this surface element. Let us calculate the charge intersecting the element dS when a polarization $\mathbf{P}$ appears. In order to simplify the formulas, we shall assume that only positive charges are displaced. We denote the dipole charge by q, the dipole arm corresponding to the polarization $\mathbf{P}$ by l and the charge concentration by N. The area element dS (see Fig. 79) is intersected upon polarization $\mathbf{P}$ by all positive charges which were present in volume $dV = h\, dS = l \cos\theta\, dS$ of an oblique cylinder with base dS before displacement was caused by polarization. Consequently,

$$dQ = Nql \cos\theta\, dS = P\, dS \cos\theta = \mathbf{P}\cdot d\mathbf{S}. \tag{17.14}$$

Let us now consider a certain volume V (Fig. 80). As a result of polarization, the surface S bounding this volume is crossed by charges. Depending on the

www.EngineeringBooksPDF.com

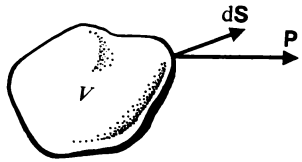


Fig. 80. To the derivation of the expression for the bound space charge

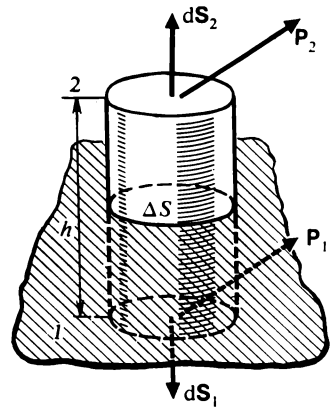


Fig. 81. To the derivation of the expression for the surface density of bound charges

balance of charges entering and leaving the volume, a bound charge of volume density ρ_b appears. Taking (17.14) into account, we can write the law of charge conservation in volume V in the form

$$\int_V \rho_b \, dV = -\int_S \mathbf{P} \cdot d\mathbf{S}. \tag{17.15}$$

The minus sign indicates that the charge induced in the volume is opposite in sign to the charge flowing through the surface bounding it. Applying Gauss' theorem to the right-hand side of Eq. (17.15), we can write it in the following form:

$$\int_V (\rho_b - \operatorname{div} \mathbf{P}) \, dV = 0. \tag{17.16}$$

If Eq. (17.16) is identically satisfied for all values of V, the integrand will be identically equal to zero. Consequently,

$$\boxed{\rho_b = -\operatorname{div} \mathbf{P}.} \tag{17.17}$$

Thus, *bound space charges appear only in the case when the polarization* $\mathbf{P}$ *changes from point to point.* This is clear even without calculations since in the case of a uniform polarization the charges moving on to new places occupy the vacancies created by the same number of charges. As a result, the corresponding volume of the dielectric remains electrically neutral.

Surface charges appear at the interface between two different dielectrics. This is apparent from the following considerations. The dielectric polarization is different at different points for the same electric field strength. Consequently,

www.EngineeringBooksPDF.com

a different number of polarization charges cross the boundary surface from different dielectrics. As a result, a certain bound charge, called the **bound surface charge**, is concentrated near the interface between two dielectrics. Let us denote the surface density of this charge by σ_b. In order to find this charge, it is best to proceed from formula (17.17). We construct at the interface between the two dielectrics a right cylinder with a base of area ΔS and height h (Fig. 81), and integrate both sides of Eq. (17.17) over the volume of this cylinder:

$$\int_V \rho_b \, dV = -\int_V \operatorname{div} \mathbf{P} \, dV. \tag{17.18}$$

The left-hand side of this equality represents the total charge in the cylinder, i.e. the surface charge $\sigma_b \Delta S$. With the help of the Gauss theorem, we can transform the right-hand side of this equation into a surface integral:

$$\int_V \operatorname{div} \mathbf{P} \, dV = \int_S \mathbf{P} \cdot d\mathbf{S} = \int_{S_2} \mathbf{P}_2 \cdot d\mathbf{S}_2 + \int_{S_1} \mathbf{P}_1 \cdot d\mathbf{S}_1, \tag{17.19}$$

where subscripts 1 and 2 correspond to the first and second dielectrics on different sides of the interface. The polarization flux of the vector $\mathbf{P}$ is the sum of fluxes through the bases and lateral surfaces of the cylinder. The flux through the lateral surface of the cylinder is taken equal to zero, since in the limit the height h of the cylinder tends to zero. For the positive normal to the interface, we choose the direction from the first dielectric to the second. Consequently, $d\mathbf{S}_2$ is directed along the positive direction of the normal, while $d\mathbf{S}_1$ is along the negative direction. Consequently,

$$\int_S \mathbf{P} \cdot d\mathbf{S} = P_{2n} \Delta S - P_{1n} \Delta S. \tag{17.20}$$

It should be recalled that the integral over the lateral surface is not taken into account. Considering the value of the integral on the left-hand side of Eq. (17.18), we finally obtain

$$\boxed{\sigma_b = -(P_{2n} - P_{1n}).} \tag{17.21a}$$

Hence, denoting the unit normal vector directed towards the second medium by $\mathbf{n}_2$, we can represent formula (17.21a) in the form

$$\boxed{\sigma_b = -\mathbf{n}_2 \cdot (\mathbf{P}_2 - \mathbf{P}_1).} \tag{17.21b}$$

It is worth noting that vacuum can also be treated as a dielectric with a polarization equal to zero. Formula (17.21a) can be applied to the boundary between a dielectric and the vacuum. In this case, we consider the outward normal to the dielectric as the positive normal, i.e. we assume the dielectric in (17.21a) to be medium *1* and put $P_{2n} = 0$. Consequently, we get [see (17.21a)]

$$\sigma_b = P_n, \tag{17.22}$$

www.EngineeringBooksPDF.com

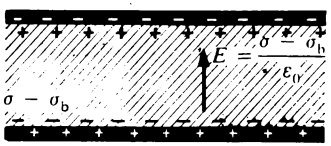


Fig. 82. Field in a capacitor in the presence of a dielectric

where P_n is the normal component of the dielectric polarization on its boundary with vacuum.

With the help of formulas (17.17) and (17.21), we can completely take into account the influence of a dielectric on an electric field. The strength of the field created by bound charges is calculated by the same formulas that are used for determining the strength of the field created in vacuum by free charges. In particular, the potential φ_d, created by the bound charges in a dielectric, is given by formulas (14.35) and (14.36) where the free charges have been replaced by bound charges:

$$\varphi_d = \frac{1}{4\pi\varepsilon_0}\int\limits_V \frac{\rho_b\,dV}{r} + \frac{1}{4\pi\varepsilon_0}\int\limits_S \frac{\sigma_b\,dS}{r}$$

$$= \frac{1}{4\pi\varepsilon_0}\int\limits_V \frac{-\operatorname{div}\mathbf{P}\,dV}{r} + \frac{1}{4\pi\varepsilon_0}\int\limits_S \frac{P_{1n}-P_{2n}}{r}\,dS. \qquad (17.23)$$

This potential is added to the potential created by free charges.

It is worthwhile to formulate once again in an explicit form the basic idea behind the influence of matter on field, which was traced by considering the example of conductors and dielectrics: *in the presence of an external electric field, the matter itself becomes the source of an electric field; consequently, the external field undergoes a change.*

Let us consider this process by taking the example of the field formation in a capacitor the space between whose plates is filled with a dielectric (Fig. 82). We shall assume that the capacitor plates carry a charge with surface density σ. If the capacitor plates are in vacuum, we get $E' = \sigma/\varepsilon_0$ [see (16.12)]. As a result of polarization of the dielectric, the field strength decreases. We determine the dielectric polarization with the help of formula (17.11), considering that $E \neq \sigma/\varepsilon_0$. In view of the homogeneity of the dielectric and the uniformity of the field between the charged parallel plates, we conclude that the dielectric polarization is uniform, i.e. there are no bound space charges. There are only bound surface charges with a surface density [see (17.22)]

$$\sigma_b = \varkappa\varepsilon_0 E, \qquad (17.24)$$

where E is the component of the field vector along the outward normal to the dielectric. It is well known that the field vector is directed from the positively charged plate to the negatively charged plate of the capacitor. Hence, it follows from (17.24) that the surface density of a bound charge is negative on the boundary with the positively charged plate and positive on the boundary with the negatively charged plate. Consequently, the strength of the field in a dielectric between the plates of a capacitor is identical to that between the same plates in vacuum, but with a surface density of charge equal to $\sigma - \sigma_b$. Thus, we can write the following equation for determining the unknown quantity:

$$E = (\sigma - \sigma_b)/\varepsilon_0 = (\sigma - \varkappa\varepsilon_0 E)/\varepsilon_0. \qquad (17.25)$$

www.EngineeringBooksPDF.com

The solution of this equation has the form

$$E = \sigma/[\varepsilon_0 (1 + \varkappa)]. \tag{17.26}$$

Electric displacement. Considering that bound charges are responsible for the field, we can obviously write Eq. (13.19) in the form

$$\operatorname{div} \mathbf{E} = \rho/\varepsilon_0 + \rho_b/\varepsilon_0. \tag{17.27}$$

Substituting into this equation the expression (17.17) for ρ_b, we get

$$\operatorname{div} (\varepsilon_0 \mathbf{E} + \mathbf{P}) = \rho. \tag{17.28}$$

The vector

$$\boxed{\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}} \tag{17.29}$$

is called the **displacement vector**. *It is not purely a field vector since it takes into account the polarization of the medium.* With the help of this vector, we can write Eq. (17.28) in the form

$$\boxed{\operatorname{div} \mathbf{D} = \rho.} \tag{17.30}$$

Recalling the meaning of the divergence of a vector, it can be concluded from (17.30) that it is advantageous to use **D**. It can be seen that *the only source of* **D** *are the free charges at the head and tail of this vector*. This vector is continuous at points without free charges, including points with bound charges. The variations in the field introduced by bound charges have already been taken into consideration in the vector **D** [see (17.29)].

Expressing **P** in (17.29) with the help of formula (17.11), we get

$$\mathbf{D} = (\varepsilon_0 + \varkappa\varepsilon_0) \mathbf{E} = \varepsilon \mathbf{E}, \quad \varepsilon = (1 + \varkappa) \varepsilon_0, \tag{17.31}$$

where ε is the **dielectric constant or permittivity.** The use of **D** considerably simplifies the analysis of the field in the presence of a dielectric. In addition to ε, another dimensionless quantity

$$\varepsilon_r = \varepsilon/\varepsilon_0, \tag{17.32}$$

called **relative permittivity** is also used.

Gauss' electrostatic theorem in the presence of dielectrics. Multiplying both sides of (17.30) by dV and integrating over the volume V, we obtain

$$\int_V \operatorname{div} \mathbf{D}\, dV = \int_V \rho\, dV. \tag{17.33}$$

The right-hand side of this equation represents the total charge Q inside the volume, while the left-hand side can be transformed into a surface integral with the help of Gauss' theorem. As a result, we get the formula

$$\boxed{\int_S \mathbf{D} \cdot d\mathbf{S} = Q,} \tag{17.34}$$

www.EngineeringBooksPDF.com

called **Gauss' theorem** for electrostatic fields in the presence of dielectrics. It is valid for any arrangement of dielectrics and boundary surfaces: a part or the entire volume may be filled with different dielectrics, and the surface S may either be in vacuum or cross the dielectrics.

Let us apply formula (17.34) to a point charge q located in an infinite homogeneous dielectric medium, and taking a sphere of radius r with its centre at the point where the point charge is located as the integration surface, we obtain Coulomb's law for a homogeneous dielectric medium:

$$\mathbf{E}=\frac{1}{4\pi\varepsilon}\frac{q}{r^2}\frac{\mathbf{r}}{r}. \tag{17.35}$$

The intensity of the field in a medium is ε_r times smaller than in vacuum. The potential of the point charge behaves in the same manner. Formula (17.26) shows that in the presence of a dielectric, the field strength between the plates of a capacitor also drops to $1/\varepsilon_r$ of its value in vacuum. The capacitance of the capacitor increases ε_r times.

Boundary conditions. Boundary conditions are relations between field vectors on different sides of the interface between two regions. This surface may separate substances with different properties, be the boundary between a body and vacuum or, in general, may just be an imaginary surface in a homogeneous medium. In all cases, boundary conditions can be used to determine the change in the field vector upon crossing this boundary. These conditions are derived with the help of field equations.

Boundary conditions for the normal component of vector D. Let us derive this condition in the same way as the boundary condition (17.21). In this case, however, we must proceed from Eq. (17.30) rather than from (17.17):

$$\boxed{\mathbf{D}_{2n}-\mathbf{D}_{1n}=\sigma,} \qquad \boxed{\mathbf{n}_2\cdot(\mathbf{D}_2-\mathbf{D}_1)=\sigma,} \tag{17.36}$$

where σ is the surface charge density at the boundary. The normal $\mathbf{n}_2$ is directed towards medium *2*. In particular, Eq. (17.36) can be used to obtain the field strength on the surface of a charged conductor. Taking the outward normal to the surface as the positive one, we must consider vacuum as medium *2* and conductor as medium *1* in formula (17.36). The field $\mathbf{E}$ inside the conductor is equal to zero, i.e. $D_{1n} = 0$. Consequently,

$$D_n = \sigma \tag{17.37}$$

or

$$E_n = \sigma/\varepsilon. \tag{17.38}$$

This formula is identical with formula (16.12) for vacuum, where ε_0 has been replaced by ε, i.e. the field strength on the surface of a conductor decreases in the presence of a dielectric to $1/\varepsilon_r$, i.e. $\varepsilon_0/\varepsilon$ of its initial value.

Formula (17.38) also provides a solution to the problem of the field in a parallel-plate capacitor, which is expressed by relation (17.26). In this case, it is not necessary to explicitly take into account the bound surface charges in the dielectric between the capacitor plates, as was done while deriving Eq. (17.26).

www.EngineeringBooksPDF.com

Boundary conditions for the tangential component of vector E. Let us construct a closed contour near the interface between dielectrics *1* and *2* (Fig. 83). In view of the potential nature of the electric field, the circulation of **E** around the closed circuit is equal to zero:

$$\oint_{ABCDA} \mathbf{E}\cdot d\mathbf{l} = 0. \tag{17.39}$$

The integrals over the segments BC and DA are arbitrarily small, since AB and CD are infinitely close to the interface. The signs of integrals over AB and

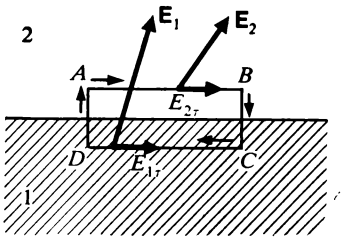


Fig. 83. To the derivation of the boundary condition for the tangential component of vector **E**

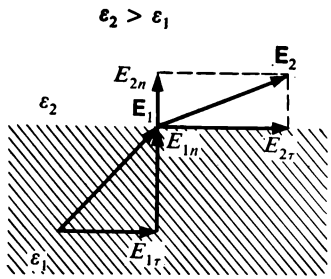


Fig. 84. Refraction of field lines at the interface between two dielectrics

CD are opposite in view of the fact that the integration is carried out in opposite directions. Hence [see (17.39)]

$$\boxed{E_{2\tau} - E_{1\tau} = 0.} \tag{17.40}$$

Refraction of field lines at the interface between dielectrics. Suppose that there are no free charges at the interface between two dielectrics. In this case,

$$\varepsilon_1 E_{1n} = \varepsilon_2 E_{2n}, \quad E_{1\tau} = E_{2\tau}. \tag{17.41}$$

If $\varepsilon_2 > \varepsilon_1$, then $E_{2n} < E_{1n}$, and hence the field lines behave as shown in Fig. 84. It can be seen in this figure that the field lines deviate from the normal when entering the dielectric with a higher permittivity.

Signs of bound charges at the interface between dielectrics. Let us consider the normal components of field and polarization vectors at the interface between dielectrics. With the help of formula (17.31) we can write formula (17.11) for dielectrics on both sides of the boundary in the following form (Fig. 85):

$$P_{2n} = (\varepsilon_2 - \varepsilon_0) E_{2n}, \quad P_{1n} = (\varepsilon_1 - \varepsilon_0) E_{1n}. \tag{17.42}$$

Let us write formula (17.21) for surface charge density by taking into account Eq. (17.32):

$$\sigma_b = P_{1n} - P_{2n} = \varepsilon_1 E_{1n} - \varepsilon_2 E_{2n} - \varepsilon_0 (E_{1n} - E_{2n}). \tag{17.43}$$

www.EngineeringBooksPDF.com

If there are no free charges on the surface, we get $\varepsilon_1 E_{1n} - \varepsilon_2 E_{2n} = 0$, and formula (17.43) can be simplified as follows:

$$\sigma_b = -\varepsilon_0 (E_{1n} - E_{2n}). \tag{17.44}$$

For the sake of definiteness, we shall assume as before that $\varepsilon_2 > \varepsilon_1$ and **E** is directed from the first medium into the second. It should be recalled that the normal directed towards the second medium is taken as the positive normal. In this case, E_{1n} and E_{2n} in (17.44) are positive and $E_{1n} > E_{2n}$. Hence the bound

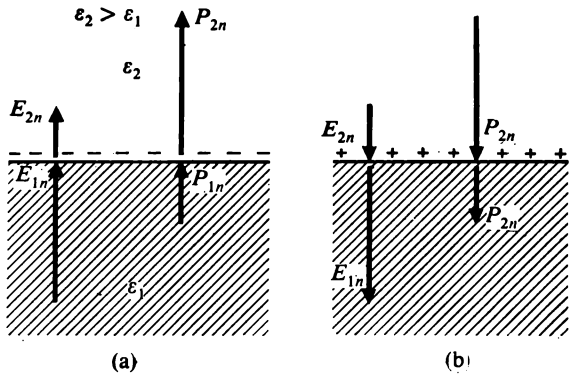


Fig. 85. Sign of the surface charge and the behaviour of normal components of the electric field and polarization vectors when the interface is crossed in opposite directions

charge at the boundary is negative (Fig. 85*a*). The quantities P_{1n} and P_{2n} are also positive and, consequently, $P_{2n} > P_{1n}$, as can be seen from (17.43) for $\sigma_b < 0$ (Fig. 85*a*).

Similar arguments can be applied to study the variation of normal components of the field and polarization vectors, as well as the sign of the surface charge density when the field vector is directed towards the dielectric with lower permittivity (Fig. 85*b*).

Method of images. When applied to dielectrics, the idea of this method is the same as for conductors (see Sec. 16).

Suppose that we have two very long dielectric media (having permittivities ε_1 and ε_2) with a plane interface. A point charge q is located in the first medium at a distance d from the interface. It is stated that the potential in the first medium is the same as due to charge q and its image $q' = q(\varepsilon_1 - \varepsilon_2)/(\varepsilon_1 + \varepsilon_2)$ located in the second medium at a distance d from the interface (see Fig. 86*a*). The calculations are carried out as if the permittivity of the media were equal to ε_1. The potential in the second medium is equal to the potential due to the charge $q'' = 2\varepsilon_2 q/(\varepsilon_1 + \varepsilon_2)$ located in place of charge q in the first medium (Fig. 86*b*), the calculations being carried out as if the permittivities of both

www.EngineeringBooksPDF.com

media were equal to ε_2. Thus, the potentials in the first and second media are

$$\varphi_1 = \frac{q}{4\pi\varepsilon_1}\left\{\frac{1}{\sqrt{(x+d)^2+y^2}} + \frac{\varepsilon_1-\varepsilon_2}{\varepsilon_1+\varepsilon_2}\frac{1}{\sqrt{(x-d)^2+y^2}}\right\},$$
$$\varphi_2 = \frac{q}{4\pi\varepsilon_2}\frac{2\varepsilon_2}{\varepsilon_1+\varepsilon_2}\frac{1}{\sqrt{(x+d)^2+y^2}}. \tag{17.45}$$

It can be easily verified that φ_1 and φ_2 satisfy the Laplace equation and the boundary conditions

$$\varepsilon_1\frac{\partial\varphi_1}{\partial x}\Big|_{x=0} = \varepsilon_2\frac{\partial\varphi_2}{\partial x}\Big|_{x=0}, \quad \frac{\partial\varphi_1}{\partial y}\Big|_{x=0} = \frac{\partial\varphi_2}{\partial y}\Big|_{x=0} = 0, \tag{17.46}$$

which express the continuity of the normal components of **D** and of the tangential components of **E**. Besides, the requirement that the potential be finite is also met:

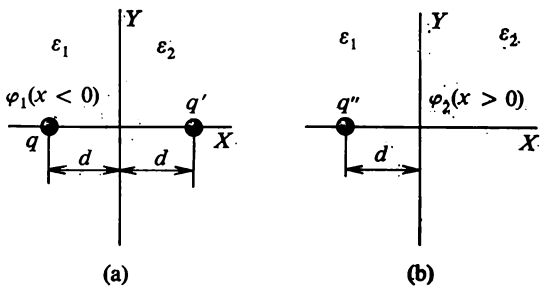


Fig. 86. Method of images applied to dielectrics

$$\varphi_1|_{x\to-\infty}\to 0, \quad \varphi_2|_{x\to+\infty}\to 0. \tag{17.47}$$

In accordance with the uniqueness theorem, formulas (17.45) give the required solution.

The force acting on charge q is equal to the force of interaction of this charge with the image $[(\varepsilon_1 - \varepsilon_2)/(\varepsilon_1 + \varepsilon_2)]\,q$, located at a distance $2d$ from q:

$$|F| = \frac{1}{4\pi\varepsilon_1}\left(\frac{\varepsilon_1-\varepsilon_2}{\varepsilon_1+\varepsilon_2}\right)\frac{q^2}{4d^2}. \tag{17.48}$$

For $\varepsilon_1 < \varepsilon_2$, F is negative, i.e. q is attracted towards the interface between the dielectrics. If $\varepsilon_1 > \varepsilon_2$, F is positive and hence q is repelled from the interface.

Dielectric sphere in a uniform field. With the help of the Laplace equation, let us find the field strength when a dielectric sphere is brought into an initially homogeneous electric field. If the linear dimensions of the plates in a parallel-plate capacitor are quite large, the field will be homogeneous to a high degree of accuracy in the inner parts away from the edges even for large separations between the plates. If the size of the plates is increased to infinity and at the same time the distance between them is also increased to infinity for a constant

www.EngineeringBooksPDF.com

surface charge density on the plates, a uniform electric field is created in the entire space. We place a conducting dielectric sphere in this space. It is clear that as a result of polarization the field strength near the sphere will change while it will remain unchanged at infinity. Let us determine the electric field strength in the entire space including the region inside the dielectric sphere.

We assume that a sphere of radius R consists of a dielectric with permittivity ε_1, and its surrounding space is filled with a dielectric with permittivity ε_2

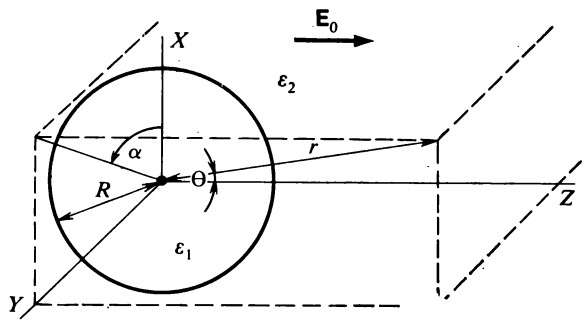


Fig. 87. Orientation of the coordinate system for a dielectric sphere in a uniform field

(Fig. 87). A homogeneous field is directed along the Z-axis. In view of the axial symmetry of the problem, it is convenient to use a spherical coordinate system with the polar axis along the Z-axis.

For a homogeneous dielectric with permittivity ε, the Poisson equation (15.14) has the form

$$\nabla^2\varphi = -\rho/\varepsilon. \tag{17.49}$$

This is obvious from a comparison of Eq. (15.10) for vacuum with Eq. (17.30) which has the following form for a homogeneous dielectric:

$$\operatorname{div} \mathbf{E} = \rho/\varepsilon. \tag{17.50}$$

In spherical coordinates, Poisson's equation can be written as follows:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\varphi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\varphi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\varphi}{\partial\alpha^2}=-\frac{\rho}{\varepsilon}, \tag{17.51}$$

where α is the axial angle. There are no free charges ($\rho = 0$) in this problem, and $\partial\varphi/\partial\alpha = 0$ on account of axial symmetry. Hence the problem is reduced to the solution of the Laplace equation

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\varphi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\varphi}{\partial\theta}\right)=0 \tag{17.52}$$

in the entire space under the following conditions:

(1) potential φ is continuous and finite everywhere;

(2) the normal components of vector $\mathbf{D} = -\varepsilon \operatorname{grad}\varphi$ are continuous at the interface, i.e. on the surface of the sphere;

www.EngineeringBooksPDF.com

(3) the tangential components of vector $\mathbf{E} = -\text{grad}\,\varphi$ are continuous on the surface of the sphere.

The quantities corresponding to the inner region of the sphere are denoted by subscript 1, while those corresponding to the space outside the sphere are denoted by subscript 2. The general solution of Eq. (17.52) is well known in mathematics. In the present case, it becomes much simpler. It can be found as a result of direct verification that the functions

$$\varphi_1 = A_1 r\cos\theta + A_2 r^{-2}\cos\theta,\quad \varphi_2 = -E_0 r\cos\theta + B_2 r^{-2}\cos\theta \tag{17.53a}$$

satisfy Eq. (17.52), where A_1, A_2 and B_2 are constants, and E_0 is the absolute value of the strength of the uniform field (at infinity).

Since φ_1 and φ_2 satisfy Eq. (17.52), they represent the potential if they satisfy all the conditions of the problem. The potential φ_1 corresponds to the inner region of the sphere, while φ_2 corresponds to the space outside it. It can be seen from (17.53a) that $\varphi_1 \to \infty$ as $r \to 0$. Hence, it can be assumed that $A_2 = 0$. The continuity equation for φ at the boundary has the form

$$A_1 R\cos\theta = -E_0 R\cos\theta + B_2 R^{-2}\cos\theta, \tag{17.53b}$$

whence

$$A_1 = B_2 R^{-3} - E_0. \tag{17.54}$$

The tangential component of the electric field vector $\mathbf{E}$ on the surface of the sphere is

$$E_\tau = E_\theta = -\left[\frac{1}{r}\frac{\partial\varphi}{\partial\theta}\right]_{r=R}. \tag{17.55}$$

The condition $E_{1\theta} = E_{2\theta}$ is satisfied if Eq. (17.53b) is valid, i.e. A_1 and B_2 are connected through relation (17.54).

The normal components of the electric field vector are given by

$$\begin{aligned} E_{1n} &= E_{1r} = -(\partial\varphi_1/\partial r)_{r=R} = -A_1\cos\theta, \\ E_{2n} &= E_{2r} = -(\partial\varphi_2/\partial r)_{r=R} = E_0\cos\theta + 2B_2 R^{-3}\cos\theta. \end{aligned} \tag{17.56}$$

It follows from the condition $\varepsilon_1 E_{1r} = \varepsilon_2 E_{2r}$ that

$$A_1 = -(\varepsilon_2/\varepsilon_1)\,(E_0 + 2B_2 R^{-3}). \tag{17.57}$$

The solution of the system of equations (17.54) and (17.57) is

$$A_1 = -\frac{3\varepsilon_2}{\varepsilon_1 + 2\varepsilon_2} E_0,\quad B_2 = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + 2\varepsilon_2} R^3 E_0. \tag{17.58}$$

The potentials inside and outside the sphere are

$$\varphi_1 = -\frac{3\varepsilon_2}{\varepsilon_1 + 2\varepsilon_2} E_0 r\cos\theta, \tag{17.59}$$

$$\varphi_2 = -\left(1 - \frac{R^3}{r^3}\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + 2\varepsilon_2}\right) E_0 r\cos\theta. \tag{17.60}$$

www.EngineeringBooksPDF.com

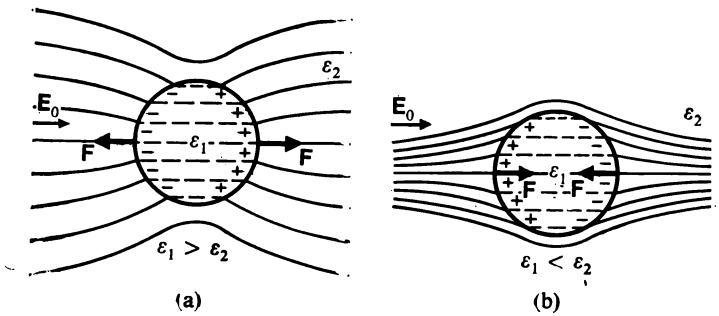


Fig. 88. Field lines of displacement vector **D** for a dielectric sphere in a uniform external field

Obviously, the field inside the sphere is constant and parallel to the Z-axis:

$$E_{1z} = -\frac{\partial \varphi_1}{\partial z} = -\frac{\partial \varphi_1}{\partial (r \cos \theta)} = \frac{3\varepsilon_2}{\varepsilon_1 + 2\varepsilon_2} E_0. \tag{17.61}$$

This field is the sum of the external field and the field created by the bound charges appearing on the surface of the sphere. Consequently, the strength of the field created inside the sphere by bound charges is

$$E_b = E_{1z} - E_0 = (\varepsilon_2 - \varepsilon_1) E_0/(\varepsilon_1 + 2\varepsilon_2). \tag{17.62}$$

This field is constant and parallel to the Z-axis. The charge distribution on the surface of the sphere, which leads to a constant field strength inside the sphere, is given by formula (16.75). Hence it can be concluded that the field (17.62) is created by bound charges on the surface of the sphere, the charge density varying with angle θ in the same way as in formula (16.79), i.e. $\sigma \propto \cos \theta$.

It can be seen from (17.62) that for $\varepsilon_1 > \varepsilon_2$ the field $\mathbf{E}_b$ is directed against $\mathbf{E}_0$. Consequently, the field inside the sphere is less than the initial uniform field. If $\varepsilon_2 > \varepsilon_1$, the field $\mathbf{E}_b$ is in the same direction as $\mathbf{E}_0$, and the field inside the sphere is stronger than in the surrounding medium. The lines of vector **D** for the case (a) $\varepsilon_1 > \varepsilon_2$ and (b) $\varepsilon_1 < \varepsilon_2$ are shown in Fig. 88, as well as the signs of the bound charges appearing on the surface of the sphere. It should be noted that Fig. 88 shows the lines of vector **D** rather than **E**, since it is the vector **D** that is continuous in the absence of bound charges. While plotting the lines of vector **E**, we must change their density on the surface of the sphere where bound charges exist.

Polarization (or bound) charges appear at the sites where the polarization changes.
In the presence of an external electric field, material bodies themselves become the sources of an electric field, as a result of which the field changes. In this case, electric fields behave with respect to their sources as if the latter were in vacuum and there were no material bodies.
Polarization is the process of formation of dipole moments in macroscopic volumes of a dielectric.

The normal component of an electric field vector undergoes a discontinuity at the interface between different dielectrics. Hence, the field lines break.

www.EngineeringBooksPDF.com

Example 17.1. *Find the bound charges, polarization and the field induced by a point charge q placed at the centre of two concentric spheres of radii a_1 and a_2 respectively. The spherical layer is filled with a substance having permittivity* ε (Fig. 89).

The field is spherically symmetric. Choosing for S the surface of a sphere of radius r having its centre at the point where charge q is located, we determine from the Gauss formula

$$\int_S \mathbf{D}\cdot d\mathbf{S} = D_r 4\pi r^2 = q$$

the electric displacement

$$D_r = \frac{1}{4\pi}\frac{q}{r^2},$$

This displacement is continuous in the entire space. The strength of the electric field is given by

$$\begin{aligned} E_r &= \frac{D_r}{\varepsilon_0} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2} && \text{for } r < a_1, \\ E_r &= \frac{D_r}{\varepsilon} = \frac{1}{4\pi\varepsilon}\frac{q}{r^2} && \text{for } a_1 < r < a_2, \\ E_r &= \frac{D_r}{\varepsilon_0} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2} && \text{for } a_2 < r \end{aligned} \tag{17.63}$$

and undergoes a discontinuity at the surfaces of the spherical layer for $r = a_1$ and $r = a_2$.

The polarization is given by the expressions

$$P_r = D_r - \varepsilon_0 E_r = \begin{cases} 0 & \text{for } r < a_1, \\ \dfrac{(\varepsilon-\varepsilon_0)\,q}{4\pi\varepsilon r^2} & \text{for } a_1 < r < a_2, \\ 0 & \text{for } a_2 < r. \end{cases} \tag{17.64}$$

Consequently, the surface density of bound charges is

$$\begin{aligned} \sigma_{b1} &= -P_r\,(r=a_1) = -(\varepsilon-\varepsilon_0)\,q/(4\pi\varepsilon a_1^2), \\ \sigma_{b2} &= P_r\,(r=a_2) = (\varepsilon-\varepsilon_0)\,q/(4\pi\varepsilon a_2^2). \end{aligned} \tag{17.65}$$

The bound charges on the surface of the spherical layer are calculated from the formula

$$q_{b1} = 4\pi a_1^2\sigma_{b1} = -(\varepsilon-\varepsilon_0)\,q/\varepsilon,\quad q_{b2} = 4\pi a_2^2\sigma_{b2} = (\varepsilon-\varepsilon_0)\,q/\varepsilon.$$

These charges are equal in magnitude but opposite in sign.

The volume density of bound charges is equal to zero everywhere, since

$$\rho_b = -\operatorname{div}\mathbf{P} = -\frac{1}{r_2}\frac{\partial}{\partial r}(r^2 P_r) = 0. \tag{17.66}$$

The field inside the spherical layer is created by the point charge q and the bound charge q_{b1} on the inner surface of the layer. The bound charge on the outer surface of the layer does not create an electric field in the volume bounded by it. Hence the field strength of the point charge q in the spherical layer is reduced by the value of the field created by the bound charge $q_{b1} = -(\varepsilon-\varepsilon_0)\,q/\varepsilon$. As $a_1 \to 0$, we find that the point charge q in the dielectric behaves as an effective point charge

$$q_{eff} = q + q_{b1} = \varepsilon_0 q/\varepsilon. \tag{17.67}$$

This results in a weakening of the electric field in the dielectric.

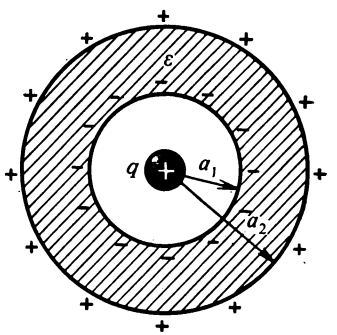


Fig. 89. A point charge surrounded by a concentric layer of a dielectric

www.EngineeringBooksPDF.com

Sec. 18. Energy of Electrostatic Field

The interaction energy is considered as well as the self-energy of charges and its relation with the energy density of an electric field. Formulas for the energy of charged conductors and of a dielectric in an external field are derived.

Energy of interaction between discrete charges. Let us suppose that we have charged spheres of a very small diameter which is less than the distance between the centres of the spheres. The charge distribution in the spheres is spherically symmetric. From the physical meaning of formula (14.32) we may conclude that the quantity

$$W' = \frac{1}{4\pi\varepsilon_0}\frac{Q_1 Q_2}{r} \tag{18.1}$$

is equal to the work done upon increasing the distance between the charges Q_1 and Q_2 from r to infinity. This work is positive when the charges have the same sign and repulsive forces are acting between them. Unlike charges attract each other, and the work is negative. In the latter case the work must be accomplished at the expense of external energy sources. Consequently, in accordance with the general definition, (18.1) is the energy of interaction between the charged spheres. Since both charges appear in formula (18.1) symmetrically, it is expedient to write it in the form

$$W' = \frac{1}{2}\left(\frac{1}{4\pi\varepsilon_0}\frac{Q_2}{r}Q_1 + \frac{1}{4\pi\varepsilon_0}\frac{Q_1}{r}Q_2\right) = \frac{1}{2}\left(\varphi_1' Q_1 + \varphi_2' Q_2\right), \tag{18.2}$$

where φ_1' is the potential created by the second charge at the centre of the first sphere, while φ_2' is the potential due to the first charge at the centre of the second sphere.

Formula (18.2) can be easily generalized for the case of several charged spheres with charges Q_i:

$$W' = \frac{1}{2}\sum_{i \neq j}\frac{1}{4\pi\varepsilon_0}\frac{Q_i Q_j}{r_{ij}} = \frac{1}{2}\sum_i \varphi_i' Q_i'. \tag{18.3}$$

It gives the interaction energy of a system of charges.

Energy of interaction for a continuous distribution of charges. Suppose that a volume element dV contains a charge $dQ = \rho\, dV$. In order to find the energy of interaction between elements of charge dQ, we can apply formula (18.3),

www.EngineeringBooksPDF.com

going over from the summation to the integration in it:

$$\boxed{W=\frac{1}{2}\int_V \varphi\rho\, dV,} \tag{18.4}$$

where φ is the potential at a point in the volume element dV.

Self-energy. At first sight, formula (18.4) seems to be similar to (18.3). However, these two formulas differ in principle. *Formula* (18.3) *takes into account only the energy of interaction between charged spheres and disregards the energy of interaction between elements of charge in each sphere.* Formula (18.4) takes into consideration both the energy of interaction between the spheres and the energy of interaction between elements of charge in each sphere, called the self-energy (intrinsic energy) of a charged sphere. In calculating the energy of interaction between charged spheres, formula (18.4) is reduced to the integrals over volumes V_i of the spheres:

$$W=\frac{1}{2}\int_V \varphi\rho\, dV=\sum_i \frac{1}{2}\int_{V_i} \varphi_i\rho\, dV. \tag{18.5}$$

At any point in the volume of the ith sphere, the potential φ_i is the sum of two components: the potential $\varphi_i^{(1)}$ created by the charges of other spheres and the potential $\varphi_i^{(self)}$ created by the charges of the ith sphere:

$$\varphi_i=\varphi_i^{(1)}+\varphi_i^{(self)} \tag{18.6}$$

In this case, [see (18.5)]

$$W=\sum_i \frac{1}{2}\int_{V_i} \varphi_i^{(1)}\rho\, dV+\sum_i \frac{1}{2}\int_{V_i} \varphi_i^{(self)}\rho\, dV. \tag{18.7}$$

Since the charge distribution in the spheres is spherically symmetric, we have

$$\int_{V_i} \varphi_i^{(1)}\rho\, dV=\varphi_i' Q_i, \tag{18.8}$$

where φ_i' is the potential at the centre of a sphere and $Q_i=\int_{V_i}\rho\, dV$ is the total charge of the sphere. The proof of (18.8) is similar to that of the equivalence of the electric field generated by a spherically symmetric charge distribution in a sphere and the field created by the corresponding point charge located at the centre of the sphere (for the region outside the sphere). Now we can write (18.7) in the form

$$W=\frac{1}{2}\sum_i \varphi_i' Q_i+\frac{1}{2}\sum_i \int_{V_i} \varphi_i^{(self)}\rho\, dV=W'+\sum_i W_i^{(self)} \tag{18.9}$$

where W_i' is given by formula (18.3).

www.EngineeringBooksPDF.com

The self-energies $W_i^{(self)}$ of the spheres depend on the laws of charge distribution in the spheres and on the magnitudes of the charges. Let, for example, the charge Q be uniformly distributed over the surface of a sphere. In this case, the potential is defined by formula (16.28), and hence

$$W^{(self)} = \frac{1}{8\pi\varepsilon_0}\frac{Q^2}{R}. \tag{18.10}$$

As $R \to 0$, the value of $W^{(self)} \to \infty$. This means that the **self-energy of a point charge is equal to infinity.** This creates serious difficulties when the concept of point charges is being used.

Thus, *formula* (18.3) *can be applied for analyzing interaction between point charges since it does not contain their infinite self-energies.* Formula (18.4) for a continuous charge distribution takes into account the entire interaction energy, while formula (18.3), only a part of this energy. Therefore, formula (18.4) is more complete and informative in comparison with formula (18.3).

Energy density of a field. Using the equation

$$\operatorname{div} \mathbf{D} = \rho, \tag{18.11}$$

we write (18.4) in the form

$$W = \frac{1}{2}\int_V \varphi \operatorname{div} \mathbf{D}\, dV. \tag{18.12}$$

Taking into account the formula

$$\varphi \operatorname{div} \mathbf{D} = -\mathbf{D} \operatorname{grad} \varphi + \operatorname{div} (\varphi \mathbf{D}), \tag{18.13}$$

of vector calculus, we represent (18.12) as the sum of two integrals:

$$W = \frac{1}{2}\int_V \mathbf{E}\cdot\mathbf{D}\, dV + \frac{1}{2}\int_V \operatorname{div} (\varphi \mathbf{D})\, dV, \tag{18.14}$$

where $\mathbf{E} = -\operatorname{grad} \varphi$. The second integral in (18.14), in accordance with the Gauss theorem, is written as

$$\int_V \operatorname{div} (\varphi \mathbf{D})\, dV = \int_S \varphi \mathbf{D}\cdot d\mathbf{S}, \tag{18.15}$$

where S is a closed surface enveloping the volume V. It is assumed that all charges are located in a finite region of space. At large distances r from the charges, $\varphi \propto 1/r$, $D \propto 1/r^2$, i.e. $\varphi D \propto 1/r^3$. The area S of the surface increases in proportion to r^2. Consequently, integral (18.15) is of the order of $\varphi DS \propto 1/r$ and tends to zero as the surface of integration approaches infinity. Hence, for the entire space, formula (18.14) becomes

$$\boxed{W = \frac{1}{2}\int \mathbf{E}\cdot\mathbf{D}\, dV.} \tag{18.16}$$

The energy W calculated by formulas (18.16) and (18.4) has the same value, but the physical contents of these formulas are quite different. Suppose that the

www.EngineeringBooksPDF.com

charges are located in thin surface layers of the spheres. In this case, integral (18.4) is reduced to the sum of integrals over the surface layers of the spheres, since in the space between the spheres it is equal to zero. On the other hand, integral (18.16) is reduced to the integral over the volume between the spheres where the field **E** is contained. Consequently, in (18.4) the carriers of energy are charges, and the energy is assumed to be localized on charges. In (18.16), the carrier of energy is the electric field, and the energy is assumed to be localized in the entire space containing the electric field. The density of electric energy [see (18.16)] is given by

$$\boxed{w = \frac{1}{2}\mathbf{E}\cdot\mathbf{D}.} \tag{18.17}$$

Thus, the energy density in (18.17) is positive since $\mathbf{E}\cdot\mathbf{D} = \varepsilon E^2 > 0$. Consequently, the total energy in (18.16) and (18.4) is positive. However, the interaction energy (18.3) of discrete charges can be either positive or negative. The reason behind this is clear from equality (18.9), which can be represented in the form

$$W' = W - \sum_i W_i^{(\text{self})}. \tag{18.18}$$

Thus, *the interaction energy of discrete charges is positive if their intrinsic energy (which is always positive) is less than the total energy of the field and negative if their intrinsic energy is higher than the total energy of the field.*

Suppose that all charges, with the exception of one, are fixed at their places. Then the energy of interaction between this charge and the remaining charges is called its potential energy. It follows from what was said above that it is a part of the energy of the electric field. A change in the potential energy is associated with the change in the energy of the field. *The law of conservation of energy for a particle in a potential field, which states that the sum of its kinetic and potential energies is constant, indicates that a decrease in the kinetic energy of the particle is accompanied by the corresponding increase in the energy of the field, and vice versa.*

Expression (18.17) is formulated in the local form and defines the energy density as a function of the electric field strength and properties of the medium at a given point, which are taken into account by the displacement **D**. Obviously, *the validity of this formula cannot depend on the way in which the field is created at a given point. Hence, expression* (18.17) *is valid not only for constant fields but also for varying fields.* In other words, this formula expresses the **energy density of an electric** as well as an electrostatic **field.**

Energy of the field of surface charges. Since formula (18.17) does not depend on the nature of charges which are the sources of the field, it is also valid for surface charges. Formula (18.16) also gives the total energy of the field irrespective of the nature of the charges creating this field. Therefore, formula (18.16) takes into account surface charges as well as volume charges.

www.EngineeringBooksPDF.com

For surface charges, formula (18.4) assumes a somewhat different form. This change, however, is self-evident. The integrand in (18.4) is equal to $\varphi\rho\, dV = \varphi\, dq$ and has the meaning of the potential energy possessed by the element of charge dq located at a point with potential φ. This potential energy does not depend on whether dq is the element of a volume charge or a surface charge. Consequently, expression (18.4) is applicable to surface charges as well, but in this case $dq = \sigma\, dS$, and we must integrate over all surfaces S which contain charges. Thus, for surface charges, formula (18.4) becomes

$$\boxed{W=\frac{1}{2}\int_V \varphi\rho\, dV+\frac{1}{2}\int_S \varphi\sigma\, dS.} \tag{18.19}$$

All that has been said about the interaction energy and self-energy is also valid for surface charges. We must only take into account their contribution to the total as well as to the intrinsic energy. This circumstance has already been used in deriving the formula for self-energy [see (18.10)].

Energy of charged conductors. Since there are only surface charges on conductors and the potential at different points on the surface of a conductor has the same constant value, formula (18.18) assumes the form

$$W=\frac{1}{2}\int_S \varphi\sigma\, dS=\frac{1}{2}\sum_i \int_{S_i} \varphi_i\sigma_i\, dS_i$$

$$=\frac{1}{2}\sum_i \varphi_i \int_{S_i} \sigma_i\, dS_i=\frac{1}{2}\sum_i \varphi_i Q_i. \tag{18.20a}$$

Substituting expression (16.42) into this formula, we obtain the relation

$$W=\frac{1}{2}\sum_{i,j} \alpha_{ij}Q_iQ_j. \tag{18.20b}$$

Using (16.45), we transform (18.20a) as follows:

$$W=\frac{1}{2}\sum_{i,j} C_{ij}\varphi_i\varphi_j. \tag{18.20c}$$

From (18.20a), we have

$$W=\frac{1}{2}Q(\varphi_1-\varphi_2)=\frac{1}{2}\frac{Q^2}{C}, \tag{18.20d}$$

where $C = Q/(\varphi_1 - \varphi_2)$ is the capacitance of a capacitor and Q is the charge on one of the plates.

Energy of a dipole in an external field. This energy is equal to the sum of the energies of dipole charges (see Fig. 77):

$$W = q[\varphi(\mathbf{r}+\mathbf{l}) - \varphi(\mathbf{r})]. \tag{18.21}$$

www.EngineeringBooksPDF.com

Let us expand $\varphi(\mathbf{r}+\mathbf{l})$ into a series in $\mathbf{l}$:

$$\varphi(\mathbf{r}+\mathbf{l}) = \varphi(\mathbf{r}) + l_x\frac{\partial\varphi}{\partial x} + l_y\frac{\partial\varphi}{\partial y} + l_z\frac{\partial\varphi}{\partial z} + \cdots$$
$$= \varphi(r) + (l_xE_x + l_yE_y + l_zE_z) = \varphi(\mathbf{r}) - \mathbf{l}\cdot\mathbf{E}, \qquad (18.22)$$

where we retained only the first-order terms in l due to an extremely small value of l. Formula (18.21) assumes the form

$$\boxed{W = -\mathbf{p}\cdot\mathbf{E}.} \qquad (18.23)$$

Energy of a dielectric in an external field. The dipole moment of the volume element dV of a body is $d\mathbf{p} = \mathbf{P}\,dV$. The energy of this element in an external field $\mathbf{E}$ is [see (18.23)] $dW = -\mathbf{P}\cdot\mathbf{E}\,dV$. It may seem that the energy of a dielectric body is equal to the integral of dW over the volume of the body. This, however, is wrong. As a matter of fact, each polarized volume element dV of a dielectric body becomes a source of electric field and thus appears twice during the calculation of energy: once as a dipole in an external field and the second time as a source of the field in which other dipoles are located.

Hence, in determining the energy of the field, it is convenient to proceed from its total energy. Besides, let us assume that the dielectric is homogeneous and fills the entire space, which considerably simplifies mathematical calculations.

Let an electrostatic field be created by a certain charge distribution in free space. As usual, we assume that the charges are located in a finite region of space. We denote by $\mathbf{E}_0$ and $\mathbf{D}_0 = \varepsilon_0\mathbf{E}_0$ the vectors of the field created by the distribution of charges in free space. The total energy of the field [see (18.16)] is

$$W_0 = \frac{1}{2}\int \mathbf{E}_0\cdot\mathbf{D}_0\,dV, \qquad (18.24)$$

where the integral is extended over the entire space. Let us now suppose that the entire space is filled with a dielectric medium, the charges as the sources of field remaining unchanged. The field in the entire space varies. We denote by ε, $\mathbf{E}$ and $\mathbf{D} = \varepsilon\mathbf{E}$ the permittivity and the field vectors in the medium. After the space has been filled with the dielectric, the total energy becomes

$$W = \frac{1}{2}\int \mathbf{E}\cdot\mathbf{D}\,dV. \qquad (18.25)$$

Consequently, the energy of the dielectric placed in an external field $\mathbf{E}_0$ is given by

$$W_d = W - W_0 = \frac{1}{2}\int (\mathbf{E}\cdot\mathbf{D} - \mathbf{E}_0\cdot\mathbf{D}_0)\,dV. \qquad (18.26)$$

Upon filling the entire space with a homogeneous dielectric having permittivity ε, the field strength at all points of the field decreases by a factor of $\varepsilon/\varepsilon_0$. Consequently,

$$\mathbf{E} = \varepsilon_0\mathbf{E}_0/\varepsilon. \qquad (18.27)$$

www.EngineeringBooksPDF.com

Hence, the integrand in (18.26) can be transformed as follows:

$$\mathbf{E}\cdot\mathbf{D}-\mathbf{E}_0\cdot\mathbf{D}_0=\varepsilon E^2-\varepsilon_0 E_0^2=-(\varepsilon-\varepsilon_0)\frac{\varepsilon_0}{\varepsilon}E_0^2=-\mathbf{P}\cdot\mathbf{E}_0, \tag{18.28}$$

where

$$(\varepsilon-\varepsilon_0)\frac{\varepsilon_0}{\varepsilon}\mathbf{E}_0=(\varepsilon-\varepsilon_0)\mathbf{E}=\mathbf{P}. \tag{18.29}$$

Then [see (18.26)]

$$W_{\mathrm{d}}=-\frac{1}{2}\int \mathbf{P}\cdot\mathbf{E}_0\,dV. \tag{18.30}$$

It can be shown that formula (18.30) is also valid for the energy of a dielectric of a finite size in an external field $\mathbf{E}_0$.

Formula (18.30) can be used to obtain the energy of a dielectric body of permittivity ε_2, located in the medium whose permittivity is ε_1. Let us write formula (18.30) for the energy of a dielectric body with permittivity ε_1:

$$W_{\mathrm{d}1}=-\frac{1}{2}\int(\varepsilon_1-\varepsilon_0)\mathbf{E}_1\cdot\mathbf{E}_0\,dV, \tag{18.31}$$

where $\mathbf{E}_1$ is the field strength in the body. In order to simplify calculations, we assume that the dielectric fills the entire space. The energy of the dielectric having permittivity ε_2 is expressed, by analogy with (18.31), by

$$W_{\mathrm{d}2}=-\frac{1}{2}\int(\varepsilon_2-\varepsilon_0)\mathbf{E}_2\cdot\mathbf{E}_0\,dV. \tag{18.32}$$

Hence it follows that the difference in the energies of dielectrics having permittivities ε_2 and ε_1 is

$$W_{\mathrm{d}21}=W_{\mathrm{d}2}-W_{\mathrm{d}1}=-\frac{1}{2}\int[(\varepsilon_2-\varepsilon_0)\mathbf{E}_2\cdot\mathbf{E}_0-(\varepsilon_1-\varepsilon_0)\mathbf{E}_1\cdot\mathbf{E}_0]\,dV. \tag{18.32a}$$

Transforming the integrand with the help of formulas

$$\mathbf{E}_2=\varepsilon_0\mathbf{E}_0/\varepsilon_2,\quad \mathbf{E}_1=\varepsilon_0\mathbf{E}_0/\varepsilon_1, \tag{18.33}$$

we find

$$\begin{aligned}(\varepsilon_2-\varepsilon_0)\mathbf{E}_2\cdot\mathbf{E}_0-(\varepsilon_1-\varepsilon_0)\mathbf{E}_1\cdot\mathbf{E}_0&=\left[\frac{\varepsilon_0}{\varepsilon_2}(\varepsilon_2-\varepsilon_0)-\frac{\varepsilon_0}{\varepsilon_1}(\varepsilon_1-\varepsilon_0)\right]E_0^2\\&=(\varepsilon_2-\varepsilon_1)\frac{\varepsilon_0^2}{\varepsilon_1\varepsilon_2}E_0^2=(\varepsilon_2-\varepsilon_1)\mathbf{E}_2\cdot\mathbf{E}_1.\end{aligned} \tag{18.34}$$

Then (18.32) becomes

$$\boxed{W_{\mathrm{d}21}=-\frac{1}{2}\int(\varepsilon_2-\varepsilon_1)\mathbf{E}_2\cdot\mathbf{E}_1\,dV,} \tag{18.35}$$

where $W_{\mathrm{d}21}$ is the energy of the dielectric of permittivity ε_2 placed in the medium of permittivity ε_1, in which the field $\mathbf{E}_1$ is created by fixed free charges. It can be

www.EngineeringBooksPDF.com

shown that this formula is valid for a finite dielectric as well, if in (18.35) we assume that integration is performed over the volume of the dielectric. In this case, $\mathbf{E}_1$ is the strength of the field which would exist in the volume of the dielectric if its permittivity were equal to the permittivity ε_1 of the surrounding medium, and $\mathbf{E}_2$ is the field in the volume of the dielectric after it has been introduced into the field provided that the charges creating the field are fixed. Formula (18.35) is important for understanding the nature of forces acting on dielectrics.

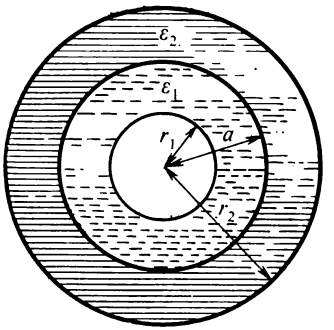


Fig. 90. A two-layer cylindrical or spherical capacitor

Formula (18.35) leads to the following important conclusion: *an increase in the permittivity of the medium leads to a decrease in the total energy of the field.* The proof of this statement can be obtained as follows. Suppose that the strength of the initial field is $\mathbf{E}_1 = \mathbf{E}$ and the permittivity of the medium is ε_1. If the permittivity of the medium increases by $\delta\varepsilon = \varepsilon_2 - \varepsilon_1$, the field strength becomes $\mathbf{E}_2 = \mathbf{E} + \delta\mathbf{E}$, and hence the change in the energy is given by

$$\delta W = -\frac{1}{2}\int \delta\varepsilon E^2\, dV \tag{18.36}$$

(we neglected the term $\delta\varepsilon\delta\mathbf{E}\cdot\mathbf{E}$ of the higher order of smallness). Formula (18.36) proves the above statement.

The intrinsic [self-] energy of a charge is the energy of interaction of different elements of the charge with each other. The intrinsic energy of a point charge is infinite.
The energy of interaction of discrete charges is the total energy of the field minus the intrinsic energy of the charges. It is positive when their intrinsic energy [which is always positive] is less than the total energy of the field and negative in the opposite case.
The law of conservation of energy for a particle in a potential field, which establishes the constancy of the sum of its kinetic and potential energies, indicates that a decrease in the kinetic energy of the particle is accompanied by the corresponding increase in the field energy, and vice versa. An increase in the permittivity of the medium leads to a decrease in the total energy of the field.

What determines the difference between the coefficients in the formulas for the dipole energy [see (18.23)] and the energy of a dielectric [see (18.30)]?

Example 18.1. *Find the energy accumulated in a cylindrical two-layer capacitor over length l. The parameters of the capacitor are given in* Fig. 90.

Assuming that the inner capacitor plate contains the charge Q over the length l and applying the Gauss theorem to the cylindrical surface of radius r, coaxial with the capacitor, we find the expression for the radial component of the field:

$$E_r = \begin{cases} \dfrac{1}{2\pi l\varepsilon_1}\dfrac{Q}{r} & \text{for } r_1 < r < a, \\ \dfrac{1}{2\pi l\varepsilon_2}\dfrac{Q}{r} & \text{for } a < r < r_2, \\ 0 & \text{for } r_2 < r < \infty. \end{cases}$$

www.EngineeringBooksPDF.com

The energy of the field is found by the formula

$$W = \frac{1}{2} \int \mathbf{E} \cdot \mathbf{D} \, dV$$

which in this case assumes the form

$$W = \frac{1}{2} \int_0^l dl \int_{r_1}^a \left(\frac{Q}{2\pi l} \right)^2 \frac{1}{\varepsilon_1} \frac{1}{r^2} 2\pi r \, dr + \frac{1}{2} \int_0^l dl \int_a^{r_2} \left(\frac{Q}{2\pi l} \right)^2 \frac{1}{\varepsilon_2} \frac{1}{r^2} 2\pi r \, dr$$

$$= \frac{Q^2}{4\pi l} \left(\frac{1}{\varepsilon_1} \ln \frac{a}{r_1} + \frac{1}{\varepsilon_2} \ln \frac{r_2}{a} \right).$$

Sec. 19. Forces in an Electric Field

Forces acting on charges, conductors and dielectrics in an electric field are considered. The emergence of volume and surface forces is analyzed.

Nature of forces. All forces in an electrostatic field are ultimately forces acting on a charge.

Force acting on a point charge. This force is equal to

$$\boxed{\mathbf{F} = q\mathbf{E} = -q \operatorname{grad} \varphi.} \tag{19.1}$$

Force acting on a continuously distributed charge. This force is equal to

$$\boxed{d\mathbf{F} = \rho \mathbf{E} \, dV.} \tag{19.2}$$

Consequently, the volume charge density is equal to

$$\boxed{\mathbf{f} = \frac{d\mathbf{F}}{dV} = \rho \mathbf{E} = -\rho \operatorname{grad} \varphi.} \tag{19.3}$$

Force acting on a dipole. This force is equal to the sum of forces applied to the dipole charges (Fig. 91):

$$\mathbf{F} = \mathbf{F}_{(+)} + \mathbf{F}_{(-)} = q \, [\mathbf{E}(\mathbf{r} + \mathbf{l}) - \mathbf{E}(\mathbf{r})]. \tag{19.4}$$

www.EngineeringBooksPDF.com

Here, we can express $\mathbf{E}(\mathbf{r}+\mathbf{l})$ as a series in l_x, l_y, l_z, confining ourselves to linear terms only:

$$\mathbf{E}(\mathbf{r}+\mathbf{l}) = \mathbf{E}(\mathbf{r}) + l_x \frac{\partial \mathbf{E}(\mathbf{r})}{\partial x} + l_y \frac{\partial \mathbf{E}(\mathbf{r})}{\partial y} + l_z \frac{\partial \mathbf{E}(\mathbf{r})}{\partial z} + \cdots = \mathbf{E}(\mathbf{r}) + (\mathbf{l}\cdot\nabla)\,\mathbf{E}(\mathbf{r}), \quad (19.5)$$

where

$$(\mathbf{l}\cdot\nabla) = l_x \frac{\partial}{\partial x} + l_y \frac{\partial}{\partial y} + l_z \frac{\partial}{\partial z}.$$

Taking (19.5) into account, we can write formula (19.4) in the form

$$\boxed{\mathbf{F} = (\mathbf{p}\cdot\nabla)\,\mathbf{E}.} \quad (19.6)$$

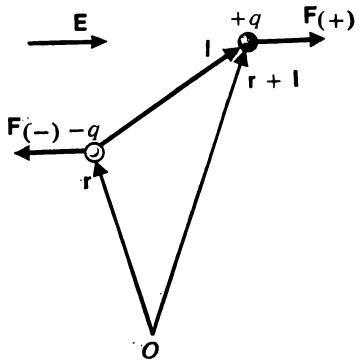


Fig. 91. Force and the moment of force acting on a dipole

The force acting on a dipole in a uniform field is equal to zero, since the dipole charges are subjected to equal and opposite forces.

Moment of force acting on a dipole. Forces applied to the charges of a dipole (see Fig. 91) form a couple with moment

$$\boxed{\mathbf{M} = \mathbf{p}\times\mathbf{E}.} \quad (19.7)$$

Volume forces acting on a dielectric. The forces applied to a volume element dV of a dielectric is equal to the sum of forces acting on the elementary dipoles in this volume. Consequently, formula (19.6) assumes the form

$$d\mathbf{F} = \sum_{\Delta V} \mathbf{F}_i = \sum_{\Delta V} (\mathbf{p}_i\cdot\nabla)\,\mathbf{E}_i, \quad (19.8)$$

where ΔV indicates that the summation is carried out over all the elementary dipoles in the volume ΔV. On the macroscopic scale, field strength $\mathbf{E}$ is assumed to be a slowly varying quantity. Hence we can replace $\mathbf{E}_i$ in the sum (19.8) by the quantity $\mathbf{E}$ which is the same for all the terms in the sum. In this case, the summation in (19.8) is reduced to the computation of

$$\sum_{\Delta V} \mathbf{p}_i = \mathbf{P}\,\Delta V. \quad (19.9)$$

Consequently, we get from (19.8) the following expression for the volume density of the force acting in a dielectric:

$$\mathbf{f} = \frac{d\mathbf{F}}{\Delta V} = (\mathbf{P}\cdot\nabla)\,\mathbf{E}. \quad (19.10)$$

Considering that

$$\mathbf{P} = \varkappa\varepsilon_0\mathbf{E} = (\varepsilon - \varepsilon_0)\,\mathbf{E}$$

and using the well-known vector identity

$$(\mathbf{E}\cdot\nabla)\,\mathbf{E} = 1/2\ \mathrm{grad}\ E^2 - \mathbf{E}\times\mathrm{curl}\ \mathbf{E}, \quad (19.11)$$

www.EngineeringBooksPDF.com

in which curl $\mathbf{E} = 0$ in view of the potential nature of the electric field, we get

$$\boxed{\mathbf{f} = \frac{\varepsilon - \varepsilon_0}{2} \operatorname{grad} E^2.} \tag{19.12}$$

This formula is valid both for absolutely rigid dielectrics and compressible ones under the condition that their polarization is a linear function of their mass density, i.e. the dipole moments of individual molecules and atoms do not change due to compression or extension of a volume element, while the dipoles due to displacement of ions are either absent or do not contribute significantly to the polarization. These conditions are satisfied for gases and for most of the liquids.

This formula is quite illustrative, since it shows that the volume elements of a dielectric are subjected to forces which tend to displace these elements towards the highest rate of increase in the absolute value of the electric field. This is sometimes expressed in the form of the statement that the volume element of a dielectric is drawn towards the increasing field strength.

The formula for the volume density of forces, which is valid for isotropic compressible dielectrics, has the form [(see (19.41)]

$$\mathbf{f} = -\frac{1}{2} E^2 \operatorname{grad} \varepsilon + \frac{1}{2} \operatorname{grad} \left[\rho_m \left(\frac{\partial \varepsilon}{\partial \rho_m} \right)_T E^2 \right], \tag{19.13}$$

where ρ_m is the mass density of the dielectric. This formula is also valid when $\varepsilon \neq \text{const}$. If P depends linearly on ρ_m, we get $\varepsilon = D/E = \varepsilon_0 + P/E$, $P \propto \rho_m$, whence $\rho_m \left(\frac{\partial \varepsilon}{\partial \rho_m} \right) = \varepsilon - \varepsilon_0$, and formula (19.13) is transformed into (19.12). If a dielectric has free charges inside it and is subjected to a hydrostatic pressure, Eq. (19.13) is supplemented by the volume density $\rho\mathbf{E}$ of forces acting on the free charges, and by the hydrostatic pressure.

Let us apply these formulas for determining the forces acting on a dielectric sphere in a uniform field (see Fig. 88). In order to apply formula (19.12), we must assume that the transition from the external region having permittivity ε_2 to the internal region of permittivity ε_1 is accomplished not abruptly on the surface of the sphere, but continuously over a certain thin spherical layer. In this layer, field strength $\mathbf{E}$ changes continuously from its value outside the sphere to its value inside the sphere. Formula (19.12) can be used to calculate the force at each point in the spherical layer.

If $\varepsilon_1 > \varepsilon_2$, the field inside the sphere is weaker than outside it. Hence the force at each point of the layer is directed outwards. In view of symmetry, the resultants of these forces on opposite sides of the sphere tend to stretch it along the external field vector (see Fig. 88*a*). However, the resultant of all the forces is equal to zero and the sphere as a whole remains at rest. For $\varepsilon_1 < \varepsilon_2$, the forces in the transient spherical layer are directed inwards and their resultants on both sides of the sphere tend to compress this sphere along the external field vector. As before, the resultant force acting on the sphere as a whole is equal to zero (Fig. 88*b*).

www.EngineeringBooksPDF.com

If, however, the external field is not uniform, the resultant force acting on the sphere as a whole is not equal to zero. It can be easily seen that for $\varepsilon_1 > \varepsilon_2$ this force is directed towards increasing field strength in the medium. This explains why light dielectrics are attracted by electrified bodies: for air $\varepsilon_2 = \varepsilon_0$, and the condition $\varepsilon_1 > \varepsilon_0$ is always satisfied. If, however, $\varepsilon_1 < \varepsilon_2$, the resultant force is directed oppositely, i.e. towards decreasing field strength in the medium. Hence in a medium with a fairly high permittivity, dielectrics with low permittivity are repelled by electrified bodies.

While investigating the behaviour of the electric field at the interface between two dielectrics (see Figs. 84 and 85), it was remarked that E^2 always increases towards the dielectric with a lower permittivity. Hence, by applying

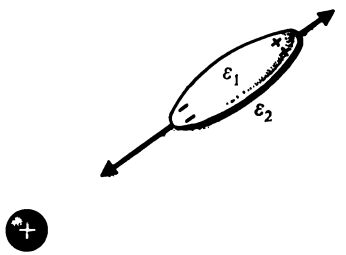


Fig. 92. Emergence of a force of attraction due to a charge on neutral dielectric bodies

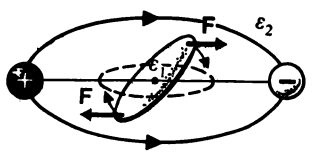


Fig. 93. A dielectric body having the shape of a prolate ellipsoid is oriented so that its major axis is along the field

arguments similar to those considered in the case of a dielectric sphere, we conclude from Eq. (19.12) that at the uncharged interface between two dielectrics, the force is always directed towards the dielectric with lower permittivity. This explains several phenomena. For example, dielectrics (pieces of paper, etc.) are attracted by a charge. Of course, the forces in any part of the surface of the object (a piece of paper, etc.) are directed outwards, but these forces are larger in the regions in the vicinity of a charge. This results in an overall force of attraction (Fig. 92).

Such a behaviour of dielectrics can be understood by proceeding from Eq. (18.35) for the energy of a dielectric having permittivity ε_2, placed in a medium of permittivity ε_1. It is obvious that this energy is negative if $\varepsilon_2 > \varepsilon_1$. It decreases with increasing ε_2 and $\mathbf{E}_1$ and with decreasing ε_1. Since the system strives to attain minimum energy position, the body will be drawn for $\varepsilon_2 > \varepsilon_1$ towards the region with a higher field strength or the lower permittivity ε_1. If, however, $\varepsilon_2 < \varepsilon_1$, the dielectric with permittivity ε_2 will be repelled from the region with a higher field strength to a region where the field is weaker.

Let us assume that a dielectric body in the form of a prolate ellipsoid is placed in a field shown in Fig. 93. Since at all points of the ellipsoid surface the forces directed outwards are stronger at points where the square of the field strength has a larger gradient, a moment of force is created and tends to rotate the ellip-

www.EngineeringBooksPDF.com

soid in such a way that its major axis is parallel to the field lines. This becomes especially clear if we recall that all parts of the dielectric are drawn towards the region with the highest field strength.

If the permittivity of a body is lower than that of the medium, the forces in the surface layer are directed outwards. Consequently, the direction of the resultant force changes. Instead of being attracted by an electrified body, dielectrics (pieces of paper, etc.) are now repelled by it. The pattern of forces in

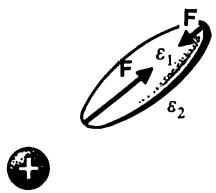


Fig. 94. Emergence of the repulsive force exerted by a charge on a neutral dielectric body placed into a dielectric medium with permittivity higher than that of the body

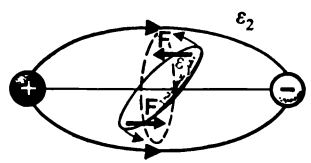


Fig. 95. A prolate ellipsoid in a medium with permittivity higher than its own permittivity is oriented so that its major axis is transverse to the electric field

this case is shown in Fig. 94. A prolate dielectric ellipsoid, when placed in a medium whose permittivity is higher than its own, is oriented with its major axis not along the field lines, but perpendicular to them (Fig. 95). In this case, parts of the dielectric are repelled from the region with a stronger field to regions with a weaker field.

Forces acting on a conductor. The charge $dq = \sigma\, dS$ on the surface element dS of a conductor is acted upon by only half the field at the surface of the conductor, since the other half is created by the charge of the surface element itself and cannot act on it (see Sec. 16, Fig. 39). Consequently, the surface density of force is equal to

$$\mathbf{f}_s = \frac{d\mathbf{F}}{dS} = \frac{\sigma \mathbf{E}}{2} = \frac{\sigma^2}{2\varepsilon}\mathbf{n}, \tag{19.14}$$

where $\mathbf{n}$ is the unit outward normal to the surface of the conductor, and ε is the permittivity of the medium adjoining the conductor [see (17.28)]. Thus, *the force on the surface of a conductor always acts in the direction of the outward normal and as if tends to increase its volume.*

The resultant force acting on the conductor as a whole [see (18.24)] is

$$\mathbf{F} = \frac{1}{2}\int_S \frac{\sigma^2}{\varepsilon}\mathbf{n}\, dS = \frac{1}{2}\int_S \frac{\sigma^2}{\varepsilon}\, dS, \tag{19.15}$$

where S is the surface area of the conductor.

www.EngineeringBooksPDF.com

This expression can be used to calculate the force on the surface element S of the plate of a parallel-plate capacitor filled with a dielectric:

$$F=\frac{1}{2}\frac{\sigma^2}{\varepsilon}S, \tag{19.16}$$

since the field is uniform in this case, i.e. σ and ε in the integrand of (19.15) are constant. This force is directed into the capacitor.

Surface forces acting on a dielectric. In the equilibrium state, volume (body) forces of electrostatic nature do not lead to a displacement of the corresponding volume element. *Such forces cause a deformation of the medium, thus creating elastic body forces which completely balance the volume forces of electrostatic nature.* A similar balancing takes place in the volume of a liquid in a gravitational field. Each volume element is acted upon by a force of gravity due to the liquid in this volume; this force, however, is balanced by the force generated by the pressure exerted on the surface of the volume element by neighbouring regions. Electrostatic volume forces lead to a displacement of volume elements only for quite rapid variations of the fields, when elastic forces do not balance electric forces at each instant of time. The resultant of all volume elements is applied to the dielectric as a whole and may cause its displacement if it is not balanced by some other force.

Besides volume forces, a dielectric is also subjected to *surface forces which emerge in its surface layer.* These forces act together with the volume forces. We shall derive these forces from the first law of thermodynamics.

The thermodynamic potential in isothermal processes is the free energy F which is connected with work through the relation

$$dA = -dF. \tag{19.17}$$

Since the thermodynamic relations in the absence of an electric field were described in molecular physics, we shall now take into account only those quantities which depend on the electric field. Hence, we consider in (19.17) only the work and change in the free energy due to the electric field. The work and change in the free energy due to deformations and elastic forces are not taken into account, i.e. the dielectric is assumed to be nondeformable. Besides, we confine the analysis to isotropic dielectrics only.

The part of the internal energy which is not bound in the system and is available for obtaining work is considered as the free energy. Its magnitude depends on the conditions under which a process is carried out.

Let us consider the plane interface between two dielectrics having permittivities ε_1 and ε_2. By way of a specific model for the physical system, we can consider a parallel-plate capacitor, the space between whose plates is filled with liquid dielectrics with a plane interface. The interface may be parallel or perpendicular to the capacitor plates. This model can be used to obtain an expression for the surface density of forces acting at the interface between dielectrics. Since the relations obtained in this case are of a local nature, they will be independent of the specific shape of the nonlocal model under which they have been obtained. In other words, these relations will be of a general nature.

www.EngineeringBooksPDF.com

Let us consider the plane interface parallel to the capacitor plates (Fig. 96). The field vector **E** is perpendicular to the interface. The normal directed towards the second dielectric is taken as positive. Upon an infinitely small displacement of the interface, work is performed due to a change in the free energy. Having independently calculated the work and the variation in the free energy, we can

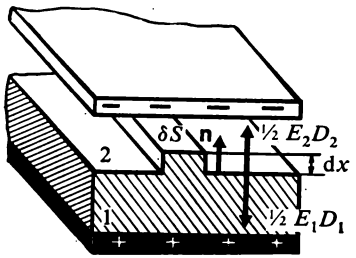


Fig. 96. Emergence of Maxwellian stresses

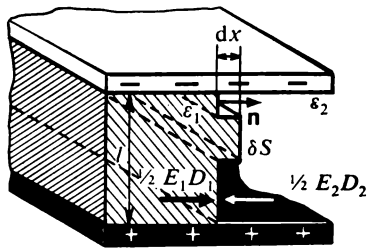


Fig. 97. Emergence of Maxwellian pressures

find the surface density of forces from (19.17). Of course, the displacement dx should be considered as a virtual displacement, i.e. not necessarily realizable in practice.

The work done in displacing the surface element ΔS by dx along the normal is

$$dA = \Delta S f_s \, dx, \tag{19.18}$$

where f_s is the surface density of the force.

In order to calculate dF, we consider that $D_2 = D_1$ at the interface between the dielectrics, i.e. at $D = \text{const}$ the interface is displaced. This corresponds to the condition that the charge at the plates of the capacitor is constant, since $D = \sigma$. Consequently, we must calculate dF at a constant charge q on the plates, i.e. $(dF)_{T,q}$. As a result of the displacement of the interface by dx, the volume $\Delta S \, dx$, which was initially filled with an electric energy of density $E_2D_2/2$, will now be filled with energy whose density is $E_1D_1/2$. There are no other energy factors participating in doing work. Consequently, the difference in energy in the volume $\Delta S \, dx$ before and after the displacement of the interface is just the variation in the free energy:

$$(d\mathbf{F})_{T,q} = \left(\frac{1}{2} D_{1n}\mathbf{E}_{1n} - \frac{1}{2} D_{2n}F_{2n}\right) \Delta S \, dx \tag{19.19}$$

where the subscript n means that the normal components of **D** and **E** are considered.

Taking (19.18) and (19.19) into account, we can write (19.17) in the form

$$f_s = \frac{1}{2} E_{2n}D_{2n} - \frac{1}{2} E_{1n}D_{1n}. \tag{19.20}$$

www.EngineeringBooksPDF.com

The surface density of force is directed along the normal to the interface. It can be seen from (19.20) that the surface force density f_s is composed of two parts:

(1) the surface density

$$f_{2s} = \frac{1}{2} E_{2n} D_{2n}, \tag{19.21}$$

of the force emerging under the effect of the electric field of the second medium and directed towards the second medium;

(2) the surface density

$$f_{1s} = -\frac{1}{2} E_{1n} D_{1n}, \tag{19.22}$$

of the force emerging under the effect of the electric field of the first medium and directed towards the first medium.

Thus, *in this case the electric fields on both sides of the interface as if attract the interface with a surface density of force equal to the volume density of the electric energy along the normal component of the field.*

The resultant of the two forces applied to the interface from both sides of the fields is the total force acting on the interface. Since $D_{2n} = D_{1n} = D_n$, we get [see (19.20)]

$$f_s = \frac{1}{2} D_n^2 \left(\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1} \right). \tag{19.23}$$

For $\varepsilon_2 < \varepsilon_1$, the surface density f_s of force is greater than zero. This means that the force at the interface acts in the direction of the dielectric having a lower permittivity, i.e. in the direction of the higher volume density of the electric energy. It should be noted that the volume density of the force [see (19.12)] is also directed towards increasing volume density of the electric energy.

Let us now consider the dielectrics the plane interface between which is perpendicular to the plates of a parallel-plate capacitor (Fig. 97). In this case, the condition $E_{2\tau} = E_{1\tau} = E_\tau$ is satisfied, since the field vector is parallel to the interface. The subscript τ indicates the components of the vectors tangential to the interface. The displacement of the boundary takes place under the condition $E_\tau = \text{const}$, i.e. at a constant potential difference. Consequently, it is necessary to calculate the variation in the free energy $(dF)_{T,\varphi}$. In order to maintain a constant potential difference, we must change the charge density in that part of the capacitor plates which corresponds to a displacement of the interface by dx. This involves an expenditure of energy equal to $dq\,(\varphi_2 - \varphi_1) = dq\,E_\tau l$, where E_τ and l are the field strength and the distance between the capacitor plates. The surface charge densities in the region where the plates come in contact with the first and second dielectrics are respectively equal to $\sigma_1 = \varepsilon_1 E_1 = \varepsilon_1 E_\tau$ and $\sigma_2 = \varepsilon_2 E_2 = \varepsilon_2 E_\tau$. The depth of the dielectric in the direction perpendicular to the plane of Fig. 97 is equal to $\Delta S/l$. Consequently,

$$dq = (\sigma_1 - \sigma_2)\,(\Delta S/l)\,dx. \tag{19.24}$$

www.EngineeringBooksPDF.com

Under the conditions described above, work can be performed only due to the difference between the field energy and the energy spent to keep the potentials constant. Consequently, the change in the free energy is equal to

$$(\mathrm{d}F)_{T,\varphi}=\left(\frac{1}{2}E_{1\tau}D_{1\tau}-\frac{1}{2}E_{2\tau}D_{2\tau}\right)\Delta S\,\mathrm{d}x-(\sigma_2-\sigma_1)(\Delta S/l)\,\mathrm{d}x\,E_\tau l. \quad (19.25)$$

Since $\sigma_2 = \varepsilon_2 E_\tau$ and $\sigma_1 = \varepsilon_1 E_\tau$, we get

$$(\mathrm{d}F)_{T,\varphi}=-\left(\frac{1}{2}E_{1\tau}D_{1\tau}-\frac{1}{2}E_{2\tau}D_{2\tau}\right)\Delta S\,\mathrm{d}x. \quad (19.26)$$

Taking (19.18) and (19.26) into account, we can write (19.17) in the form

$$f_s=-\frac{1}{2}E_{2\tau}D_{2\tau}+\frac{1}{2}E_{1\tau}D_{1\tau}. \quad (19.27)$$

This surface density of force is also directed along the normal to the interface. It can be seen from (19.27) that it is composed of two parts:

(1) the surface density

$$f_{2s}=-\frac{1}{2}E_{2\tau}D_{2\tau}, \quad (19.28)$$

of the force exerted at the interface by the electric force of the second medium in the direction of the first medium. It should be recalled that the positive normal is chosen in the direction from the first medium to the second, and hence the minus sign in (19.28) indicates that the force is directed from the second medium to the first;

(2) the density

$$f_{1s}=\frac{1}{2}E_{1\tau}D_{1\tau}, \quad (19.29)$$

of the force exerted at the interface by the electric field of the first medium in the direction of the second medium.

Thus, *the tangential component of the electric field strength as if exerts a pressure on the interface in contact with it. This pressure is equal to the volume density of the energy corresponding to the tangential component of the field strength.*

The resultant of the pressure forces exerted by the fields on both sides of the interface is the total force applied to it. Since $E_{1\tau} = E_{2\tau} = E_\tau$, formula (19.27) assumes the form

$$f_s=\frac{1}{2}E_\tau^2(\varepsilon_1-\varepsilon_2). \quad (19.30)$$

For $\varepsilon_2 < \varepsilon_1$, the force density $f_s > 0$. Consequently, the surface density of the force is directed towards the dielectric having lower permittivity. Thus, *irrespective of the orientation of the field with respect to the interface, the surface density of the force is always directed towards the dielectric with lower permittivity* [see (19.12)]. The validity and generality of this statement follows also from Eq. (18.36) if we consider that the system tends to go over to a state with minimum energy.

www.EngineeringBooksPDF.com

Volume forces acting on a compressible dielectric. We proceed from formula (18.36) in which $\delta\varepsilon$ is due to the strain which changes the mass density. The processes are assumed to be isothermal (T = const). The permittivity changes from point to point, being a function of $\mathbf{r}$. Besides, it may be a function of the mass density ρ_m of the dielectric, i.e. $\varepsilon = \varepsilon(\mathbf{r}, \rho_m)$. Suppose that as a result of deformation, the volume element dV is displaced by $\mathbf{l}$ and that the mass density of the dielectric changes in the process. The volume element which is at the point with radius vector $\mathbf{r}$ after displacement was at the point $\mathbf{r} - \mathbf{l}$ before displacement. Hence,

$$d\varepsilon = -\mathbf{l}\,\mathrm{grad}\,\varepsilon + \frac{\partial\varepsilon}{\partial\rho_m}\delta\rho_m, \tag{19.31}$$

where $\delta\rho_m$ is the change in the mass density of the dielectric.

It can be shown that after deformation the volume element dV' becomes equal to

$$dV = (1 + \mathrm{div}\,\mathbf{l})\,dV'. \tag{19.32}$$

The law of mass conservation has the following form for the volume element:

$$\rho_m\,dV = \rho_m'\,dV' \tag{19.33}$$

or

$$\rho_m(1 + \mathrm{div}\,\mathbf{l})\,dV' = \rho_m'\,dV', \tag{19.34}$$

where ρ_m' and ρ_m are the mass densities after and before deformation. From (19.34) we obtain the following expression for an infinitely small displacement:

$$\delta\rho_m = \rho_m - \rho_m' = -\rho_m\,\mathrm{div}\,\mathbf{l}. \tag{19.35}$$

Substituting (19.31) and (19.35) into (18.36), we obtain

$$\delta W = \frac{1}{2}\int\left[E^2\mathbf{l}\cdot\mathrm{grad}\,\varepsilon + E^2\rho_m\frac{\partial\varepsilon}{\partial\rho_m}\mathrm{div}\,\mathbf{l}\right]dV. \tag{19.36}$$

From formula (A.12), we have

$$E^2\rho_m\frac{\partial\varepsilon}{\partial\rho_m}\mathrm{div}\,\mathbf{l} = \mathrm{div}\left(E^2\rho_m\frac{\partial\varepsilon}{\partial\rho_m}\mathbf{l}\right) - \mathbf{l}\cdot\mathrm{grad}\left(E^2\rho_m\frac{\partial\varepsilon}{\partial\rho_m}\right). \tag{19.37}$$

This gives [see (19.36)]

$$\delta W = \frac{1}{2}\int\left[E^2\,\mathrm{grad}\,\varepsilon - \mathrm{grad}\left(E^2\rho_m\frac{\partial\varepsilon}{\partial\rho_m}\right)\right]\cdot\mathbf{l}\,dV + \frac{1}{2}\int\mathrm{div}\left(E^2\rho_m\frac{\partial\varepsilon}{\partial\rho_m}\mathbf{l}\right)dV. \tag{19.38}$$

On the basis of the usual assumptions about the continuity of integrands, we can transform the second of the above integrals into an integral over the surface bounding the volume under consideration with the help of the Gauss theorem. Assuming for the sake of simplicity that the dielectric occupies the entire space while the charges generating the field are distributed over a finite region of space, we find that the second integral is equal to zero, since $E^2 \propto 1/r^4$, r being the

www.EngineeringBooksPDF.com

distance between the charge and the integration surface. Consequently,

$$\int \operatorname{div}\left(E^2\rho_m \frac{\partial\varepsilon}{\partial\rho_m}\mathbf{l}\right) dV = \int\limits_{S\to\infty} E^2\rho_m \frac{\partial\varepsilon}{\partial\rho_m}\mathbf{l}\cdot d\mathbf{S} \to 0. \tag{19.39}$$

The volume density $\mathbf{f}$ of the forces describes the action of an electric field on a dielectric. The volume density of the work done by this force upon deformation is equal to $\mathbf{f}\cdot\mathbf{l}$. Consequently, the law of energy conservation for deformation can be expressed in the following form if we take into account Eqs. (19.38) and (19.39):

$$\int \mathbf{f}\cdot\mathbf{l}\, dV = -\frac{1}{2}\int \left[E^2 \operatorname{grad}\varepsilon - \operatorname{grad}\left(E^2\rho_m \frac{\partial\varepsilon}{\partial\rho_m}\right)\right]\cdot\mathbf{l}\, dV. \tag{19.40}$$

Since this relation is valid for any displacement $\mathbf{l}$, we obtain

$$\boxed{\mathbf{f} = -\frac{1}{2}E^2 \operatorname{grad}\varepsilon + \frac{1}{2}\operatorname{grad}\left(E^2\rho_m \frac{\partial\varepsilon}{\partial\rho_m}\right).} \tag{19.41}$$

This formula is valid for isotropic compressible dielectrics for any dependence of ε on the mass density ρ_m [see (19.13)].

If polarization depends linearly on the volume density of the mass, we get

$$\rho_m \frac{\partial\varepsilon}{\partial\rho_m} = \varepsilon - \varepsilon_0 \tag{19.42}$$

and Eq. (19.41) is transformed into (19.12). Consequently, formula (19.12) is valid not only for rigid dielectrics, but also for compressible dielectrics with $P \propto \rho_m$.

Although for the sake of simplicity in transformation of (19.39) formula (19.41) was derived under the assumption that the dielectric occupies the entire space, it is actually valid under any condition, since it is a differential relation whose validity is independent of the processes occurring at other points in space.

Calculation of forces from the expression for energy. In order to transfer a charge dq to a point with potential φ, we must perform work $\varphi\, dq$. Consequently, the total change in the energy of a system of charges upon a change in the charge by dq_i is

$$\sum_j \varphi_j\, dq_j. \tag{19.43}$$

This change is accompanied by a change in the energy of the electric field by dW and by performance of work by the charges. If the system configuration is characterized by the parameters ξ_i then, by definition, the generalized force connected with this parameter is the quantity F_i, such that $F_i\, d\xi_i$ is the work done by the system when the parameter ξ_i changes by $d\xi_i$. The law of conservation of energy has the form

$$\sum_j \varphi_j\, dq_j = dW + \sum_i F_i\, d\xi_i. \tag{19.44}$$

www.EngineeringBooksPDF.com

To begin with, let us consider virtual processes in which charges remain constant, i.e. $dq_i = 0$. In this case, Eq. (19.44) assumes the form

$$0 = (dW)_q + \sum_i F_i \, d\xi_i. \tag{19.45a}$$

Here, $(dW)_q$ depends only on ξ_i. Consequently,

$$(dW)_q = \sum_i \left(\frac{\partial W}{\partial \xi_i}\right)_q d\xi_i. \tag{19.45b}$$

Considering that $d\xi_i$ is an independent quantity, a comparison of (19.45a) and (19.45b) gives

$$\boxed{F_i = -\left(\frac{\partial W}{\partial \xi_i}\right)_q,} \tag{19.46}$$

where the subscript q on the partial derivative in an explicit form indicates that the force is calculated at a constant value of the charges. *In order to use this formula, we must express the energy W as a function of charges and parameters* ξ_i.

The generalized force can also be expressed in terms of the derivative at a constant potential. For this purpose, we take into consideration the expression

$$W = \frac{1}{2} \sum_i \varphi_i q_i. \tag{19.47}$$

The change in energy under a constant potential is equal to

$$(dW)_\varphi = \frac{1}{2} \sum_i \varphi_i \, dq_i. \tag{19.48}$$

Consequently [see (19.45a)],

$$0 = (dW)_\varphi - \sum_i F_i \, d\xi_i. \tag{19.49}$$

Taking into account the independence of $d\xi_i$, we obtain

$$\boxed{F_i = \left(\frac{\partial W}{\partial \xi_i}\right)_\varphi,} \tag{19.50}$$

where the subscript φ on the partial derivative in an explicit form shows that differentiation is carried out at a constant potential. *In order to use this formula, we must express the energy W as a function of potentials* φ_i *and parameters* ξ_i. Obviously, formulas (19.46) and (19.50) are equivalent and are obtained from each other. The choice of a particular formula depends on the circumstances.

For example, suppose that it is required to calculate the force of attraction between the plates of a parallel-plate capacitor. The energy of such a capacitor

www.EngineeringBooksPDF.com

is equal to

$$W = Q^2/(2C) = (\Delta\varphi)^2 C/2,$$

where $C = \varepsilon_0 S/x$, S and x being the surface area of a plate and the distance between the plates respectively.

Calculation of the force from formulas (19.46) and (19.50) gives

$$F_x = -\frac{\partial}{\partial x}\left(\frac{Q^2}{2C}\right)_Q = -\frac{Q^2}{2}\frac{\partial}{\partial x}\left(\frac{1}{C}\right) = \frac{Q^2}{2C^2}\frac{\partial C}{\partial x}; \tag{19.51}$$

$$F'_x = \frac{\partial}{\partial x}\left[\frac{(\Delta\varphi)^2 C}{2}\right]_\varphi = \frac{(\Delta\varphi)^2}{2}\frac{\partial C}{\partial x}. \tag{19.52}$$

Taking into consideration the definition of the capacitance $C = Q/\Delta\varphi$, we conclude that $F'_x = F_x$.

The forces in an electric field are ultimately the forces acting on the charges, although the value of charges does not always appear in the expression for the force.
The formula for the force acting on perfectly rigid dielectrics is also valid for compressible dielectrics provided that their polarization depends linearly on the density of the material.
The forces acting on a dielectric depend on the ratio of the permittivity of the body and the permittivity of the surrounding medium. At the interface between dielectrics, the force is always directed towards the dielectric having lower permittivity.

The field component normal to the interface between dielectrics as if attracts the surface with the surface density of force equal to the volume density of electric energy of the field associated with this component.
The field component tangential to the interface between dielectrics as if exerts a pressure on the surface, the pressure being equal to the volume density of electric energy of the field associated with this component.
The surface force acts towards the dielectric with lower permittivity in all cases, irrespective of the field orientation.

Example 19.1 *Proceeding from the solution of Example* 16.3 *find the moment of the force of attraction between the plates of the capacitor shown in* Fig. 73.

The energy of the capacitor is equal to [see (16.109)]

$$W = \frac{U_0^2 C}{2} = \frac{U_0^2 l\varepsilon \ln(b/a)}{2\alpha_0}. \tag{19.53}$$

The generalized force for the angle of rotation is the moment M of the force with respect to the axis coinciding in the present case with the line of intersection of the capacitor plates. Hence, taking into account Eq. (19.50), we obtain

$$M = \left(\frac{\partial W}{\partial \alpha_0}\right)_\varphi = -\frac{U_0^2 l\varepsilon \ln(b/a)}{2\alpha_0^2} \tag{19.54}$$

where the minus sign indicates that the moment of force tends to reduce the angle α_0. In other words, forces of attraction exist between the plates of the capacitor. Of course, attractive forces always exist between the capacitor plates and formula (19.54) just states that the moment of force is obtained with the minus sign. Such a verification of the correctness of the result is useful when generalized coordinates and generalized forces are used, in which case these variables cannot be interpreted graphically.

We can obtain this result in another way. The surface density of the force acting on a conductor is $f = \sigma^2/(2\varepsilon)$. Consequently, a layer of length l between r and $r + dr$ is subjected to

www.EngineeringBooksPDF.com

a force

$$dF = -fl\,dr = -\frac{\varepsilon U_0^2}{2\alpha_0^2 r^2}\,l\,dr, \tag{19.55}$$

where the value (16.107b) is used for σ. The minus sign indicates that this force tends to reduce the angle α_0. The resultant force acting on the plate is equal to

$$F = \int_a^b dF = -\frac{\varepsilon U_0^2 l}{2\alpha_0^2}\int_a^b \frac{dr}{r^2} = \frac{\varepsilon U_0^2 l}{2\alpha_0^2}\left(\frac{1}{b}-\frac{1}{a}\right). \tag{19.56}$$

The line of application of forces is at a distance r_0 from the axis of rotation, which is determined from the condition

$$r_0 F = \int_a^b r\,dF = -\frac{\varepsilon U_0^2 l}{2\alpha_0^2}\ln\frac{b}{a}, \tag{19.57}$$

whence

$$r_0 = \frac{ab}{b-a}\ln\frac{b}{a}. \tag{19.58}$$

The moment of force with respect to the axis of rotation is equal to

$$M = r_0 F = -\frac{\varepsilon U_0^2 l}{2\alpha_0^2}\ln\frac{b}{a}, \tag{19.59}$$

which is identical to (19.54).

Problems

2.1. Find the strength of the electric field in a spherical cavity of radius a inside a uniformly charged sphere of radius R. The volume charge density is equal to ρ (Fig. 98).

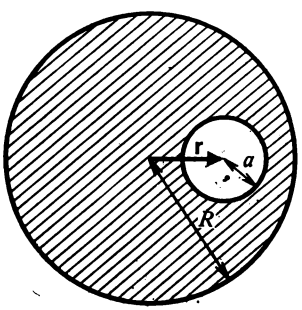


Fig. 98. A cylindrical cavity in a cylinder or a spherical cavity in a sphere

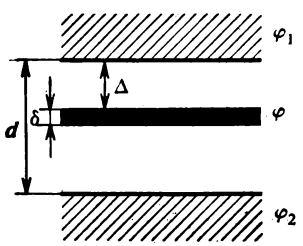


Fig. 99. A conducting plate in a parallel-plate capacitor

2.2. Find the field strength in an infinite cylindrical cavity whose axis is parallel to the axis of a very long uniformly charged circular cylinder. The volume charge density is equal to ρ (Fig. 98).

2.3. The distance between the plates of a parallel-plate capacitor is equal to d. A metallic strip of thickness δ, with its surface parallel to the capacitor plates, is brought into the space between the plates whose potentials are φ_1 and φ_2 respectively (Fig. 99). Find the potential of the metallic strip.

www.EngineeringBooksPDF.com

2.4. Find the force acting on a charge q placed at a distance d from the centre of an uncharged isolated conducting sphere of radius r_0 ($r_0 < d$).

2.5. Find the force acting on a charge q placed in a metallic sphere at a distance r from its centre. The radius of the sphere is equal to a.

2.6. A point charge q is placed between two concentric conducting spheres of radii r_1 and r_2 ($r_1 < r_2$) at a distance d from their common centre ($r_1 < d < r_2$). Find the charges induced on the spheres.

2.7. A point charge q is placed at a distance d from the centre of a grounded sphere. Find the ratio f of the charge induced on the part of the sphere seen from the point of location of the charge q to the charge on the part which cannot be seen from this point. The radius of the sphere is equal to a, $d > a$.

2.8. Two capacitors having capacitances C_1 and C_2 and charges q_1 and q_2 (q_1 and q_2 are the absolute values of the charge on the plates of the first and second capacitors) are connected in parallel. Calculate the change in the energy of the capacitors and explain the reasons behind the result.

2.9. The permittivity of the medium between the plates of a parallel-plate capacitor (surface area of plates is equal to S) changes uniformly from ε_1 to ε_2. The distance between the plates is equal to d. Find the capacitance of the capacitor.

2.10. A cylindrical capacitor with plates of radii r_1 and r_2 is immersed vertically into a dielectric liquid having permittivity ε. The bottom of the capacitor is in the liquid while the top is in the air whose permittivity is taken as ε_0. The mass density of the liquid is ρ. Find the height h to which the liquid rises between the capacitor plates if the potential difference between them is U.

2.11. A conducting sphere of density ρ_1 floats in a liquid of density ρ_2 ($\rho_2 > 2\rho_1$) and permittivity ε. The sphere is submerged in the liquid to less than half its diameter. What charge must be imparted to the sphere so that it is half submerged into the liquid? The radius of the sphere is equal to a.

2.12. A parallel-plate capacitor has square plates with side a. The distance and the potential difference between the plates are respectively equal to d and U. A square strip with side a and thickness Δ is partially introduced into the space between the plates. The surfaces and edges of the strip are parallel to the surfaces and edges of the plates, and its permittivity is equal to ε. Find the force with which the strip is drawn into the space between the capacitor plates.

2.13. A uniformly charged very long filament is placed at a distance d from the axis of an infinite conducting cylinder of radius r, the cylinder axis being parallel to the filament. The linear charge density of the filament is τ. Find the force acting on the length l of the filament ($d > r$).

2.14. Using the method of images, find the force acting on length l of each of two infinite conducting cylinders the distance between whose parallel axes is equal to d. The radii of the cylinders are equal to r_1 and r_2, and one of the cylinders is charged and has a linear charge density τ.

2.15. Find the dipole moment of a charge distributed uniformly over the surface of a sphere of radius a. One hemisphere has a charge Q while the charge on the other is equal to $-Q$.

2.16. A point dipole with a moment $\mathbf{p}$ lies at a distance d from the centre of a grounded conducting sphere of radius a. Find the induced dipole moment of the sphere.

2.17. A constant potential difference U_0 is applied to the square plates of a parallel-plate air capacitor with side l. Find the force which must be applied in order to displace one of the plates parallel to itself in a direction perpendicular to any side of the square so that the distance d between the plates remains unchanged.

2.18. There is a conducting sphere of radius r_1 and a concentric spherical conducting layer whose inner surface has a radius r_2 ($r_2 > r_1$), while its outer surface has a radius equal to r_3 ($r_3 > r_2$). The space between the spheres of radii r_1 and r_2 is empty. The sphere and the layer carry charges equal to Q_1 and Q_2 respectively but, unlike the case of a capacitor, Q_1 and $-Q_2$ are not equal in this case. Find the energy of this system of charges.

2.19. Find the electric field strength at the centre of a right circular cylinder of length l and radius a, whose polarization $\mathbf{P}$ is uniform and parallel to the axis.

2.20. The polarization $\mathbf{P}$ in Problem 2.19 is perpendicular to the axis of the cylinder. Find the field strength at the centre of the cylinder.

www.EngineeringBooksPDF.com

2.21. An infinite conducting cylinder of circular cross section of radius a and a conducting plane at a distance d from the axis of the cylinder form a capacitor. Find the capacitance of such a capacitor of length l.

2.22. Using the solution of problem 2.21, find the force exerted by an earthed infinite plane on a segment of length l of a rectilinear charged filament parallel to the plane. The linear charge density of the filament is equal to τ.

2.23. A molecule is represented by a charge $-2|q|$ at the origin and two charges $|q|$ at the points characterized by the radius vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, where $|\mathbf{r}_1| = |\mathbf{r}_2| = l$. The angle between $\mathbf{r}_1$ and $\mathbf{r}_2$ is denoted by θ. Find the effective charge $|q|_{\text{eff}}$ for a water molecule for which $l = 0.958 \times 10^{-10}$ m, $\theta = 105°$, and $p = 6.14 \times 10^{-30}$ C·m.

2.24. A point charge q is placed between two infinite earthed parallel conducting planes separated by a distance d. The distance between the point charge and one of the plates is equal to x. Finding the images of the charge q, calculate the force acting on it.

Answers

2.1. $E = \rho r/(3\varepsilon_0)$. 2.2. $E = \rho r/(2\varepsilon_0)$. 2.3. $\varphi = \varphi_1 - \frac{\Delta}{d-\delta}(\varphi_1 - \varphi_2)$. 2.4. $F = -\frac{q^2 r_0^3}{4\pi\varepsilon_0 d^3} \times \left[\frac{2d^2 - r_0^2}{(d^2 - r_0^2)^2}\right]$. 2.5. $F = \frac{q^2 a r}{4\pi\varepsilon_0 (a^2 - r^2)^2}$. 2.6. $q_1 = -\frac{r_1 (r_2 - d)}{d (r_2 - r_1)} q$, $q_2 = -\frac{r_2 (d - r_1)}{d (r_2 - r_1)} q$.

2.7. $f = \sqrt{(d+a)/(d-a)}$. 2.8. $\Delta W = (C_2 q_1 - C_1 q_2)^2/[2C_1 C_2 (C_1 + C_2)]$. 2.9. $C = \frac{S}{d} \times \frac{\varepsilon_2 - \varepsilon_1}{\ln(\varepsilon_2/\varepsilon_1)}$. 2.10. $h = \frac{(\varepsilon - \varepsilon_0) U^2}{(r_2^2 - r_1^2) \ln(r_2/r_1)} \frac{1}{\rho g}$. 2.11. $C = 4\pi(\varepsilon + \varepsilon_0) \sqrt{\frac{a^5 g (\rho_2 - 2\rho_1)}{3(\varepsilon - \varepsilon_0)}}$.

2.12. $F = \frac{\varepsilon_0}{2} \frac{(\varepsilon - \varepsilon_0)\Delta}{(d-\Delta)\varepsilon + \Delta\varepsilon_0} \frac{a}{d} U^2$. 2.13. $f = -\tau^2 dl/[2\pi\varepsilon_0 (d^2 - r^2)]$. 2.14. $f = -\frac{\tau^2 dl}{2\pi\varepsilon_0} \times [d^2 - (r_1 + r_2)^2]^{-1/2} [d^2 - (r_1 - r_2)^2]^{-1/2}$. 2.15. $p = Qa$. 2.16. $\mathbf{p}_{\text{ind}} = \mathbf{p} a^3/d^3$. 2.17. $F = -\frac{1}{2} \frac{\varepsilon_0 l}{d} U_0$. 2.18. $W = \frac{1}{8\pi\varepsilon_0} \left[\left(\frac{1}{r_1} - \frac{1}{r_2} + \frac{1}{r_3}\right) Q_1^2 + \frac{2Q_1 Q_2 + Q_2^2}{r^3}\right]$. 2.19. $E = -(1/\varepsilon_0) P (1 - l/\sqrt{4a^2 + l^2})$. 2.20. $E = -\frac{[1/(2\varepsilon_0)] lP}{\sqrt{4a^2 + l^2}}$. 2.21. $C = \frac{2\pi\varepsilon_0 l}{\ln[(d + \sqrt{d^2 - a^2})/a)]}$; for $a \ll d$ we have $C \approx \frac{2\pi\varepsilon_0 l}{\ln(2d/a)}$. 2.22. $F = -\left(\frac{\partial W}{\partial d}\right)_Q = \left(\frac{\partial W}{\partial d}\right)_\varphi = \frac{1}{2} U^2 \frac{\partial C}{\partial d} = \frac{U^2}{d} \times \frac{\pi\varepsilon_0 l}{(\ln 2d/a)^2} = \frac{U^2 C^2 l}{4\pi\varepsilon_0 d} = \frac{\tau^2 l}{4\pi\varepsilon_0 d}$. 2.23. $\mathbf{p} = |q|_{\text{eff}} (\mathbf{r}_1 + \mathbf{r}_2)$, $p = 2|q|_{\text{eff}} l \cos(\theta/2)$, $|q|_{\text{eff}} = 5.26 \cdot 10^{-20}$ C $= 0.328|e|$. 2.24. $F = -\frac{q^2}{16\pi\varepsilon_0} \left\{\frac{1}{x^2} + \sum_{n=1}^{\infty} \left[\frac{1}{(nd + x)^2} - \frac{1}{(nd - x)^2}\right]\right\}$.

www.EngineeringBooksPDF.com

CHAPTER 3

Dielectrics

The electric dipole moment of atoms and molecules is the basic physical factor determining the nature of interaction between a dielectric and an electric field.
The principal mechanisms of polarization are associated with the appearance of induced dipole moments of atoms and molecules or with the spatial reorientation and rearrangement of available dipole moments.
Ionic lattice polarization also takes place.

Sec. 20. Local Field

The reasons behind the difference between the local and external fields are discussed, and the local field strength is calculated for simplest conditions.

The difference between a local field and an external field. As a result of polarization, a dielectric placed in an external field becomes a source of an electric field. Consequently, *the field inside a dielectric, which acts on dielectric molecules, differs from the external field.* This field is called **local field**. The difference between the local field and the external field is especially significant for dielectrics with a high density, viz. liquids and solids.

Calculation of local field strength. Let us isolate in the volume of a dielectric a physically small sphere at whose centre the local field strength is being calculated (Fig. 100). The field appearing at the centre of the sphere as a result of polarization of the dielectric consists of the field $\mathbf{E}_1$ generated by the part of the dielectric located outside the volume bounded by the sphere and the field $\mathbf{E}_2$ created by the part of the dielectric contained in the volume bounded by the sphere.

While calculating $\mathbf{E}_1$, we can assume that the dielectric is a continuum because the distance from the centre of the sphere at which the local field strength is being calculated to the sources of the field is comparatively large. Since the volume of the sphere is physically small, we can assume that the medium near its outer surface is polarized uniformly. In the volume bounded by the sphere, we should take into account the atomic structure of the dielectric, i.e. calculate the contribution of the dipole moment of each individual atom to the local field strength and assume that the sphere is the interface between the surrounding

www.EngineeringBooksPDF.com

medium and the vacuum in the volume bounded by it.

At the centre of the sphere, the field is created by bound charges at its surface as on the interface between two media with different permittivities. The surface density of bound charges is given by [see (17.21)]

$$\sigma_b = (P_{2n} - P_{1n}) = -P_{1n}, \qquad (20.1)$$

where P_{1n} is the normal component of polarization from the outer side of the surface of the sphere and $P_{2n} = 0$ from the inner side. In this formula, the outward normal of the sphere is positive. Directing the Z-axis along the vector $\mathbf{P}$ of constant polarization, we obtain

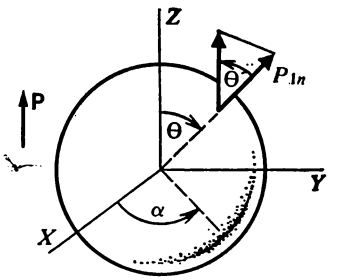


Fig. 100. To the calculation of local field

$$\sigma_b = -P_{1n} = -P \cos\theta. \qquad (20.2)$$

The surface charge contained within a solid angle $d\Omega$ is given by

$$dQ = \sigma_b r^2 \, d\Omega, \qquad (20.3)$$

where r is the radius of the sphere. At the centre of the sphere this charge creates a field

$$dE_z = -\frac{1}{4\pi\varepsilon_0} \frac{dQ}{r^2} \cos\theta \qquad (20.4)$$

in the direction of the Z-axis.

It can be seen that only the component of the field along the Z-axis differs from zero. Combining (20.4) and (20.3), we obtain

$$E_z = E_1 = \frac{1}{4\pi\varepsilon_0} P \int \cos^2\theta \, d\Omega = E_1$$

$$= \frac{1}{4\pi\varepsilon_0} P \int_0^{2\pi} d\alpha \int_0^{\pi} \cos^2\theta \sin\theta \, d\theta = \frac{1}{3\varepsilon_0} P \qquad (20.5)$$

or, in vector form,

$$\boxed{\mathbf{E}_1 = \frac{1}{3\varepsilon_0} \mathbf{P}.} \qquad (20.6)$$

This formula is valid only for an infinite homogeneous dielectric. If a dielectric is finite, the field in it generally depends on its size and shape. In homogeneous dielectrics, volume polarization charges are equal to zero since $\rho_b = -\text{div}\,\mathbf{P} = -\varkappa\varepsilon_0 \,\text{div}\,\mathbf{E} = 0$. Therefore, the difference between the field of a finite dielectric and the field $\mathbf{E}_1'$ of an infinite dielectric is due to the field created by bound charges appearing on the outer surface of the body. This field is sometimes called a depolarization field since it reduces the field strength.

www.EngineeringBooksPDF.com

The field $\mathbf{E}_2$ depends on the distribution of dipole moments of molecules inside the small isolated sphere and cannot be represented by any universal formula. Let us calculate the field for the case when molecules are located at the sites of a cubic crystal lattice and all dipole moments are oriented along the same direction in space. This condition is satisfied for induced dipole moments. We must find the field $\mathbf{E}_2$ at the point of location of a molecule, i.e. at a lattice site. We fix the origin of coordinates at this point and direct the X-, Y-, and Z-axes along the edges of the lattice. Let us use formula (16.85) which has the following form for the x-component:

$$E_{2x}=\frac{p_x}{4\pi\varepsilon_0}\sum_i\frac{-r_i^2+3x_i^2}{r_i^5}+\frac{p_y}{4\pi\varepsilon_0}\sum_i\frac{3x_iy_i}{r_i^5}+\frac{p_z}{4\pi\varepsilon_0}\sum_i\frac{3x_iz_i}{r_i^5}. \quad (20.7)$$

The summation is carried out over all molecules in a small volume inside the sphere. Similar formulas can also be written for the y- and z-components of the field.

In formula (20.7), we can first calculate the sum over all molecules contained in a small spherical layer of radius r and then calculate the sum over the spherical layers corresponding to different r's. In view of cubic symmetry, for the first summation we have

$$\sum_i x_i^2=\sum_i y_i^2=\sum_i z_i^2=\frac{1}{3}\sum_i r_i^2,$$
$$\sum_i x_iy_i=\sum_i y_iz_i=\sum_i z_ix_i=0. \quad (20.8)$$

Consequently, expression (20.7) assumes the form

$$E_{2x}=0. \quad (20.9)$$

Similarly, we can prove that $E_{2y}=E_{2z}=0$. Hence, we finally obtain

$$E_2=0. \quad (20.10)$$

Thus, the strength of the local field acting on a molecule inside a dielectric is

$$\mathbf{E}^*=\mathbf{E}+\mathbf{P}/(3\varepsilon_0). \quad (20.11)$$

This formula must be treated only as a first approximation, since a real dielectric differs from the model used for obtaining this formula. In particular, electric fields of molecules may considerably differ from the fields of dipoles, the crystal lattice of a dielectric may have a different symmetry, dipole moments of the molecules may have different directions, etc.

The local field acting on the molecules of a dielectric differs from the external field because the dielectric placed in an external field itself becomes a source of an additional field.

www.EngineeringBooksPDF.com

Molecular dielectric susceptibility does not depend noticeably on the density of the material or on temperature.
The permittivity of a nonpolar dielectric may depend on temperature only implicitly, through the temperature dependence of molecular concentration.
The local field acting on the molecules of a dielectric differs from the external field since the dielectric itself in the external field becomes a source of an additional field.

Which basic factors are responsible for the difference in dielectric properties of rarefied and dense gases? What are these differences?
What physical factors determine the independence of the permittivity of nonpolar dielectrics from temperature over a sufficiently wide range?

Sec. 21. Nonpolar Dielectrics

Basic properties of nonpolar dielectrics are described.

Molecular dielectric susceptibility. It follows from the mechanism of creation of the induced dipole moment of a molecule (see Sec. 17) that its direction coincides with the direction of the electric field. To a first approximation, the dipole moment of a molecule can be considered proportional to the field strength:

$$\mathbf{p} = \alpha\varepsilon_0\mathbf{E}^*, \tag{21.1}$$

where α characterizes the "polarizability" of a molecule (or atom) and is called the molecular (or atomic) dielectric susceptibility. It is determined by the intrinsic properties of the molecule. In view of strong intrinsic electric fields in the molecule, the molecular dielectric susceptibility is small and does not noticeably depend on the density of the substance and temperature. The magnitude of α can be estimated by proceeding from the following model of molecular polarization. A molecule is represented in the form of a conducting sphere whose radius is approximately equal to the molecular radius ($a = 10^{-10}$ m). In a constant field $\mathbf{E}^*$, this sphere acquires a dipole moment [see (16.82)] given by

$$\mathbf{p} = 4\pi\varepsilon_0 a^3\mathbf{E}^*. \tag{21.2}$$

Comparing (21.2) with (21.1), we obtain the following expression for the dielectric susceptibility:

$$\alpha = 4\pi a^3. \tag{21.3}$$

If for molecular radii we use the values, obtained in the kinetic theory, formula (21.3) *gives a slightly exaggerated but still correct (in order of magnitude) value of* α. Hence, this model of molecular polarization is quite suitable for order-of-magnitude estimates.

Using (21.1), we find the polarization

$$\mathbf{P} = \frac{1}{\Delta V}\sum_{\Delta V}\alpha\varepsilon_0\mathbf{E}^* = \alpha\varepsilon_0\mathbf{E}^*\frac{1}{\Delta V}\sum_{\Delta V}1 = \alpha\varepsilon_0 N\mathbf{E}^*. \tag{21.4}$$

www.EngineeringBooksPDF.com

Here

$$\sum_{\Delta V} 1 = \Delta V N, \tag{21.5}$$

where N is the molecular concentration.

Rarefied gases. In this case, the local field strength $\mathbf{E}^*$ differs but slightly from the external field strength $\mathbf{E}$. Hence [see (21.4)], we have

$$\mathbf{P} = \alpha \varepsilon_0 N \mathbf{E}. \tag{21.6}$$

Comparing (21.6) and (17.11), we conclude that the dielectric susceptibility is

$$\varkappa = \alpha N. \tag{21.7}$$

Taking (17.31) into account, the relative permittivity $\varepsilon_r = \varepsilon/\varepsilon_0$ can then be presented in the form

$$\varepsilon_r = 1 + \alpha N. \tag{21.8}$$

The value of ε_r differs from unity by the value of αN, which is very small for gases. For example, the molecular concentration of air under normal conditions is $N = 2.6 \times 10^{25}$ m^{-3}. Considering, in accordance with (21.3), that $\alpha \simeq 10^{-29}$ m^3 for molecules, we find

$$\alpha N \simeq 10^{-3}. \tag{21.9}$$

The value of α, and hence αN, increases with molecular size, remaining small in order of magnitude.

The value of ε_r may depend on temperature only implicitly through the temperature dependence of N. We denote by N_A, ρ_m and M the Avogadro constant, gas density and molar mass respectively and write the obvious equality

$$N = N_A \rho_m / M. \tag{21.10}$$

Using this relation, we can write (21.8) in the form

$$(\varepsilon_r - 1)\, M/\rho_m = \alpha N_A. \tag{21.11}$$

Consequently, $(\varepsilon_r - 1)/\rho_m$ *is a constant quantity independent of temperature and pressure only if the pressure is sufficiently low.* As the pressure rises, the density increases, and we must take into account the difference between local and external fields.

Dense gases. In this case, we must use expression (20.11) in formula (21.4) for $\mathbf{E}^*$:

$$\mathbf{P} = \alpha \varepsilon_0 N [\mathbf{E} + \mathbf{P}/(3\varepsilon_0)], \tag{21.12}$$

whence

$$\mathbf{P} = \frac{\alpha \varepsilon_0 N}{1 - \alpha N/3} \mathbf{E}. \tag{21.13}$$

Substituting (21.13) into (17.29), we obtain

$$\mathbf{D} = \varepsilon \mathbf{E} = \varepsilon_0 \mathbf{E} + \frac{\alpha \varepsilon_0 N}{1 - \alpha N/3} \mathbf{E}, \tag{21.14}$$

www.EngineeringBooksPDF.com

whence

$$\frac{3(\varepsilon_r-1)}{\varepsilon_r+2}=\alpha N. \qquad (21.15)$$

This expression is called the Clausius-Mosotti formula. It can be used to represent (21.10) in the form

$$\frac{3(\varepsilon_r-1)}{\varepsilon_r+2}\frac{M}{\rho_m}=\alpha N_A. \qquad (21.16)$$

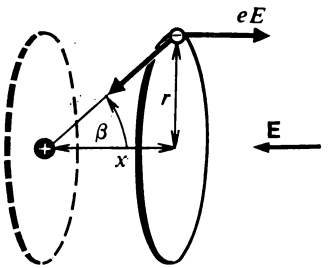


Fig. 101. To the calculation of atomic dielectric susceptibility of hydrogen

The left-hand side of this expression does not depend on temperature and pressure within the limits in which the molecular susceptibility remains constant. For gases, such pressures can be high (of the order of 100 MPa). In liquids and solids having high densities, the value of α depends on pressure. Formula (21.16) was experimentally verified for a wide range of pressure. For example, for carbon dioxide, which is a nonpolar gas, the validity of the Clausius-Mosotti relation (21.16) was checked to a high degree of accuracy up to pressures of the order of 100 MPa at 100°C. Over the entire pressure range, the relative deviation of the left-hand side from the constant value does not exceed a few hundredths. In this case up to pressures of about 20 MPa, a small increase is observed in the value of the left-hand side of (21.16), while above this value, the left-hand side of this expression slightly decreases. The relative permittivity ε_r here varies considerably (by a factor of 1.5) in the pressure range from 1 to 100 MPa.

Example 21.1. *Estimate the atomic dielectric susceptibility α of a hydrogen atom. The electric field is directed normally to the plane in which the electron moves* (Fig. 101).

We write the equilibrium condition for an electron moving in an external field:

$$eE=\frac{e^2}{4\pi\varepsilon_0(x^2+r^2)}\cos\beta=\frac{e^2}{4\pi\varepsilon_0}\frac{x}{(x^2+r^2)^{3/2}}. \qquad (21.17)$$

For $x \ll r$, we obtain $x/(x^2+r^2)^{3/2} = x/r^3$, and hence [see (21.17)]

$$ex = 4\pi\varepsilon_0 r^3 E = p,$$

whence

$$\alpha = 4\pi r^3 \simeq 1.57\cdot10^{-30}\ \mathrm{m}^3,$$

which gives the correct order of the dielectric susceptibility of a hydrogen atom.

Sec. 22. Polar Dielectrics

Basic properties of polar dielectrics are described.

Temperature dependence of polarization. The constant dipole moment for most molecules is of the order of 10^{-29}-10^{-30} C·m. For example, it is equal to 0.36×10^{-30} C·m for CO, 5.3×10^{-30} C·m for SO_2, and 3.5×10^{-29} C·m for KCl.

www.EngineeringBooksPDF.com

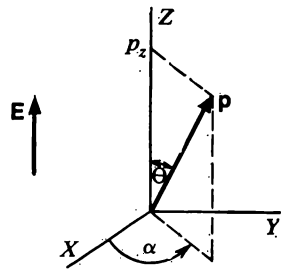


Fig. 102. Dipole orientation in a spherical coordinate system

The dipole moments of many molecules have been measured and compiled in tabular form.

The potential energy corresponding to a dipole with a dipole moment **p** in an electric field **E** is given by

$$W = -\mathbf{p}\cdot\mathbf{E}. \tag{22.1}$$

This quantity attains its minimum value when the direction of the dipole orientation coincides with the direction of the electric field. Since the state of a system with minimum energy is stable, *the dipole moments of polar molecules tend to orient themselves along the electric field vector.* The required rotation is realized by the couple acting on the dipole (see Fig. 91). Thermal motion, however, disturbs the ordering action of the electric field. As a result, a certain equilibrium sets in.

Let us direct the Z-axis along the electric field **E** (Fig. 102). The potential energy (22.1) of molecules depends on the angle between the directions of their dipole moment and the field vector:

$$W = -pE\cos\theta = -p_zE \tag{22.2}$$

and hence the angular distribution of dipole moments is characterized in this case by the Boltzmann distribution The number dn of molecules whose dipole moments are contained within the solid angle $d\Omega$ is given by

$$dn = Ae^{\frac{pE\cos\theta}{kT}}d\Omega = Ae^{\frac{pE\cos\theta}{kT}}d\alpha\sin\theta\,d\theta. \tag{22.3}$$

Then the average value of the Z-component of the dipole moments is

$$\langle p_z\rangle = \frac{\int p_z\,dn}{\int dn} = \frac{Ap\int_0^{2\pi}d\alpha\int_0^{\pi}e^{\beta\cos\theta}\cos\theta\sin\theta\,d\theta}{A\int_0^{2\pi}d\alpha\int_0^{\pi}e^{\beta\cos\theta}\sin\theta\,d\theta}, \tag{22.4}$$

where $p_z = p\cos\theta$ and the notation

$$\beta = pE/(kT) \tag{22.5}$$

has been introduced.

First of all, we must calculate the internal integral in the denominator of (22.4):

$$I = \int_0^{\pi}e^{\beta\cos\theta}\sin\theta\,d\theta, \tag{22.6}$$

www.EngineeringBooksPDF.com

since the internal integral in the numerator is given by

$$\int_0^{\pi} e^{\beta \cos\theta} \cos\theta \sin\theta \, d\theta = \partial I/\partial\beta. \tag{22.7}$$

Integral (22.6) can easily be evaluated:

$$I = \int_0^{\pi} e^{\beta \cos\theta} \sin\theta \, d\theta = -\frac{1}{\beta} e^{\beta \cos\theta}\Big|_0^{\pi} = \frac{2}{\beta} \sinh\beta, \tag{22.8}$$

whence

$$\frac{\partial I}{\partial\beta} = \frac{2}{\beta}\left(\cosh\beta - \frac{1}{\beta}\sinh\beta\right). \tag{22.9}$$

Thus, formula (22.4) combined with (22.8) and (22.9) assumes the form

$$\langle p_z \rangle = pL(\beta), \tag{22.10}$$

where $L(\beta) = \coth\beta - 1/\beta$ is the **Langevin function** (Fig. 103).

For not very strong fields, when $pE \ll kT$, i.e. $\beta \ll 1$, we expand the hyperbolic cotangent into the series

$$\coth\beta = 1/\beta + \beta/3 - \beta^3/45 + \ldots \tag{22.11}$$

and confine ourselves to the term linear in β in the expression for $L(\beta)$:

$$L(\beta) = \beta/3. \tag{22.12}$$

This gives

$$\langle p_z \rangle = p^2E/(3kT). \tag{22.13}$$

Saturation field. As the field strength increases, the dipole moments are oriented in the direction of the field more and more intensely, and when $pE \gg kT$, i.e. for $\beta \gg 1$, we can assume that all dipole moments are parallel to each other and are directed along the field. Consequently,

$$\langle p_z \rangle = p. \tag{22.14}$$

This relation can be obtained from (22.10) if we take into account that, for $\beta \gg 1$, the function $L(\beta)$ is close to unity:

$$L(\beta \to \infty) \to 1. \tag{22.15}$$

Maximum possible polarization is attained when condition (22.14) *is satisfied, and any further increase in the field does not lead to a higher polarization.* The field at which the maximum possible polarization is reached is called the **saturation field**. Considering that the order of magnitude of dipole moments is equal to 10^{-29} C·m, we conclude that for $T = 300$ K the saturation field strength is

$$E_s \approx kT/p \simeq 4.2 \cdot 10^8 \text{ V/m}. \tag{22.16}$$

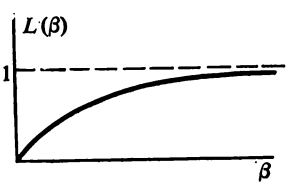


Fig. 103. The Langevin function

www.EngineeringBooksPDF.com

Hence it follows that the condition $pE \ll kT$ under which formula (22.13) is valid is fulfilled up to field strengths of the order of million volts per metre. Therefore, in most practically important cases, we can use formula (22.13).

Rarefied gases. In this case, the local field strength can be assumed to be equal to the external field strength and the polarization [see (22.13)] can be represented in the form

$$\mathbf{P} = Np^2\mathbf{E}/(3kT). \tag{22.17}$$

Further, in complete analogy with the calculations carried out by formulas (21.6)-(21.8), we find that the relative permittivity is

$$\varepsilon_r = 1 + Np^2/(3kT\varepsilon_0). \tag{22.18}$$

In addition to polarization due to a reorientation of constant dipole moments, polar dielectrics also possess polarization due to induced dipole moments, described by formula (21.8). Hence, taking into account both polarization mechanisms, we obtain the following expression for ε_r of polar gaseous dielectrics under a moderate pressure:

$$\varepsilon_r = 1 + N\,[\alpha + p^2/(3kT\varepsilon_0)]. \tag{22.19}$$

It can be seen from (21.3) that $\alpha = 10^{-29}\ \mathrm{m}^3$. On the other hand, at room temperature $kT \simeq 4 \times 10^{-21}\ \mathrm{J}$, and hence for $p \simeq 10^{-29}\ \mathrm{C \cdot m}$ $p^2/(3kT\varepsilon_0) \simeq 10^{-27}\ \mathrm{m}^3$, i.e. *the contribution from induced dipole moments to polarization amounts to about* 1/100 *of the value due to constant dipole moments and hence can be ignored. However, it is possible at present to make precise measurements which allow us to distinguish between the contributions to polarization from constant and induced dipole moments.* For this purpose, ε_r is measured over a wide range of temperature and formula (22.19) is employed. The ε_r vs. $1/T$ dependence is represented on the graph by a straight line. Its intersection with the axis of ordinates for $1/T = 0$ gives $\varepsilon_r = 1 + \alpha N$. Hence we can calculate $\alpha = (\varepsilon_r - 1)/N$. After this, we can use the results of measurement for other values of $1/T$ and calculate the constant dipole moment with the help of formula (22.19) since all other quantities in this equation are known.

Quantum interpretation of polarization of polar gaseous dielectrics. In quantum theory, as well as in the classical one, the polarization of polar dielectrics is explained by the predominant orientation of constant magnetic moments of molecules in the direction of the electric field. The permittivity of dielectrics is described by formula (22.19). *However, there is an essential difference in the interpretations of the reorientation of constant dipole moments according to classical and quantum theories.*

In the quantum theory, it is necessary to take into account the rotation of molecules. The angular momenta of rotating molecules are oriented in various directions in space and their projections onto any particular direction form a discrete set of values, the mean value of the projection being equal to zero. The electric dipole moment is rigidly connected with a molecule and changes its spatial orientation as a result of molecular rotation.

www.EngineeringBooksPDF.com

The dipole moment of a molecule can be decomposed into two components: along the axis of rotation and normal to it. As a result of rotation, the latter component changes its spatial orientation in the plane perpendicular to the axis of rotation of the molecule. In the coordinate system in which this molecule rotates, the mean value of this component is equal to zero. The mean value of the dipole moment component along the axis of molecular rotation is also equal to zero due to the fact that the moment of inertia of the molecule is quantized, and the mean value of its projection onto any direction is equal to zero regardless of whether or not the electric field is present. Consequently, *molecules with nonzero angular momentum do not contribute to polarization. Polarization is created only by nonrotating molecules with zero angular momentum as a result of reorientation of their constant electric dipole moments.* The projections of dipole moments onto the direction of the electric field form a discrete set of values with a nonzero mean, which explains the appearance of the polarization.

Dense gases. In this case, we must bear in mind the difference between the local field and the external field and different orientations of dipole moments, which depends on the interaction between dipoles. This considerably complicates the calculations.

Assuming that the local field is much weaker than the saturation field, it is expedient to write the following formula for polarization instead of (22.17):

$$\mathbf{P}=\frac{Np^2}{3kT}\mathbf{E}^*. \tag{22.20}$$

However, the local field $\mathbf{E}^*$ in it cannot be expressed in terms of the external field through formula (20.11). This can be verified by the following considerations.

Suppose that we place a dipole $\mathbf{p}$ at the centre of a spherical cavity of radius a, formed in a dense dielectric having the relative permittivity ε_r. The field of this dipole polarizes the medium outside the sphere. Thus, an additional field

$$\mathbf{E}_{\text{add}}=\frac{2(\varepsilon_r-1)}{2\varepsilon_r+1}\frac{\mathbf{p}}{4\pi\varepsilon_0 a^3} \tag{22.21}$$

appears in the spherical cavity, i.e. there appears a constant field coinciding in direction with the dipole moment. *This additional field creates an additional induced dipole moment which coincides in direction with the constant dipole moment, and hence cannot reorient it. Consequently, polarization cannot be interpreted as a reorientation of dipole moments in the local field.*

Considering (20.11), we can give formula (22.20) the form

$$\mathbf{P}=\frac{Np^2}{3kT}\left[\mathbf{E}+\frac{\mathbf{P}}{3\varepsilon_0}\right] \tag{22.22}$$

whence

$$\mathbf{P}=\frac{Np^2/(3kT)}{1-Np^2/(9kT\varepsilon_0)}\mathbf{E}. \tag{22.23}$$

When $T_0 = Np^2/(9k\varepsilon_0)$, the denominator on the right-hand side vanishes. For $T > T_0$, the polarization $\mathbf{P}$ has a finite value, while for $T = T_0$ it tends to

www.EngineeringBooksPDF.com

infinity. This means that for $T \leqslant T_0$ the corresponding quantity must possess a spontaneous polarization. For example, analyzing formula (22.23) we can expect that water vapour should be spontaneously polarized, which is obviously wrong. Similar erroneous results are obtained for other materials also. *Consequently, different models are required for the description of dense gases with polar molecules and polar liquids.*

Polar liquids. Onsager proposed a model for polar liquids which is in better agreement with experiment although it gives rather approximate numerical results. In this model, each dipole is assumed to be at the centre of a **real** spherical cavity whose volume is equal to the mean volume per molecule. The model takes into account the dipole orientation by long-range forces and the appearance of an additional dipole moment under the effect of the field (22.12). As a result, the following relation was obtained:

$$\frac{(\varepsilon_r - \varepsilon_{\text{ind}})(2\varepsilon_r + \varepsilon_{r\,\text{ind}})}{\varepsilon_r(\varepsilon_{r\,\text{ind}} + 2)^2} = \frac{Np^2}{9kT\varepsilon_0} \tag{22.24}$$

where ε_r is the relative permittivity, and $\varepsilon_{r\,\text{ind}}$ is the relative permittivity due to induced dipole moments. For water, $\varepsilon_{r\,\text{ind}} = 4.9$, $p = 2.16 \times 10^{-29}$ C·m, and formula (22.24) gives $\varepsilon_r = 105$ at $T = 273$ K. The experimental value of $\varepsilon_r = 88$. We can hardly expect a better agreement with experiment.

A better quantitative agreement with experimental results was obtained for highly dilute solutions of polar dielectrics in a nonpolar solvent. In this case, polar molecules of a dissolved substance are located at sufficiently large distances from each other, and the interaction between them can be ignored. Using the Onsager model, it is possible to take into account the interaction between polar molecules and a nonpolar solvent. This leads to the theory which is fairly in agreement with experiment.

Ionic crystals. These materials can be represented as combinations of two sublattices with positive and negative ions. Under the action of an external electric field, these lattices are displaced relative to each other, as a result of which a considerable polarization appears. This gives comparatively high values of the relative permittivity ε_r. For example, $\varepsilon_r = 6$ for common salt, 5 for KCl, etc.

Under usual conditions, the saturation fields (viz. the fields at which the polarization of a polar dielectric attains the maximum possible value) amount to hundreds of millions volts per metre.

The contribution from induced dipole moments to the polarization is about one hundredth of that from permanent dipole moments and in most cases can be ignored.

Taking into account the local field, the mechanism of polarization of dense polar gases and liquids cannot be interpreted as the reorientation of dipole moments in this field.

Why do the dipole moments of polar molecule tend to align with the electric field?

Under what conditions does the polarization of polar dielectrics attain saturation?

To what distances between elementary charges do the permanent dipole moments of molecules correspond?

Do the modern experimental techniques allow us to separate the contributions of permanent and induced dipole moments to the polarization? Explain how this can be done in principle.

What physical factors make it impossible to treat the polarization of dense polar dielectrics as the result of reorientation of dipole moments in the local field?

www.EngineeringBooksPDF.com

Sec. 23. Ferroelectrics

The physical properties of ferroelectrics and the nature of ferroelectricity are considered.

Definition. Ferroelectrics are polar dielectrics which are spontaneously polarized in a certain temperature interval. In other words, they possess polarization in the absence of electric field. As a result of phase transition, a ferroelectric is transformed into a polar dielectric at the boundaries of this temperature interval.

Relative permittivity of ferroelectrics is extremely high ($\varepsilon_r \simeq 10^4$) and *depends on the field strength, although it is not a single-valued function of the field.* The value of ε_r depends on the variation of the field strength before a given value has been attained.

The term "ferroelectrics" is explained by a formal analogy existing between their properties and the properties of ferromagnetics. Examples of ferroelectrics are Rochelle salt $NaKC_4H_4O_6 \cdot 4H_2O$ and barium titanate $BaTiO_3$.

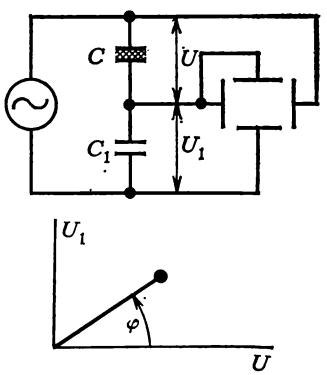


Fig. 104. Schematic diagram of the circuit for eliminating the hysteresis loop: $\tan \varphi = \varepsilon/\varepsilon_0 = D/\varepsilon_0 E$

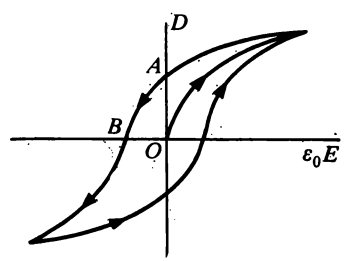


Fig. 105. Hysteresis loop

Hysteresis loop. Since ε depends on E, $D = \varepsilon E$ depends nonlinearly on E. Besides, since ε is determined by the past history of variation of E, D depends ambiguously on E. Let us place a ferroelectric between the plates of a capacitor and measure ε depending on the field strength E which varies according to a harmonic law.

The schematic diagram of the circuit is shown in Fig. 104. The outer terminals of two series-connected parallel-plate capacitors are connected to a

www.EngineeringBooksPDF.com

generator which creates a harmonically varying potential difference across the terminals. The potential difference is distributed between the capacitor C containing a ferroelectric and the capacitor C_1 with no material between its plates. Assuming that the areas of all capacitor plates are equal and denoting by d the distance between the plates, we have

$$E = \sigma/\varepsilon, \quad E_1 = \sigma/\varepsilon_0, \tag{23.1}$$

whence

$$U = Ed = \sigma d/\varepsilon, \quad U_1 = E_1 d = \sigma d/\varepsilon_0 \tag{23.2}$$

and

$$\tan \varphi = U_1/U = \varepsilon/\varepsilon_0 = \varepsilon E/(\varepsilon_0 E). \tag{23.3}$$

Therefore, if the voltage U is applied to the horizontal sweep and U_1 to the vertical sweep of an oscillograph, the variation of E will be registered on the screen by a curve the abscissa of whose points is equal to $\varepsilon_0 E$ on a certain scale, and the ordinate is equal to $\varepsilon E = D$ on the same scale. This curve is called the **hysteresis loop** (Fig. 105). The arrows on the curve indicate the direction of motion of a point along the curve upon the change in the field strength. The segment OA characterizes the residual polarization, i.e. the polarization of the sample for zero external field. The segment OB characterizes the field vector directed against the polarization, at which the sample is completely depolarized, i.e. its residual polarization vanishes. The larger the magnitude of $|OA|$, the higher the residual polarization of a ferroelectric. The larger the value of $|OB|$, the better the residual polarization is retained by the ferroelectric.

Curie point. As the temperature of a ferroelectric *exceeds a certain value* T_C, *typical of each material, its ferroelectric properties vanish and it becomes an ordinary polar dielectric.* The point marking the transition between ferroelectric phase and polar electric phase is called the **Curie point**, and the temperature T_C corresponding to it is known as the **Curie temperature**. In certain cases there are two Curie points, i.e. ferroelectric properties vanish with decreasing temperature also. For example, Rochelle salt has two Curie points corresponding to the temperatures $T_{C\,up} = 24°C$ and $t_{C\,l} = -18°C$. The number of ferroelectrics with two Curie points is comparatively small. Most ferroelectrics have only the upper point which is simply called the Curie point.

At the Curie point, a dielectric goes over from the ferroelectric state to the state of a polar dielectric. The permittivity in this process varies continuously from the value corresponding to the ferroelectric phase to the value corresponding to the phase of polar dielectric. The law of variation of dielectric susceptibility $\varkappa$ in the vicinity of the Curie point has the form

$$\boxed{\varkappa = \frac{A}{T - T_0}} \tag{23.4}$$

where A is a constant and T_0 is the **Curie-Weiss temperature** which is close to the Curie temperature T_C (in formula (23.4), T_C is often used instead of T_0, which does not introduce any significant error in the value of $\varkappa$ for temperatures

www.EngineeringBooksPDF.com

differing from T_C). The law expressed by formula (23.4) is called the **Curie-Weiss law.**

If the lower Curie point also exists, the Curie-Weiss law in the vicinity of this point has the form

$$\varkappa = \frac{A'}{T_0' - T}. \tag{23.5}$$

As was mentioned above, crystals exhibit different dielectric properties in different directions, and hence their dielectric susceptibility is characterized by the dielectric susceptibility tensor $\varkappa_{ij}$ instead of the scalar dielectric susceptibility $\varkappa$. However, the temperature dependence of tensor components is the same as in (23.4) and (23.5).

Molecular mechanism of spontaneous polarization. The theory of ferroelectricity is beyond the scope of the course of general physics. For this reason, we shall confine ourselves to only a qualitative description of processes on molecular level. A very strong interaction between dipole moments of molecules may lead to the appearance of a finite polarization **P** at as small a field strength **E** as desired or, which is the same, the polarization **P** may exist in the absence of an electric field. In other words, *a very strong interaction between dipole moments of molecules causes spontaneous polarization which is characterized by the same orientation of separate dipole moments.* Taking into account the fact that permanent dipole moments are much larger than the induced ones [see (22.19)], we conclude that spontaneous polarization is characterized by a very high polarization. Consequently, the corresponding susceptibility $\varkappa$ and permittivity ε have considerably higher values than those observed for polar and nonpolar dielectrics. *The state of spontaneous polarization is precisely the ferroelectric state.* A transition from the ferroelectric phase to the polar dielectric phase is the transition from the state of spontaneous polarization to the state when spontaneous polarization vanishes and the substance becomes an ordinary dielectric whose molecules have permanent dipole moments. In other words, this is the transition to the polar dielectric state. *The physical factors responsible for this transition ultimately weaken the interaction between the dipole moments of molecules.*

Dielectric domains. Spontaneous polarization is a source of strong electric fields. Consequently, if a macroscopic volume of a ferroelectric is polarized spontaneously in a certain direction, a very strong electric field appears arround this volume, and a high energy is associated with it. Such a state is disadvantageous from the point of view of energy. *The system tends to go over to a state characterized by a spontaneous polarization on the one hand and the minimum field energy on the other. This can be realized as a result of division of the volume of the ferroelectric into small regions each of which is spontaneously polarized in a certain direction. These directions are different for different regions.* The average polarization of the volume containing a sufficient number of small regions with different direction of spontaneous polarization is equal to zero, and hence the strength of the external electric field created by this volume is close to zero. *Small regions of spontaneous polarization are called dielectric domains, or simply domains.* Thus,

www.EngineeringBooksPDF.com

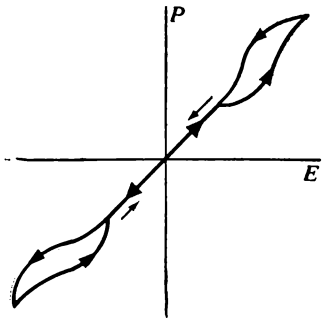


Fig. 106. Double hysteresis loops for antiferroelectrics which become ferroelectrics in strong fields

a nonpolar ferroelectric is an aggregate of domains with randomly oriented spontaneous polarizations.

Obviously, in order to decrease the electric energy, it is expedient to decrease the volumes of domains. *However, the process of decreasing the domain size is hampered by another factor associated with the presence of the surface energy at the boundary between neighbouring domains.* Obviously, the total surface area of domain boundaries increases upon a decrease in the domain volume, and hence the surface energy also increases. For this reason, the volumes of domains may decrease only to certain limits corresponding to a decrease in the total energy of the system. *Upon a further decrease in the domain volume, the total energy increases rather than decreases at the expense of the surface energy. This determines the domain sizes.* These sizes are of the order of thousands of intermolecular distances. The existence of domains is proved in experiments involving direct observations with the help of polarized light, as well as in the experiments on etching of a ferroelectric, since different parts of a domain are destroyed at different rates upon etching.

The variation of polarization of a ferroelectric in an external electric field involves the reorientation of dipole moments of individual domains, a change in the domain volumes, and displacement of domain boundaries. These processes are thoroughly investigated since ferroelectrics have wide practical applications. More than one hundred pure ferroelectrics and a very large number of ferroelectric solid solutions are known to date.

Antiferroelectrics. Under certain conditions, two spontaneous polarizations with opposite orientations appear in a crystal. One of them appears due to the orientation of dipole moments of molecules of one crystal sublattice in a certain direction, while the other is due to the orientation of dipole moments of the other crystal sublattice in the opposite direction. In this case, the total polarization of any physically small volume of such a crystal is equal to zero. Thus, there are no domains with different orientations of spontaneous polarization, although spontaneous polarization exists in any physically small volume. Such materials are called **antiferroelectrics**. In their structure, they are similar to antiferromagnetics.

In sufficiently weak fields, antiferroelectrics behave as ordinary dielectrics with a linear dependence of polarization on the external field strength. In sufficiently strong fields, the substance may go over to the ferroelectric phase with all the consequences following from this. In particular, a hysteresis loop is observed. The transition occurs in a strong electric field. For this reason, if we replace an antiferroelectric by a ferroelectric in the circuit depicted in Fig. 104, two hysteresis loops will be observed (Fig. 106) at large amplitudes of voltage oscillations.

www.EngineeringBooksPDF.com

The Curie-Weiss temperature does not coincide with the Curie temperature but is close to it. In many cases, there is no need to distinguish between them.
Most ferroelectrics have only one (upper) Curie point. However, there are several ferroelectrics with two Curie points.

Sec. 24. Piezoelectrics

The mechanisms of piezoelectric and inverse piezoelectric effects are described. The relationship between inverse piezoelectric effect and electrostriction is analyzed. Basic information on pyroelectric materials is given.

Properties of piezoelectrics. There are many crystals whose surfaces acquire *electric charges upon deformation.* Such crystals are called **piezoelectrics.** Since deformation itself cannot alter the total charge of the crystal, ***the surface charges induced as a resuls of deformation have opposite signs on different parts of the surface.*** Piezoelectrics include quartz, tourmaline, Rochelle salt, and many other materials.

Experience shows that charges appear on the surface of a piezoelectric as a result of a uniform compression or extension in quite definite directions called the **polar axes of the piezoelectric.** On opposite faces perpendicular to a polar axis, charges of opposite signs appear under a uniform strain. The signs of the charges are reversed together with the sign of the deformation. If, for example, the compression along a polar axis has resulted in the appearance of a positive charge on a given face, this face will acquire a negative charge as a result of extension along the same axis. Piezoelectric effect is observed not only due to a pure compression or extension along a polar axis, but also upon any deformation of the crystal, which is accompanied by an extension or compression along a polar axis.

Since charges of opposite sign appear on different faces perpendicular to a polar axis, *different directions along the polar axis are not equivalent.* This means that if a crystal is rotated through 180° around an axis perpendicular to the polar axis, the polar axis will coincide with itself but the crystal will not. Consequently, crystals having a centre of symmetry cannot be piezoelectrics. The necessary condition for the piezoelectric effect to exist upon uniform deformation is the absence of a centre of symmetry in the crystal. Polar axes are determined by the symmetry properties of the crystal lattice. Generally, a crystal has several polar axes.

Piezoelectric properties depend on temperature. If at a certain temperature the crystal lattice is rearranged so that a centre of symmetry is formed, piezoelectric properties of the crystal vanish at this temperature. For example, piezoelectric properties in quartz change insignificantly up to the temperature of 200 °C. Then, up to the temperature of 576°C, these properties slowly become less and

www.EngineeringBooksPDF.com

less pronounced. At 576 °C, the crystal lattice in quartz is rearranged, as a result of which piezoelectric properties disappear. As the temperature decreases, the piezoelectric properties of quartz change in the reverse order.

Longitudinal and transverse piezoelectric effects. The appearance of charges on the faces perpendicular to a polar axis upon uniform deformation along this axis is called the **longitudinal piezoelectric effect.** However, the charges can be induced on the same surfaces by compressing or extending the crystal in a direction perpendicular to the polar axis only if extension or compression along this axis is observed. This phenomenon is called the **transverse piezoelectric effect**. *It owes its existence to the relation between longitudinal and transverse deformations of a rigid body.*

Mechanism of piezoelectric effect. Only ionic crystals may possess piezoelectric properties. *Piezoelectric effect appears when the crystal sublattice of positive ions is deformed by external forces not as the crystal sublattice of negative ions.* As a result, the positive and negative ions are displaced relative to each other, which leads to the polarization of the crystal and the appearance of surface charges. To a first approximation, the polarization is directly proportional to the strain which, in turn, is proportional to the force. Consequently, polarization is proportional to the applied force. The potential difference appearing between oppositely charged faces can be measured, and its value can be used to estimate the strain and applied forces. This relation finds numerous practical applications. For example, piezoelectric transducers are used for measuring rapidly varying pressures. Piezoelectric microphones are well known and piezoelectric transducers are widely used in automation and telemetry.

Inverse piezoelectric effect. It consists in the deformation of a piezoelectric introduced into an external electric field. *This effect owes its existence to the direct effect and to the law of conservation of energy.* When a piezoelectric is deformed, work is required to increase the energy of elastic deformation and the energy of the electric field appearing as a result of the piezoelectric effect. Consequently, in deforming a piezoelectric, it is necessary to overcome besides the elastic force of the crystal an additional force, which hampers the deformation and is a factor responsible for the inverse piezoelectric effect. In order to compensate for this additional force, we should apply an external electric field opposite to that appearing as a result of the piezoelectric effect. Consequently, in order to deform the piezoelectric to a certain extent by an external field, this field must be equal and opposite to the field that would appear under the given deformation due to the direct piezoelectric effect. For example, if a certain potential difference appears between the faces of a piezoelectric, which are perpendicular to its polar axis, upon a deformation along this axis, a potential difference of the same magnitude but of opposite sign must be applied to these faces in order to attain the same deformation without applying mechanical forces.

The mechanism of the inverse piezoelectric effect is similar to that of the direct effect: under the action of an external field, the crystal sublattices of positive and negative ions are deformed differently, which causes a deformation of the crystal.

www.EngineeringBooksPDF.com

The inverse piezoelectric effect also has numerous practical applications. In particular, quartz ultrasonic vibrators are widely used.

Pyroelectrics. A sublattice of positive ions in some piezoelectrics turns out to be displaced relative to the sublattice of negative ions in the state of thermodynamic equilibrium. As a result, *such crystals are polarized in the absence of an external electric field.* Thus, these crystals possess a spontaneous electric polarization.

Usually, the presence of such a spontaneous polarization is masked by free surface charges induced on the surface of the crystal from the surrounding medium by electric field due to spontaneous polarization. This process occurs until the electric field is completely neutralized, i.e. until the presence of spontaneous polarization is totally masked. However, as the temperature of the sample changes, for example, as a result of heating, the ionic sublattices become displaced relative to one another, which causes a change in spontaneous polarization, and electric charges appear on the surface of the crystal. The appearance of these charges is called the direct pyroelectric effect, and the corresponding crystals are called pyroelectrics.

Every pyroelectric is a piezoelectric, but the converse is not true. This is due to the fact that a pyroelectric has a preferred direction along which spontaneous polarization takes place, while a piezoelectric generally does not have such a direction.

The inverse pyroelectric effect is also known to exist: a variation of the electric field in an adiabatically isolated pyroelectric is accompanied by a change in its temperature. The existence of the inverse effect can be proved on the basis of a thermodynamic analysis of the process and be demonstrated experimentally.

When conditions are suitable for spontaneous polarization, a dielectric tends to go over to such a state in which, on the one hand, spontaneous polarization exists and, on the other, the field energy is minimum. Under these conditions, domains are formed.
The factors that weaken the interaction of dipole moments of molecules cause the disappearance of spontaneous polarization and the transition from the ferroelectric state to the state of a polar dielectric.

What is the difference between the Curie and the Curie-Weiss temperature?
What is the mechanism behind the domain formation? Why cannot domains be very large?
What materials are called antiferroelectrics?

Problems

3.1. Calculate the relative permittivity of helium at $p = 101.3$ kPa, $t = 15$ °C, if its atomic dielectric susceptibility $\alpha = 2.48 \times 10^{-30}$ m^3. The experimental value of ε_r is 1.000074.

3.2. Calculate the permittivity of ammonia at $t = 27$ °C, for $\alpha = 1.37 \times 10^{-29}$ m^3, and the dipole moment $p = 0.46 \times 10^{-29}$ C·m.
Hint. Use formula (22.19).

3.3. The permanent dipole moment of water molecule is 6.2×10^{-30} C·m. Find the polarization of saturated water vapour at $t = 100$ °C under atmospheric pressure.

3.4. Air mainly consists of N_2 and O_2 molecules. Using the Clausius-Mosotti formula, calculate their atomic susceptibilities assuming for simplicity that they are equal. Find the radius of molecules.

www.EngineeringBooksPDF.com

3.5. Taking the values of α and r_0 obtained in Problem 3.4 for nitrogen molecules, calculate the change in the distance between the charges forming the dipole in a field of strength 1 MV/m.

Answers

3.1. $\varepsilon_r = 1.000067$. 3.2. $\varepsilon_r = 1.0076$. 3.3. 1.2×10^{-4} C/m^2. 3.4. $\alpha = 1.1 \times 10^{-29}$ m^3, $r_0 = 0.96 \times 10^{-10}$ m. 3.5. 0.87×10^{-16} m.

www.EngineeringBooksPDF.com

CHAPTER 4

Direct Current

A direct current cannot be generated by forces which are of purely electrostatic origin. In order to create a direct current, we must have nonelectrostatic forces called extraneous electromotive forces. The basic law for direct current is Ohm's law in local forms.

Sec. 2.5. Electric Field in the Case of Direct Currents

The peculiarities of the electric field created by direct currents and the role of surface and volume charges are discussed. The part played by various factors in ensuring the existence of direct current is analyzed.

The field in a conductor. In differential form, Ohm's law (see Sec. 16) can be written as follows:

$$\mathbf{j} = \gamma \mathbf{E}. \tag{25.1}$$

If a current is flowing, $\mathbf{j} \neq 0$ and hence $\mathbf{E} \neq 0$. Thus, an electric field exists inside a current-carrying conductor. It should be recalled that in electrostatics, there is no field inside a conductor.

Generally speaking, the distribution of direct current density over the conductor cross section is not uniform. In order to verify this, let us consider a segment of a bent conductor with a circular cross section (we are considering a homogeneous conductor for which γ = constant). We assume that the curved part of the conductor is cut from an undeformed piece of a material, since a bent conductor is under strain and, strictly speaking, the homogeneity condition is not satisfied for it. This complicates the entire pattern of the distribution of current density.

Near the surface of the conductor, the current density can be directed only along the tangent to the surface. This means [see (25.1)] that the field $\mathbf{E}$ near the surface of the conductor is tangential to the surface. Consequently, the equipotential surfaces are perpendicular to this surface. If the part of the conductor under consideration is bent, two close equipotential surfaces obviously cannot be at a fixed distance from each other at all points inside the conductor. For example, if the conductor is in the shape of a ring of circular cross section, the distance between the equipotential surfaces at the inner part of the ring will be smaller than at the outer part. Since the distance between the neighbouring equipotential surfaces changes, the electric field strength at the corresponding points on the equipotential surface also changes. Hence [see (25.1)], the density

www.EngineeringBooksPDF.com

of the direct current in a uniform conductor generally varies over its cross section. The equipotential surfaces in a very long right circular cylindrical conductor are the planes perpendicular to the cylinder axis. Consequently, the electric field strength and the current density are constant over the entire cross section of such a homogeneous conductor.

We shall be considering conductors with a very small cross-sectional area, called **linear** conductors. To a fairly high degree of accuracy, we can neglect

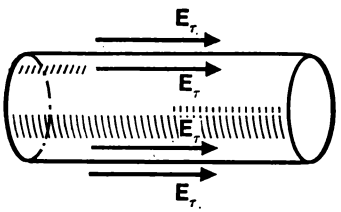


Fig. 107. Field inside a conductor and the tangential component of the field near the outer surface of the conductor

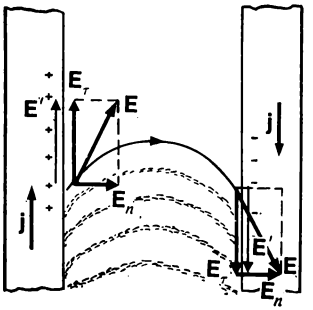


Fig. 108. Demonstration of the presence of the normal field component near the surface of a conductor

the variation in the current density over the cross section of such a conductor and assume that it is constant in magnitude at each point of this cross section and is directed along an element dl of the conductor. In this case, the current flowing through the conductor will be equal to $I = j\,\Delta S$, where ΔS is the cross-sectional area of the conductor.

Thus, the question of the electric field and density of direct current in thick conductors is quite complicated in general. The distribution of current density over the cross section depends on several factors and, in particular, on the shape of the conductor. More definite statements can be made for the field in the vicinity of the surface of a conductor. Near the surface, the field as well as the current density are directed along the tangent to the surface. There are no components of these quantities inside the conductor that are normal to the surface. From the boundary condition (17.30), we can conclude that *in the vicinity of the surface outside the conductor there exists an electric field whose tangential component* $\mathbf{E}_\tau$ *is equal to the tangential component* $\mathbf{E}_\tau$ *of the field inside the conductor* (see Fig. 107). However, it is impossible to draw any conclusions about the normal component of the field outside the conductor from here.

The sources of a field. What creates an electric field inside a conductor? In other words, what is the source of this field? Since the existence of a direct current in a circuit is ensured by a current source in the circuit, say, a galvanic cell, it has obviously got something to do with the generation of the electric field. However, this source cannot generate the field directly. Such a statement is quite obvious for a very long conductor and for parts of the circuits which are

www.EngineeringBooksPDF.com

at a very large distance from the cell, say, several hundred kilometers. The field which can be created by the charges across the terminals of this cell at such a distance is infinitely small. Consequently, the cell cannot directly act as the source of the electric field in a conductor.

The electric charge can be the only source of the electric field in a conductor. Hence, the problem under consideration is reduced to determining the charges which generate an electric field inside the conductor, as well as their location.

Field outside a conductor. In order to investigate this problem, we must consider the field outside the conductor. We place a current-carrying conductor in a plane tray with a thin layer of dielectric powder (Fig. 108). In this case, the grains of the powder are aligned in chains along the lines of force of the electric field (see Sec. 19). The figure shows two parts of the current-carrying conductor and the lines of force between them.

It can be seen that field lines are not tangential to the surface of the conductor. This means that outside the conductor near its surface, we have the tangential component $\mathbf{E}_\tau$ of the field as well as the normal component $\mathbf{E}_n$. Inside the conductor, however, $\mathbf{E}_n = 0$. Hence, we can conclude on the basis of (17.26) that the *surface of the conductor must bear charges whose surface density is given by*

$$\sigma = \varepsilon_0 E_n. \tag{25.2}$$

Here it is assumed that the conductor is in vacuum. If the conductor is in a dielectric medium, we must replace ε_0 by ε, the permittivity of the medium.

Surface charges. Thus, the surface of a conductor through which a direct current is flowing bears electric charges. *These charges are sources of the field which exists in the conductor and ensures direct current in the conductor.* The surface density of the charge may have different signs in different parts of the conductor. For example, the left and right parts of the conductor in Fig. 108 have positive and negative surface charge density respectively.

Volume charges. Only surface charges exist in homogeneous conductors. *In nonhomogeneous conductors, where the electric conductivity varies from point to point, charges appear in the bulk of the conductor also.* This follows directly from the law of charge conservation (5.24). In the stationary case under consideration $(\partial\rho/\partial t) = 0$, and Eq. (5.24) assumes the form

$$\operatorname{div} \mathbf{j} = 0. \tag{25.3}$$

In principle, the volume charge in matter can be free or bound. We are interested in the total volume charge density $\rho + \rho_b$, which is responsible for variation of the electric field along the conductor. Consequently [see (17.27)], the total volume charge density is equal to

$$\rho + \rho_b = \operatorname{div} (\varepsilon_0 \mathbf{E}) = \varepsilon_0 \operatorname{div} (\mathbf{j}/\gamma), \tag{25.4}$$

where $\mathbf{E} = \mathbf{j}/\gamma$. Considering (25.3) and the expression

$$\operatorname{div} (\mathbf{j}/\gamma) = (1/\gamma) \operatorname{div} \mathbf{j} + \mathbf{j} \cdot \operatorname{grad} (1/\gamma), \tag{25.5}$$

we obtain from (25.4)

$$\rho + \rho_b = \varepsilon_0 \mathbf{j} \cdot \operatorname{grad} (1/\gamma). \tag{25.6a}$$

www.EngineeringBooksPDF.com

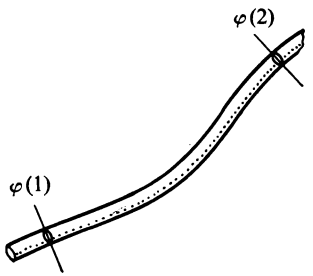


Fig. 109. To the calculation of the potential difference between two points of a current-carrying conductor

Directing the X-axis along the rectilinear part of the conductor and assuming that its properties vary only in this direction, we can write formula (25.6a) in the form

$$\rho + \dot{\rho}_b = \varepsilon_0 j \frac{\partial(1/\gamma)}{\partial x}. \qquad (25.6b)$$

If the conductivity decreases in the direction of the current flow, the volume charge density is positive. This is due to the following reason. For a constant cross-sectional area of the conductor, the current density along the conductor must be constant. If the conductivity decreases in the direction of the current, we must increase the field strength to keep the current constant. It is this increase in the field strength that is responsible for the positive volume charges. The emergence of negative volume charges when the conductivity increases in the direction of the current flow can be explained in the same way.

Mechanism of generating direct currents. The current source is called the source of extraneous electromotive forces (extraneous e.m.f.s; see Sec. 26). According to the results of its action, such a source is a process or a device separating positive charges from negative ones. After being separated, these charges move towards the electrodes and act, in accordance with Coulomb's law, on the charges of the conductor in the vicinity of the electrodes. In turn, these charges act on other charges, and so on. As a result of these collective interactions, charges are distributed on the conductor surfaces in the circuit so that the corresponding electric field appears in the conductor. Thus, *the charges at the terminals of the source of extraneous e.m.f.s do not directly create an appropriate electric field in all conductors, but ensure a surface charge distribution on the conductors such that it generates the required electric field inside them.* This is what creates a direct current. Since the interaction between charges takes place through electromagnetic forces, the flow of a direct current after the circuit is closed is characterized by the velocity of propagation of electromagnetic waves, which in turn depends on the distribution of capacitances, inductances, and other characteristics of the circuit. In the free space, the velocity of propagation of electromagnetic waves is equal to the velocity of light.

Change in potential along a current-carrying conductor. Since $\mathbf{E} \neq 0$ in a conductor carrying direct current, the potential varies along the conductor. In other words, unlike in electrostatics, the potential is not constant at all points of the conductor. *However, the field inside the conductor is created by immobile surface charges which do not vary with time. Consequently, as in electrostatics, this field is a potential field,* and the potential difference between two points on the conductor (Fig. 109) is given, in accordance with formula (14.28), by

$$\varphi(2) - \varphi(1) = -\int_{(1)}^{(2)} \mathbf{E} \cdot d\mathbf{l}, \qquad (25.7)$$

www.EngineeringBooksPDF.com

where the integral is calculated along any path joining points *1* and *2*. For the sake of convenience in calculations, it is expedient to choose for this path one of the current lines joining a certain point in the cross section *1* to the corresponding point in the cross section *2*. Along such a line, **E** and **dl** are collinear and hence $\mathbf{E} \cdot \mathbf{dl} = E\mathrm{d}l$, the positive sign indicating that the current flows in the direction from a higher to a lower potential. Besides, if the cross-sectional area of the conductor is constant, $E = \text{const}$ along the conductor. Consequently [see (25.7)],

$$\varphi(1) - \varphi(2) = El \tag{25.8}$$

where l is the length of the conductor between the cross sections *1* and *2*. The potential difference between the cross sections is called voltage and is denoted by $U_{12} = \varphi(1) - \varphi(2)$. From the differential form of Ohm's law ($\mathbf{j} = \gamma\mathbf{E}$), we get

$$E = j/\gamma = jS/(\gamma S) = I/(\gamma S), \tag{25.9}$$

where I is the current Taking this relation into account, we can write (25.8) as follows:

$$U_{12} = Il/(\gamma S) = IR_{12}, \tag{25.10}$$

where $R_{12} = l/(\gamma S)$ is the ohmic resistance of the part of the conductor between *1* and *2*. This formula represents Ohm's law for a subcircuit.

Sec. 26. Extraneous Electromotive Forces

The role of extraneous e.m.f.s in current-carrying circuits is discussed and specific sources of extraneous e.m.f.s are described.

The origin of extraneous e.m.f.s. An extraneous e.m.f. cannot be of electrostatic origin for the simple reason that electrostatic field is a potential field. Consequently, the work done by the field in a closed current-carrying loop is equal to zero. Under such a condition, there can be no current, since it must perform work in order to overcome the ohmic resistance of the conductors. The existence of a direct current proves that *extraneous electromotive forces are of a nonelectrostatic origin.*

In particular, an extraneous e.m.f. can be mechanical or electrical, but not electrostatic. An example of such an e.m.f. is the force acting on a charge in an electric field resulting from Faraday's law of electromagnetic induction (see Chap. 8).

Mechanical extraneous e.m.f. Figure 110 shows the circuit for the simplest current source in which the extraneous e.m.f. is of mechanical origin. The space between electrodes A and B is filled with a neutral medium having the same number of positive and negative charges. A nonelectrostatic extraneous force

www.EngineeringBooksPDF.com

draws positive charges to electrode B and negative charges to electrode A. As a result, electrode A acquires a negative charge, while electrode B becomes positively charged. In the outer circuit, a current flows from B to A and performs a certain amount of work. The energy required for this purpose is provided by extraneous forces which perform work to distribute the charges between electrodes A and B and to bring these charges to the respective electrodes against the forces of the electric field of strength E existing between the electrodes. The

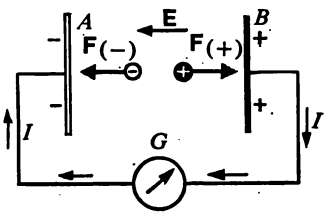


Fig. 110. Schematic diagram of extraneous e.m.f.s of mechanical origin

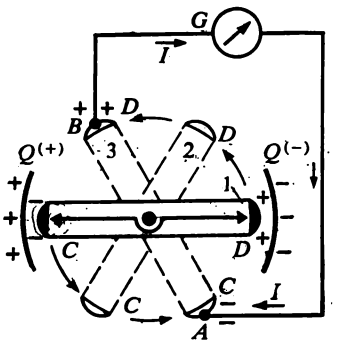


Fig. 111. Schematic diagram of the Wimshurst machine

current between electrodes A and B inside the e.m.f. source closes the external circuit. *If the direction of current is determined relative to the electrodes, the current in the external circuit flows from a positive to a negative electrode, while inside the source, the current flows from a negative to a positive electrode.*

A mechanical extraneous e.m.f. can be created with the help of the Wimshurst machine shown in Fig. 111. Charges Q^+ and Q^- generate an electrostatic field in the space between them. Mutually insulated conducting plates C and D move in a circle around an axis perpendicular to the plane of the figure under the action of extraneous mechanical forces. In the position *1*, the plates are connected through a fixed conductor (solid line with arrows at the ends). As a result of electrostatic induction, the plates C and D acquire negative and positive charges respectively in this position. Upon further rotation, their contact with the conductor is broken and they become isolated from each other in the position *2*, carrying at the same time unlike charges. In the position *3*, these strips come in contact with the electrodes A and B, and impart their charge to these electrodes. An electric current flows in the circuit BGA between the electrodes. If we have one pair of rotating conductors CD, the current in the circuit flows in pulses at the rate of two pulses per revolution. If, however, we take quite a large number of pairs of plates like C and D, so that they come in contact with the electrodes A and B successively at negligibly small time intervals, a nearly direct current will flow in the external circuit. Such a machine produces

www.EngineeringBooksPDF.com

an extraneous e.m.f. of mechanical origin, generated by mechanical forces that ensure the motion of the plates C and D in the circle.

The chain of mutual conversions of energy in this case looks like this. The extraneous mechanical forces which move the plates C and D perform work against the forces of an electrostatic field existing between the charges Q^+ and Q^- and transfer the charge from the plates to the electrodes A and B. This results in a change in the energy of the electric field, i.e. in conversion of mechanical energy to the electrical field energy. As the current flows in the circuit

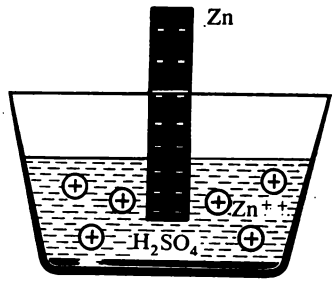


Fig. 112. Emergence of a potential difference between a solid and a liquid

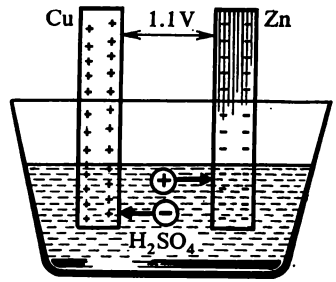


Fig. 113. Voltaic cell

BGA, this energy is converted into Joule's heat and other forms of energy on account of the work performed by the current in the external circuit.

Galvanic cells. Galvanic cells and accumulators are the most widely used sources of direct current. Electric current was discovered in 1791 by L. Galvani (1737-1798). However, Galvani could not provide a correct explanation for his experimental results. This was done by A. Volta (1745-1827) in 1792. The direct current cells which we shall be describing here were named after Galvani.

A potential difference (see Sec. 2) is created not only when two solids are brought into contact, but also when solids come in contact with liquids. This may be accompanied by chemical reactions. For example, if a zinc plate Zn (Fig. 112) is immersed in a solution of H_2SO_4, it is dissolved in the acid solution. However, not neutral zinc atoms but positive Zn^{++} ions pass into the solution. As a result, the solution becomes positively charged while the plate acquires a negative charge. A potential difference is thus created between the plate and the solution. At a certain potential of the metal with respect to the solution (called the electrochemical potential), zinc ions no longer pass into solution. This potential depends on the properties of the metal, liquid, and on the concentration of ions in the solution. Upon coming in contact with water, a metal acquires a larger negative charge than in contact with a salt solution containing ions of this metal. For a large concentration of ions in the solution, the reverse process may take place, when positive ions start depositing on the plate which thus gets positively charged. Thus, *for different combinations of metals*,

www.EngineeringBooksPDF.com

solutions and different concentrations of ions in solutions, different electrochemical potentials can appear.

Since the electrochemical potential depends on the concentration of metal ions, it is customary to take a solution whose one litre contains a mole of metal ions divided by the valency of the ions. The electrochemical potential of a metal relative to such a solution is called absolute (normal) electrochemical potential. For example, the absolute electrochemical potential for Zn in a sulphuric acid solution is equal to —0.5 V, while for Cu this value is equal to +0.6 V.

If two different metals are immersed in a solution, a potential difference equal to the difference in their electrochemical potentials is created between

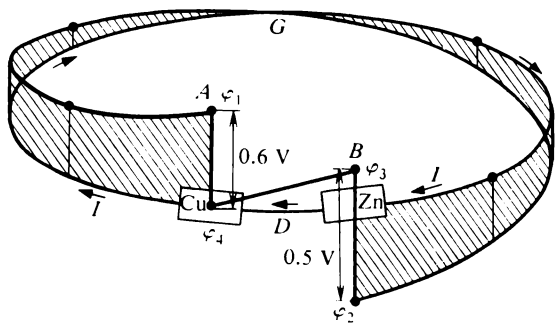


Fig. 114. Change in potential in a circuit containing a voltaic cell

them. *The set of two metals in a solution is called a galvanic cell, and the potential difference between the metals is called the electromotive force of the cell.*

Voltaic cell. A voltaic cell consists of a zinc plate and a copper plate immersed in a solution of sulphuric acid (Fig. 113). Taking into consideration the electrochemical potentials of zinc and copper, we conclude that the e.m.f. of a voltaic cell is equal to [0.6—(—0.5)] V = 1.1 V.

Range of action of extraneous e.m.f.s. It should not be thought that extraneous e.m.f.s are generated in the space between the zinc and copper plates. *In this case, we have two extraneous e.m.f.s concentrated in the surface layers of contact between the zinc and copper plates and the solution. The thickness of these layers is of molecular size. There are no extraneous e.m.f.s in the remaining volume of the solution.* If the two plates are connected by a metallic conductor, a current will flow through the latter from the copper plate, which is a positive electrode, to the zinc plate which serves as a negative electrode. Inside the solution, the current flows from the zinc plate to the copper plate. Thus, the lines of direct current are closed, as expected.

Let us consider the variation of potential in a current-carrying circuit. The potential drops across the ohmic resistance of the conductor in the direction of the current. Figure 114 shows the change in potential over a closed circuit containing a voltaic cell as the source of extraneous e.m.f. The points *A* and *B*

www.EngineeringBooksPDF.com

respectively correspond to the surface layers of the copper and zinc plates in contact with the solution over which extraneous electromotive forces act. The difference between these forces constitutes the extraneous e.m.f. of the cell and is equal to the total potential drop across the ohmic resistance of the external circuit in the section AGB and across the ohmic resistance of the electrolyte in the section BDA. The ohmic resistance of the electrolyte is called the internal resistance of the cell. We denote by $\mathscr{E}_{ext}$, R and r the extraneous e.m.f. of the cell, the resistance of the external circuit and the internal resistance of the cell respectively. For the entire circuit, Ohm's law can be written in the form

$$\mathscr{E}_{ext} = I(R + r). \tag{26.1}$$

The extraneous e.m.f. of a cell is determined by the properties of the cell and is independent of the current passing through the circuit. It can be seen from formula (26.1) that the voltage drop in the external circuit ($U = IR$) is not equal to the electromotive force of the cell and is always less. This is the voltage between the terminals of the working cell when a current flows in the circuit. As the current increases, the voltage in the external circuit decreases, the decrease being the more significant, the higher the internal resistance of the cell. While using a cell it is always desirable that the voltage in the external circuit should depend on current, i.e. on the load as little as possible. Hence the internal resistance is an important characteristic of a cell. The lower the internal resistance, the better the quality of the source of extraneous e.m.f., other conditions being equal.

Law of conservation of energy. Let us analyze the law of conservation of energy in the circuit with current shown in Fig. 114. We denote by A_1 the work done by the electric field as a charge q moves in the closed circuit, and by A_2 the work done by the extraneous e.m.f.s. The electric field performs work in the sections where the potential drops from φ_1 to φ_2 (external circuit) and from φ_3 to φ_4 (due to the ohmic resistance offered by the solution to current in the cell). This work is

$$A_1 = (\varphi_1 - \varphi_2)\, q + (\varphi_3 - \varphi_4)\, q. \tag{26.2}$$

The work done by extraneous e.m.f.s in layers of molecular thickness leads to an increase in potential from φ_4 to φ_1 (on the copper plate) and from φ_2 to φ_3 (on the zinc plate). Hence the work done by the extraneous e.m.f.s is given by

$$A_2 = (\varphi_1 - \varphi_4)\, q + (\varphi_3 - \varphi_2)\, q = (\varphi_1 - \varphi_2)\, q + (\varphi_3 - \varphi_4)\, q, \tag{26.3}$$

where the second equality is obtained as a result of regrouping the terms. It can be seen from a comparison of (26.2) and (26.3) that

$$A_1 = A_2, \tag{26.4}$$

i.e. *the work done in the circuit as the current flows in it is equal to the work done by the extraneous e.m.f.s.*

www.EngineeringBooksPDF.com

Let us derive once again Ohm's law for the entire circuit by using Ohm's law (25.10) for a part of the circuit:

$$\varphi_1 - \varphi_2 = IR, \quad \varphi_3 - \varphi_4 = Ir, \tag{26.5}$$

whence

$$IR + Ir = (\varphi_1 - \varphi_2) + (\varphi_3 - \varphi_4) = (\varphi_1 - \varphi_4) + (\varphi_3 - \varphi_4) = \mathscr{E}_{ext}. \tag{26.6}$$

Polarization of a cell. As a current flows through a circuit containing a voltaic cell, Zn^{++} ions pass into solution where they combine with negative SO_4^{--} ions, liberated along with H_2^{++} ions as a result of dissociation of sulphuric acid. The reaction $Zn^{++} + SO_4^{--} = ZnSO_4$ takes place in the solution and the reaction products precipitate from the solution. The positive hydrogen ions rush towards the copper plate where they are neutralized by the electrons of the conduction

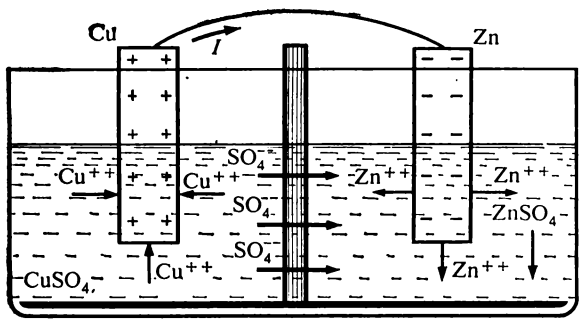


Fig. 115. Daniell cell

current in the plate. Thus a hydrogen film is formed on the copper plate. On the one hand, this film increases the internal resistance of the cell, while on the other hand, it creates an additional electrochemical potential directed against the potential which existed there before the formation of the hydrogen film. As a result of these processes, the e.m.f. of the cell drops. Such processes are called polarization of the cell.

Methods of depolarization. Various methods of depolarization are used to avoid a drop in the e.m.f.

1. Two liquids are selected in such a way that no new materials are deposited at the electrodes. A suitable liquid is chosen for each electrode. The liquids are separated by a porous partition which, on the one hand, prevents them from mixing and, on the other hand, does not obstruct the ion exchange. For example, the two liquids chosen for a Daniell cell are $CuSO_4$ and $ZnSO_4$ solutions (Fig. 115). The copper plate is immersed into the $CuSO_4$ solution while the zinc plate is immersed into the $ZnSO_4$ solution. Zinc passes into the H_2SO_4 solution in the form of Zn^{++} ions. Electrons from the copper plate pass into the copper sulphate solution and neutralize Cu^{++} ions, as a result of which copper is deposited from the solution on the copper plate. The SO_4^{--} ions remaining in the solution pass

www.EngineeringBooksPDF.com

through the partition to the other part of the cell where they combine with Zn^{++} to form an excess of $ZnSO_4$, which precipitates at the bottom. Thus, no polarization takes place during the operation of the cell, but the copper sulphate solution is slowly depleted and must be replenished from time to time.

2. Strong oxidizing agents, which combine hydrogen and oxygen to form water, are also used.

Accumulator. This is a galvanic cell in which the substances consumed during its operation as a current source are accumulated when a direct current is passed through it. Such a procedure is called the charging of the accumulator.

The most widely used accumulator is a lead-acid cell consisting of two lead plates immersed in sulphuric acid solution. In this case, $PbSO_4$ is formed at the electrodes and saturates the entire solution. The passage of current through the accumulator during charging is accompanied by the oxidation of the lead of the electrode connected with the positive terminal of the battery charger to PbO_2 and the reduction of the other electrode to pure lead. Thus, a charged accumulator has one plate with PbO_2 and the other plate made of pure lead, the electrolyte being the solution of H_2SO_4 saturated with $PbSO_4$. During operation of the accumulator, the plate with PbO_2 serves as the positive electrode and is gradually reduced, yielding $PbSO_4$. The negative plate, made of pure lead, is gradually covered by a layer of lead sulphate. As a result, the accumulator is discharged. The e.m.f. of a fully charged lead accumulator is about 2.7 V. However, it drops to about 2.2 V after a brief period of discharging, and remains at this level for quite a long time, dropping very slowly during the operation. The lowest permissible e.m.f. required for complete restoring of the properties of the accumulator as a result of charging is 1.85 V. The accumulator gets spoiled when discharged to lower values of e.m.f.

An important characteristic of an accumulator is its capacity, defined as the total charge released by the accumulator during discharging and measured in ampere-hours.

A nonelectrostatic force capable of separating charges is called an extraneous e.m.f. The work performed in a circuit during the passage of an electric current is equal to the work of extraneous electromotive forces. Generally speaking, the distribution of direct current over the cross section of a conductor is not uniform.

The surface of a current-carrying conductor contains charges which are the sources of an electric field. The field exists in the conductor and ensures the passage of a direct current.

Surface charges on different parts of a conductor may have different signs.

The role of the charges at the terminals of an extraneous e.m.f. source is not to directly create a corresponding field in all the conductors, but to ensure a distribution of surface charges on the conductors which generates the required field in them. Volume charges are induced only in nonuniform conductors.

www.EngineeringBooksPDF.com

Sec. 27. Differential Form of Joule's Law. Work Done during the Passage of Current and Power Developed

The formulas for the work done during passage of current and for the developed power are introduced. The differential form of Joule's law is given. The classical electron pattern of electric conductivity is described and its disadvantages are discussed. General features of quantum-mechanical treatment of electrical conductivity are considered.

Work performed during passage of current. Power. The amount of work performed in transferring a charge dQ between two points with a potential difference U is

$$dA = U dQ. \tag{27.1}$$

Suppose that a current I flows through a conductor. Let us consider a part of this conductor, the potential difference between whose ends is equal to U. During the time dt, a charge $dQ = I\,dt$ is transported over this part of the conductor. Consequently, the work done in this case is

$$dA = IU\,dt. \tag{27.2}$$

Hence, the power developed by the current in this part is defined by the formula

$$\boxed{P = dA/dt = IU.} \tag{27.3}$$

The form of energy liberated in this case depends on the nature of physical factors responsible for the potential drop. The potential drop across the ohmic resistance of wires is accompanied by liberation of heat, the potential drop across the terminals of a d.c. motor is due to the mechanical work performed in this case, and so on. Forumla (27.3) gives the total power developed by the current in the part of the circuit with the potential drop U. If the entire potential drop takes place on the ohmic resistance of the conductor, then, according to Ohm's law, $U = IR$, where R is the resistance of the subcircuit. In this case, the entire energy is liberated in the form of heat with the power

$$\boxed{P = IU = I^2R.} \tag{27.4}$$

Formula (27.4) expresses **Jouly's law**, discovered by J. P. Joule (1818-1889) in 1841 and subsequently investigated in detail by F. A. Lenz.

www.EngineeringBooksPDF.com

Differential form of Joule's law. Applying law (27.4) to a very small cylinder (Fig. 116) whose axis coincides with the direction of the current, we obtain

$$\Delta P = (j\Delta S)^2 \frac{1}{\gamma} \frac{\Delta l}{\Delta S}, \qquad (27.5)$$

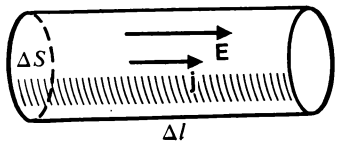


Fig. 116. To the derivation of Joule's law in differential form

where $I = j\Delta S$, j being the current density. The resistance of this cylinder $\Delta R = \Delta l/(\gamma \Delta S)$. Considering that $\Delta S \Delta l = \Delta V$ is the volume of the cylinder, we obtain from (27.5)

$$P_V = \Delta P/(\Delta l\ \Delta S) = j^2/\gamma, \qquad (27.6)$$

where P_V is the volume density of the thermal power liberated in the conductor, i.e. the heat liberated in 1 m³ of the conductor in 1 s. Formula (27.6) is the differential form of Joule's law, since all quantities refer to the same point.

With the help of the differential form of Ohm's law, we can transform (27.6) as follows:

$$\boxed{P_V = j^2/\gamma = \gamma E^2 = \mathbf{j} \cdot \mathbf{E}.} \qquad (27.7)$$

Any of these equalities with P_V on the left-hand side is the differential form of Joule's law. Although formula (27.6) has been derived for a very small cylindrical segment of a conductor, its validity is not related to the shape of the very small volume since the quantities appearing in it depend only on their values at the point and not on any other factors.

The source of energy for the work done by current. The potential drop in the circut with current is compensated by the corresponding increase in the potential as a result of the action of extraneous electromotive forces on the charges (see Sec. 26). The current passing in the circuit performs work, and energy is liberated, for example, in the form of heat. Extraneous electromotive forces perform work over the charges, imparting a certain energy to them. Hence it follows that the *entire work of current is performed at the expense of the energy of extraneous electromotive forces.*

Derivation of Ohm's law from the electron pattern of electrical conductivity. In the framework of classical concepts, the mechanism of passage of current through a conductor and its heating are described as follows.

A free electron is accelerated by the field in a conductor. For a moving electron, Newton's law has the form

$$ma = eE, \qquad (27.8)$$

where m, a and e are the mass, acceleration and charge of the electron respectively. The actual motion of the electron is quite complicated since electrons are in a random thermal motion. Under the action of an external field, all electrons acquire the same acceleration and an additional velocity in the same direction. This results in an ordered motion of electrons, i.e. an electric current. We are interested only in this ordered motion of electrons which is superimposed on the random thermal motion. Moving electrons interact with one another and with

www.EngineeringBooksPDF.com

the atoms of the crystal lattice of the conductor. During their interaction with the atoms of the crystal lattice, electrons exchange with them a small part of their energy. On the average, this energy is acquired by the electrons from the electric field, since, in the absence of an electric field, free electrons and atoms are in thermal equilibrium. This complex pattern of electrons acquiring energy from an electric field and subsequently transferring it to atoms upon interaction can be represented in the following form. Suppose that an electron is accelerated during time τ in accordance with Eq. (27.8), collides with an atom and imparts to it all the kinetic energy acquired during the motion. After this the electron is again accelerated during time τ, again collides with an atom, and so on. In other words, τ is the relaxation time of the nonequilibrium distribution of electrons to thermal equilibrium with the crystal lattice. It is assumed in the model that the mean kinetic energy of electrons increases during this time under the action of an external electric field to values higher than thir mean thermal energy. The excess energy is imparted to the crystal lattice, and thermal equilibrium is restored once again. In actual practice, however, this process takes place continuously, and its gradation is introduced in order to simplify mathematical calculations. The relaxation time τ characterizes the velocity at which the aggregate of electrons and the crystal lattice of the conductor return to thermal equilibrium if the electron equilibrium is somehow disturbed (not just by the external electric field).

In this model, the result of numerous acts of energy transfer from an electron to atoms is replaced by a single act, and hence τ has the sense of the mean interval of time between collisions. If l is the mean free path between collisions and v is the mean velocity of the electron due to its thermal motion, then, by definition,

$$\tau = l/v. \tag{27.9}$$

The path traversed by an electron from its state of rest as a result of acceleration by an electric field is equal to

$$s = \frac{a\tau^2}{2} = \frac{1}{2}\frac{eE}{m_e}\tau^2. \tag{27.10}$$

This is the path traversed on the average by an electron between collisions during time τ in the direction of the electric field. This ordered motion of electrons causes a drift with the velocity

$$v_d = s/\tau = eEl/(2m_e v). \tag{27.11}$$

The drift velocity is inversely proportional to the collision frequency v/l and therefore decreases with increasing temperature.

If n is the electron concetration, we get

$$j = env_d = e^2 ln/E\ (2m_e v). \tag{27.12}$$

A comparison of this formula with Ohm's law $j = \gamma E$ leads to the following expression for the electrical conductivity:

$$\gamma = \frac{1}{2}\frac{e^2 ln}{m_e v}. \tag{27.13}$$

www.EngineeringBooksPDF.com

Thus, we have obtained the correct dependence of current density on electric field strength and the expression for the electrical conductivity in terms of the parameters of motion of free electrons.

Derivation of Joule's law from the electron theory of electrical conductivity. The velocity lost by an electron as a result of a collision is

$$v_t = a\tau = \frac{eE}{m_e}\frac{l}{v}. \tag{27.14}$$

Therefore, the kinetic energy acquired by the electron between collisions and transferred to the conductor atoms upon collision is

$$W_k = \frac{m_e v_t^2}{2} = \frac{1}{2}\frac{e^2E^2l^2}{m_e v^2}. \tag{27.15}$$

The frequency of collisions of each electron with atoms is equal to v/l, and hence the collision frequency of n electrons is equal to nv/l. Consequently, the volume density of thermal power is given by the expression

$$P_V = W_k \frac{nv}{l} = \frac{1}{2}\frac{e^2nl}{m_e v} E^2 = \gamma E^2, \tag{27.16}$$

where we took into account Eqs. (27.13) and (27.15). Thus, we have obtained the correct expression for the differential form of Joule's law by proceeding from the electron theory of electrical conductivity.

Drawbacks of the classical theory of electrical conductivity. The classical theory of electrical conductivity is quite visual and gives a correct dependence of current density and the amount of liberated heat on the field strength. ***However, this theory does not lead to correct quantitative results.*** The main discrepancies between the theory and experiment consist in the following:

(1) In order to obtain the correct value of γ from formula (27.13), we must assume a very large value of l (exceeding the interatomic distance in the conductor by thousands of times). Classical theory fails to explain the existence of such large mean free paths.

(2) Experimental investigation of the temperature dependence of electrical conductivity γ leads to the law $\gamma \propto 1/T$. This cannot be explained by formula (27.13) since the kinetic theory of gases gives $v \propto \sqrt{T}$. The dependence $l \propto 1/\sqrt{T}$ cannot be accepted in the classical model of interaction.

(3) According to the law of equipartition of energy among the degree of freedom, a very large contribution to the specific heat of conductors should be expected from free electrons. This, however, is not observed in experiments.

Main features of quantum-mechanical interpretation of electrical conductivity. The above drawbacks of the classical concepts could be eliminated only in quantum theory. *Quantum theory takes into account wave properties of microparticles.* The diffraction of waves at obstacles is the most important characteristic of wave motion. Consequently, moving electrons as if undergo diffraction at atoms without collisions, and their mean free paths may become quite long. Since electrons obey the Fermi-Dirac statistics, only an insignificant part of electrons near the Fermi level can participate in the formation of the electronic

www.EngineeringBooksPDF.com

specific heat. Therefore, the electronic specific heat of conductors is insignificant. The solution of the quantum-mechanical problem on the motion of an electron in a metallic conductor leads to the dependence $\gamma \propto 1/T$, which is observed in actual practice. Thus, a consistent quantitative theory of electrical conductivity was constructed only in the framework of quantum mechanics.

The work done during the passage of current is not the result of a conversion of the kinetic energy of electrons into other forms of energy. The energy spent in accomplishing the work is carried by the electromagnetic field and not by electrons. Only in a particular case involving the liberation of Joule's heat is the kinetic energy of electrons the intermediate form of energy through which the energy of the electromagnetic field is converted into heat. In other cases, the kinetic energy of electrons does not play any role.

What is the meaning of the mean free time between collisions in the classical theory of electrical conductivity?
What are the principal difficulties of the classical theory of electrical conductivity? How are they eliminated in general?

Sec. 28. Linear Circuits. Kirchhoff's Laws

The laws for calculating linear circuits are formulated.

An isolated closed loop. We have already considered this case in Sec. 26 and obtained the result in the form (26.1): *if an isolated closed loop contains a source of extraneous e.m.f.s, the current in the loop must be such that the total voltage drop across the external and the internal resistance of this source is equal to the extraneous e.m.f.* If there are several sources of extraneous e.m.f.s., we must take their algebraic sum, having chosen a certain direction for the positive e.m.f.

In order to avoid confusion in signs, the following approach is usually adopted. Either clockwise or counterclockwise direction of circumvention of the circuit is taken as positive. In Fig. 117, the clockwise direction of circumvention is considered positive. The electromotive forces of the sources are denoted by $\mathscr{E}_1$, $\mathscr{E}_2$, $\mathscr{E}_3$. The direction in which the current flows is unknown beforehand. Hence, any direction can be chosen for the current. In Fig. 117, for example, it coincides with the positive direction.

Now we must adopt a sign convention. The e.m.f. is considered to be positive if on moving along the loop in the positive direction, we first arrive at the negative terminal. If, however, the positive terminal is encountered first, the corresponding e.m.f. will be negative. The current is assumed to be positive if its direction coincides with the direction of circumvention. In the opposite case the current is negative. The electromotive force and current are therefore algebraic quantities that can assume either positive or negative values. We can now easily

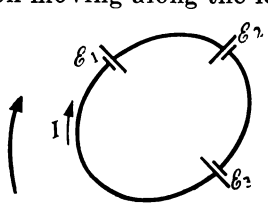


Fig. 117. Isolated closed loop

www.EngineeringBooksPDF.com

generalize Eq. (26.1) to an arbitrary number of extraneous e.m.f. sources in an isolated closed loop: *the product of the magnitude of current and the sum of external and internal resistances of all parts of a closed loop is equal to the sum of the magnitudes of extraneous e.m.f.s in this loop*:

$$\boxed{\pm I \sum_k R_k = \sum_i \pm \mathscr{E}_i,} \tag{28.1}$$

where the $\pm$ symbol in front of I and $\mathscr{E}_i$ indicates that the sign must be chosen in accordance with the above rules. For example, for the case depicted in Fig. 117, Eq. (28.1) has the form

$$I\,(R + r_1 + r_2 + r_3) = \mathscr{E}_1 - \mathscr{E}_2 + \mathscr{E}_3, \tag{28.2}$$

where r_1, r_2, r_3 are the internal resistances of the sources of extraneous e.m.f. s and R is the total resistance of all the parts of the circuit containing no e.m.f.

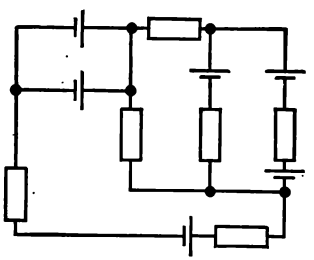


Fig. 118. Electric network

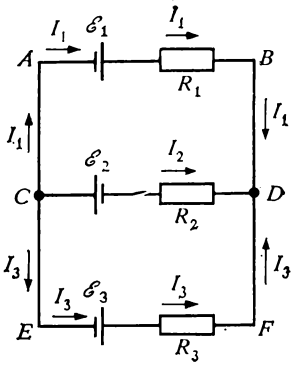


Fig. 119. To the determination of closed loops and junctions in a network

sources. If the arrow showing the direction of current were oriented oppositely for the same direction of circumvention which was taken earlier as positive, we would obtain the following equation instead of (28.2):

$$-I\,(R + r_1 + r_2 + r_3) = \mathscr{E}_1 - \mathscr{E}_2 + \mathscr{E}_3. \tag{28.3}$$

This equation must be solved for I. If the obtained value of I is positive, the current flows in the direction indicated by the arrow. Otherwise, it will flow in the opposite direction.

Branched circuits. The electric circuits encountered in most cases of practical importance are much more complicated as shown in Fig. 118. However, any complex circuit contains the elements of two simplest types:

(1) junctions where more than two conductors meet (Fig. 119, points C and D);

(2) closed loops (Fig. 119, loops *ABDCA*, *CDFEC*, *ABFEA*).

www.EngineeringBooksPDF.com

Kirchhoff's laws. Kirchhoff's laws can be used to write a system of equations from which we can obtain currents for any complex branched circuit. *They express the law of charge conservation at each junction and Ohm's law* (28.1) *for each closed loop*. The sign convention adopted for currents and e.m.f.s for each closed loop is the same as for an isolated loop [see (28.1)]. The direction of positive circumvention for all loops must be the same. The law of charge conservation at a junction requires that the sum of currents entering the junction must be equal to the sum of currents leaving it. In other words, the algebraic sum of currents at a junction must be equal to zero. While composing the sum, the currents shown by arrows pointing away from the junction are assumed, say, to be negative, while the currents shown by arrows pointing towards the junction are taken as positive. Of course, we can choose the opposite signs, but this will not alter the corresponding equations. The main thing is that the same sign rule should be applied to all the junctions.

Thus, Kirchhoff's laws state that

(1) *the sum of algebraic values of currents at each junction is equal to zero*:

$$\boxed{\sum_k (\pm) I_k = 0;} \tag{28.4}$$

(2) *the sum of products of algebraic values of currents and the resistances of the corresponding parts of each closed loop is equal to the sum of the algebraic values of extraneous e.m.f.s in each closed loop*:

$$\boxed{\sum_k (\pm) I_k R_k = \sum_i (\pm) \mathscr{E}_i.} \tag{28.5}$$

It can be shown that the system of equations thus obtained for any branched circuit is complete and can be used to determine all currents.

These laws were derived by G.R. Kirchhoff (1824-1887) who obtained the general solution of the problem on branched d.c. circuits in 1847, although the laws themselves were formulated in 1845.

Let us apply Kirchhoff's laws to the circuit depicted in Fig. 119.

1. In accordance with Kirchhoff's first law, we have

(a) $-I_1 - I_2 - I_3 = 0$ (junction C);
(b) $I_1 + I_2 + I_3 = 0$ (junction D).

2. In accordance with Kirchhoff's second law, we obtain

(a) $I_1 r_1 + I_1 R_1 - I_2 R_2 - I_2 r_2 = \mathscr{E}_1 + \mathscr{E}_2$ (loop $ABDCA$),
(b) $I_2 R_2 + I_2 r_2 - I_3 R_3 - I_3 r_3 = -\mathscr{E}_2 - \mathscr{E}_3$ (loop $CDFEC$).
(c) $I_1 r_1 + I_1 R_1 - I_3 R_3 - I_3 r_3 = \mathscr{E}_1 - \mathscr{E}_3$ (loop $ABFEA$).

Here, r_1, r_2, r_3 are the internal resistances of the extraneous e.m.f. sources. The equations for junctions are identical, while only two of the three equations for loops are independent. For example, the sum of the first two equations yields

www.EngineeringBooksPDF.com

the third equation. Thus, we have got a system of three equations in three unknown currents I_1, I_2 and I_3. The values of the currents and their actual directions can be found by solving this system of equations. However, even without solving it, we can conclude that we were obviously wrong in choosing the directions of currents in Fig. 119, since the law of charge conservation at the junctions cannot be satisfied for the chosen directions of current: a negative charge must be accumulated at the junction C, and a positive charge at the junction D. This, however, should not worry us, since the solution will automatically indicate the direction of current.

Thus, the above example shows that if Kirchhoff's laws are written for all junctions and all loops, we get a larger number of equations than required since not all the equations are independent. In order to avoid additional work, it is desirable not to write superfluous equations. For this purpose, the following rule is adopted. While writing an equation for a closed loop, we must ensure that it contains at least one quantity that did not appear in the previous equations. If all the quantities appearing in this equation have been encountered before, the equation is superfluous. The same rule is observed while writing equations for junctions. For example, in the above equations expressing Kirchhoff's second law, there is no need to write Eq. (b) since all the quantities appearing in it were already encountered in Eq. (a). Similarly, in the equations corresponding to Kirchhoff's first law, Eq. (c) is superfluous since it contains nothing new in comparison with Eqs. (a) and (b). Further control of the correctness of the obtained system of equations can be made by verifying its completeness: the number of equations must be equal to the number of unknowns.

What is the convention of signs in Kirchhoff's laws?
Which considerations must be followed in order to avoid writing superfluous Kirchhoff's equations?

Sec. 29. Currents in a Continuous Medium

The method of calculating currents in continuous media is described.

Formulation of the problem. Electric current can flow not only through conductors. For example, soil (especially damp) also conducts electric current. What is the resistance of the soil if the ends of two conductors connected to the terminals of a source of e.m.f. are thrust into the soil a certain distance apart? Or, what will be the resistance of a massive metallic plate to which two conductors from the terminals of an e.m.f. source are connected? By the resistance of a massive plate or a medium to the electric current we mean the ratio of the potential difference between the current-carrying conductors to the current. Although the electrical conductivity of the medium is known, it is not an easy

www.EngineeringBooksPDF.com

task to calculate the resistance. However, the resistance can be easily measured with the help of standard methods by determining the potential difference and current.

Derivation of the formula. Let us consider a homogeneous continuous medium into which electrodes are immersed. A current flows between the electrodes and the current density lines coincide with the electric field lines in the medium, since

$$\mathbf{j} = \gamma \mathbf{E}. \tag{29.1}$$

The current across a closed surface S surrounding one of the electrodes is

$$I = \oint_S \mathbf{j} \cdot d\mathbf{S} = \gamma \oint_S \mathbf{E} \cdot d\mathbf{S}. \tag{29.2}$$

Let us now imagine that the conducting medium is removed and the electrodes behave like the plates of a capacitor. By definition of capacitance C of a capacitor, we have

$$Q = CU, \tag{29.3}$$

where Q is the charge of an electrode and U is the potential difference between the electrodes. In accordance with the Gauss theorem, we obtain

$$\oint_S \mathbf{E} \cdot d\mathbf{S} = Q/\varepsilon_0, \tag{29.4}$$

where $\mathbf{E}$ is the field of the capacitor and S is the same surface as in (29.2). However, in view of the uniqueness of the solution of problems in electrostatics, the potential difference between given electrodes uniquely determines the field. Consequently, the field in a current-carrying conducting medium [see (29.2)] coincides with the field created in vacuum between the same electrodes for the same potential difference [see (29.4)]. Hence, taking into account Eq. (29.3), we obtain from (29.2) and (29.4):

$$I = \gamma Q/\varepsilon_0 = \gamma C U/\varepsilon_0. \tag{29.5}$$

Thus the resistance offered by a homogeneous medium to the current is given by the formula

$$R = U/I = \varepsilon_0/(\gamma C). \tag{29.6}$$

It should be noted that *all the above discussion is not applicable to nonhomogeneous media*, since the passage of current through them is accompanied by the creation of volume charges which are the sources of electric field. In this case, the electric field generated in a medium during the passage of current is not the same as the field in vacuum, although the same potential difference is maintained between the electrodes.

Conditions of applicability of Eq. (29.6). Formula (29.6) can be used to determine the resistance of a medium to the current if we know the capacitance of the capacitor formed by the electrodes. The accuracy of the results depends on the extent to which the potential of an electrode fluctuates during the passage of

www.EngineeringBooksPDF.com

current through it. If the last requirement is not fully satisfied and different points on the conductor have considerably different potentials during the passage of current, the calculation of the resistance does not boil down to the calculation of the capacitance of the capacitor, since the potential at all points on the plate of the capacitor is the same. Hence, in particular, it is necessary that the resistivity of the electrodes be small in comparison with the resistivity of the medium. This condition is not applicable if the electrodes have small surface areas.

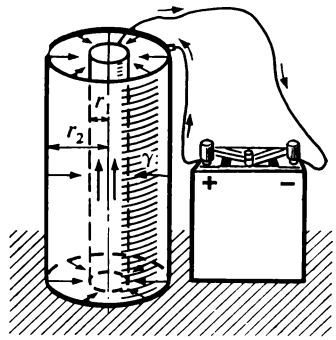


Fig. 120. To the calculation of the resistance of the medium between coaxial electrodes

Coaxial electrodes. By way of an example, let us consider two coaxial electrodes. It is required to calculate the resistance of the conducting medium between these electrodes (Fig. 120). In order to apply formula (29.6), we must assume that the conductivity of the material of the core and shell is much larger than the conductivity of the medium. The current in the medium flows over the entire volume along the radii between the core and the shell. Since the capacitance of a cylindrical capacitor is

$$C = 2\pi l\varepsilon_0/\ln (r_2/r_1),$$

the resistance of the medium is

$$R = \ln (r_2/r_1)/(2\pi l\gamma). \tag{29.7}$$

Nonhomogeneous medium. The problem becomes quite complicated if the conductivity is not constant, since in this case volume charges appear and it is necessary to take into account the electric field generated by them.

As an example, let us consider the electric currents in the atmosphere. It is shown experimentally that an electric field $E_r^0 \simeq -100$ V/m exists near the Earth's surface and is directed towards its centre. The Earth is quite a good conductor and hence it can be assumed that it has a surface charge

$$\sigma_0 = \varepsilon_0 E_r^{(0)} = -8.85 \cdot 10^{-10} \text{ C/m}^2. \tag{29.8}$$

It is borne out by measurements that the conductivity of the Earth's atmosphere increases with height. The main reason behind this is the ionization caused by cosmic radiation. Solar radiation is mainly responsible for the ionization at large heights. At a height of about 50 km, the atmosphere can be practically assumed to be an ideal conductor. Measurements show that the dependence of the conductivity on height can be expressed to a high degree of accuracy in the following form:

$$\gamma (r) = \gamma_0 + A (r - r_0)^2. \tag{29.9}$$

Here r_0 is the Earth's radius, r is the distance between the Earth's centre and the point under consideration, $\gamma_0 = \gamma (r_0)$ is the electrical conductivity at the

www.EngineeringBooksPDF.com

Earth's susface, and A is a constant. The last two quantities have the following values:

$$\gamma_0 = 3 \cdot 10^{-14} \text{ S/m}, \tag{29.10}$$

$$A = 0.5 \cdot 10^{-20} \text{ S/m}. \tag{29.11}$$

On the average, the Earth's field in the atmosphere is stationary and spherically symmetric. Hence the continuity equation for the current density assumes the form

$$\operatorname{div} \mathbf{j} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 j_r) = 0, \tag{29.12}$$

whence

$$j_r(r) = j_0 r_0^2 / r^2, \tag{29.13}$$

where j_0 is the current density near the Earth's surface ($r = r_0$), which is equal to

$$j_0 = \gamma_0 E_r^{(0)} = -3 \cdot 10^{-12} \text{ A/m}^2. \tag{29.14}$$

Since the radius of the Earth is $r_0 \simeq 6 \times 10^6$ m, the current from the atmosphere to the Earth is equal to

$$I = |j_0| \, 4\pi r_0^2 \approx 1400 \text{ A}.$$

At a distance r from the Earth's centre, the electric field is

$$E_r = \frac{j_r(r)}{\gamma(r)}. \tag{29.15}$$

Hence the potential difference U between the Earth's surface and the upper atmosphere whose conductivity is practically infinite is given by the formula

$$U = -\int_{r_0}^{\infty} E_r \, dr = -j_0 r_0^2 \int_{r_0}^{\infty} \frac{dr}{r^2 \gamma(r)}. \tag{29.16}$$

Here, the domain of integration is extended to infinity, since at distances larger than about 50 km, $\gamma(r)$ practically becomes infinite and the integrand vanishes. However, quite accurate results can also be obtained by using expression (29.9) for γ. In this case, the contribution to the integral from the domain of integration for $r > r_0 + 50$ km is quite small in comparison with the contribution from the domain of integration between r_0 and $r_0 + 50$ km and can therefore be neglected. Hence, instead of (29.16) we obtain

$$U = -j_0 r_0^2 \int_{r_0}^{\infty} \frac{dr}{r^2 [\gamma_0 + A(r - r_0)^2]}. \tag{29.17}$$

This integral can be easily evaluated in terms of elementary functions. However, the results of calculations are quite cumbersome and we shall omit them here. To within quantities of the order of $[\gamma_0 (r_0^2 A)] \ll 1$, this result can be ex-

www.EngineeringBooksPDF.com

pressed quite accurately in the form

$$U = -\frac{j_0}{Ar_0}\left[1 + \ln\frac{\gamma_0}{Ar_0^2} + \frac{\pi r_0}{2}\sqrt{\frac{A}{\gamma_0}}\right]. \qquad (29.18)$$

Substituting into this equation the values of j_0, γ_0 and A from (29.14), (29.10) and (29.11), we obtain $U \simeq 400$ kV.

Owing to a direct current of about 1400 A passing through the atmosphere, this potential difference must decrease and the charge on the surface of the Earth must be neutralized. The relaxation time for this process is of the order of $\tau = \varepsilon_0/\gamma_0 \simeq 300$ s. However, both the current and the potential difference are constant on the average. Thus, there are some reasons behind this constancy. Basically, transient atmospheric processes like storms and thunder are responsible for this.

The most important property of earthed transmission lines is that the resistance is independent of the distance between the electrodes. The main contribution to the resistance is made by the regions of the medium in the immediate vicinity of the electrodes.
The formula expressing the resistance of the medium in terms of the capacitance of a capacitor with the electrodes as its plates is valid only provided that, in the presence of a current, the potential at all the points on each plate is constant to a high degree of accuracy, and volume charges do not appear in the medium.
For this purpose, the electrical conductivity of the electrode material should be much higher than that of the medium and the latter should be electrically homogeneous.

Under what condition can the formulas for the resistance of the medium between two electrodes be applied in terms of the capacitance of the capacitor formed by the electrodes?

Sec. 30. Earthing of Transmission Lines

The physical principle behind the possibility of earthing is explained and the conditions necessary for earthing are discussed.

Formulation of the problem. Since the electrical conductivity of the soil is quite high, one can ask whether it is possible to use the Earth as a conductor of electric current. Such an electric circuit is shown in Fig. 121 (A and B are electrodes embedded into the Earth). Obviously, the expenditure on the wires can be reduced by about half in this case.

Calculation of resistance. Let us find the resistance of a continuous medium, assuming that the electrodes are spheres of radius r_0 each. We denote by d the distance between the centres of the electrodes. In order to simplify calculations, we assume that the medium is infinite (Fig. 122) and the charge distribution on the electrodes is spherically symmetric.

www.EngineeringBooksPDF.com

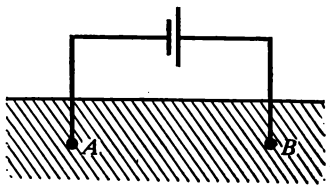


Fig. 121. Earthing of a transmission line

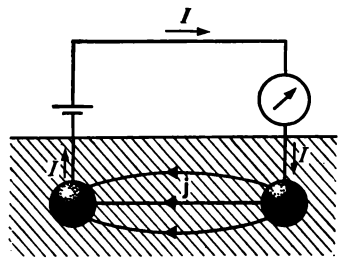


Fig. 122. To the calculation of the resistance of the medium in the case of spherical electrodes

Let x be the distance between the centre of the left electrode and a certain point on the line connecting the centres of the electrodes (Fig. 123). The field at this point is given by

$$E = E_{(+)} + E_{(-)} = \frac{Q}{4\pi\varepsilon_0}\left(\frac{1}{x^2} + \frac{1}{(d-x)^2}\right). \tag{30.1}$$

The potential difference between the electrodes is

$$U = \int_{r_0}^{d-r_0} E\,dx = \frac{Q}{4\pi\varepsilon_0}\left[-\frac{1}{x} + \frac{1}{(d-x)}\right]_{r_0}^{d-r_0}$$

$$= \frac{Q}{4\pi\varepsilon_0}\left(-\frac{1}{d-r_0} + \frac{1}{r_0} + \frac{1}{r_0} - \frac{1}{d-r_0}\right). \tag{30.2}$$

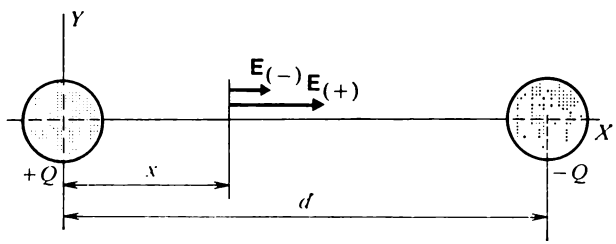


Fig. 123. To the calculation of the resistance of the medium in the case of spherical electrodes

In most practically important cases the distance between the electrodes is much larger than their size, i.e. $d \gg r$. Therefore, we can write Eq. (30.2) in the form

$$U = \frac{Q}{2\pi\varepsilon_0}\frac{1}{r_0}. \tag{30.3}$$

On the basis of what has been said in Sec. 29, we obtain

$$I = \oint_S \mathbf{j}\cdot d\mathbf{S} = \gamma \oint_S \mathbf{E}\cdot d\mathbf{S} = \gamma Q/\varepsilon_0, \tag{30.4}$$

www.EngineeringBooksPDF.com

where I is the current in the medium and S is a closed surface surrounding one of the electrodes. From (30.3) and (30.4), we obtain for the resistance of the medium

$$R = U/I = (2\pi\gamma r_0)^{-1}. \qquad (30.5)$$

The most important property of resistance (30.5) *is its independence of the distance between the electrodes.* Physically, this is explained by the fact that as the distance between the electrodes increases, the effective area of the medium through which the current is passing also increases accordingly. An increase in the distance between the electrodes increases the resistance while an increase in the area decreases it. Formula (30.5) shows that these two factors practically compensate each other and resistance is found to be independent of the distance between the electrodes.

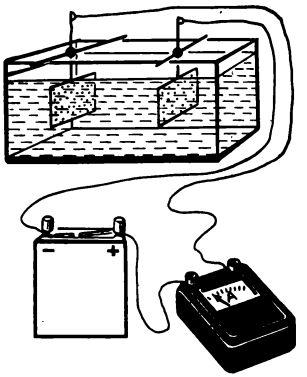


Fig. 124. Demonstration of the independence of the resistance of the medium from the distance between the electrodes

Consequently, *the main contribution to the resistance of the medium comes from the regions adjoining the electrodes.* Hence, it is especially important to ensure their good conductivity. For this reason, electrodes with large surface area are used, which are buried deeply into the Earth, where the ground water ensures a good conductivity of the soil.

Experimental verification. Two plane electrodes, connected to the terminals of an extraneous e.m.f. source, are immersed into a weakly conducting liquid, say, river water (Fig. 124). A certain current flows through the circuit. By changing the separation between the electrodes, we see that the readings of the ammeter are not altered even at large distances (as compared to the size of the electrodes). Consequently, under these conditions the resistance of the medium is independent of the distance between the electrodes.

Step voltage. Since a current flows through the medium, an electric field exists as well as a potential which varies in space.

Suppose that there is a break in a high-voltage transmission line and the free end of a wire of length L is lying on the ground. An electric current flows through the regions of soil adjoining the conductor. If a man happens to be walking nearby, a potential difference called the step voltage appears between the points where his feet touch the ground. Consequently, an electric current whose strength depends on this potential difference flows through the man.

Let us calculate the step voltage. Since the conductor is quite long, we assume that the current flows from it to the ground in a direction perpendicular to the conductor. The equipotential surfaces are the surfaces of semicylinders whose axes coincide with the conductor (Fig. 124). Suppose that the man is walking in a direction perpendicular to the conductor with a step of length l, the distance between the conductor and the foot closer to it being d. Assuming that the current flows uniformly from the conductor over the semicylindrical region we obtain the following expression for the current density at a distance r from the

www.EngineeringBooksPDF.com

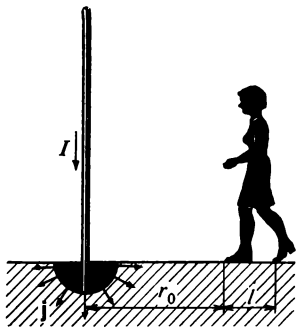


Fig. 125. To the calculation of step voltage when a hemispherical earth plate is approached

conductor:

$$j = I/(\pi r L). \tag{30.6}$$

In this case, the field strength along the radii perpendicular to the conductor is

$$E_r = j/\gamma = I/(\pi r L\gamma). \tag{30.7}$$

Consequently, the step voltage is

$$U_{st} = \int_d^{d+l} E_r \, dr = \frac{I}{\pi\gamma L} \ln \frac{d+l}{d}. \tag{30.8}$$

For example, if $I = 500$ A, $d = 1$ m, $l = 65$ cm and $L = 30$ m, we find that $U_{st} = 270$ V. Much higher voltages may appear under other conditions and other shapes of conductors. Hence, *when a part of a high-voltage transmission line falls on the ground, it creates a hazard not only because there can be a direct contact between the cable and a human being, but also because of the emergence of step voltages.*

The resistance does not depend on the distance between the electrodes since the effective cross section of the area through which the current flows is proportional to the distance between the electrodes.

Example 30.1. A hemispherical earth plate is buried into the earth in level with its surface (Fig. 125). Find the voltage which may be applied to a woman approaching this earth plate (step voltage). The current passing through the earth plate is equal to I, the length of a step is l, and the distance between the plate and the foot closer to it is r_0. Solve the numerical problem: $\gamma = 10^{-2}$ S/m, $I = 1$ A, $r_0 = 2$ m and $l = 1$ m.

The current from the earth plate is uniform in all directions and hence the current density vector is directed along the radius vector from the earth plate and is equal to

$$j_r = I/(2\pi r^2).$$

In accordance with Ohm's law, the electric field strength is equal to

$$E_r = j_r/r = I/(2\pi r^2\gamma).$$

Consequently, the step voltage is

$$U_{st} = \int_{r_0}^{r_0+l} E_r \, dr = \frac{I}{2\pi\gamma} \int_{r_0}^{r_0+l} \frac{dr}{r^2} = \frac{I}{2\pi\gamma} \left(\frac{1}{r_0} - \frac{1}{r_0+l} \right) = 2.7 \text{ V}.$$

Problems

4.1. A copper sphere of 10-cm diameter is lowered into a water-filled hemispherical copper vessel of 20-cm diameter so that the sphere and the vessel are concentric. The electrical conductivity of water is equal to $\gamma = 10^{-3}$ S/m. Find the electric resistance between the sphere and the vessel.

4.2. A small spherical electrode of radius a is immersed in a medium of conductivity γ at a distance d from another electrode in the shape of a large plate having a high conduc-

www.EngineeringBooksPDF.com

tivity. Find the resistance of the medium to the electric current passing between the electrodes.

4.3. Find the resistance of a medium to the current flowing between two concentric electrodes of radii r_1 and r_2 respectively. The conductivity of the medium is γ.

4.4. Find the resistance between points A and B of the network shown in Fig. 126. The resistance of the side of a small square is R.

4.5. Two plane electrodes of area S each, whose linear dimensions are much larger than the distance d between them, are separated by a conducting material whose conductivity

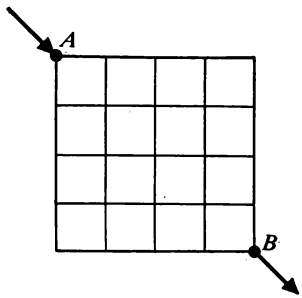


Fig. 126. To Problem 4.4

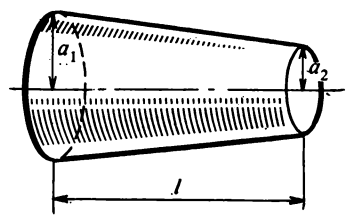


Fig. 127. To Problem 4.6

varies linearly from γ_1 at the surface of one electrode to γ_2 at the surface of the other electrode. Find the resistance of the medium between the electrodes.

4.6. Find the resistance of a conic conductor of a circular cross section, whose dimensions are shown in Fig. 127. The electrical conductivity of the conductor material is γ.

4.7. The space between two infinite plane-parallel electrodes separated by a distance d is filled with two layers of a substance with a plane interface parallel to the electrodes. The conductivities and the permittivities of the layers are γ_1, ε_1 and γ_2, ε_2 respectively, the layer thicknesses being a and $d - a$. The potentials φ_1 and φ_2 are applied to the electrodes. Find the potential and the surface charge density at the interface.

Answers

4.1. $R = 1590\ \Omega$. 4.2. $R = [1 - a/(2d)]/(4\pi\gamma a)$. 4.3. $R = \dfrac{1}{4\pi\gamma}\left(\dfrac{1}{r_1} - \dfrac{1}{r_2}\right)$. 4.4. $R_{AB} = \dfrac{47}{22} R$. 4.5. $R = \dfrac{d \ln (\gamma_2/\gamma_1)}{S(\gamma_2 - \gamma_1)}$. 4.6. $R = \dfrac{l}{\pi\gamma a_1 a_2}$. 4.7. $\varphi = \dfrac{\varphi_1\gamma_1(d-a) + \varphi_2\gamma_2 a}{\gamma_1(d-a) + \gamma_2 a}$; $\sigma = \dfrac{(\gamma_1\varepsilon_2 - \gamma_2\varepsilon_1)(\varphi_1 - \varphi_2)}{\gamma_1(d-a) + \gamma_2 a}$.

www.EngineeringBooksPDF.com

CHAPTER 5

Electrical Conductivity

The mechanisms of electrical conductivity are diverse. Their only common feature is the close relation with the motion of charges. The laws governing electrical conductivity vary over wide limits depending on the mechanism of electrical conductivity, properties of materials and conditions under which electric current flows in conductors.

Sec. 31. Electrical Conductivity of Metals

Main experimental facts associated with electrical conductivity of metals are described and their theoretical interpretaion is given.

The proof of the absence of mass transport by electric current in metals. Long before electrons were discovered, it was experimentally shown that, unlike in electrolytes, the passage of current in metals is not connected with the mass transport of metals. Experiments involved the flow of a direct current through a metal-to-metal contact, for example, between gold and silver, during a period of time reaching several months. After this, the material in the vicinity of contacts was investigated. It was demonstrated that no mass transport through the interface between two metals was observed, and the substances on both sides of the interface have the same composition as before the passage of current. This experiment proved that *atoms and molecules in metals do not participate in electric current* but they failed to give an answer to the question about the nature of charge carriers in metals.

The Tolman and Stewart experiments. These experiments, carried out in 1916, served as a direct proof that *current in metals is due to the motion of electrons.* The idea of these experiments was put forward by Mandelstam and Papaleksi in 1913.

Suppose that we have a conducting coil which can rotate about its axis. The ends of the coil are connected to a galvanometer with the help of sliding contacts (Fig. 128). If the rapidly rotating coil is abruptly stopped, free electrons in the wire continue to move by inertia, as a result of which the galvanometer should register a current pulse.

Let us denote by $\dot{v}$ the linear acceleration of the coil during braking. It is directed along the tangent to the surface of the coil. For a sufficiently dense winding and thin wires, we can assume that the acceleration is directed along the

www.EngineeringBooksPDF.com

wires. During the deceleration of the coil, the inertial force $m_e\dot{v}$ directed against the acceleration is applied to each free electron (m_e is the electron mass). Under the action of this force, an electron in metal behaves as if it were acted upon by a certain effective electric field

$$E_{\text{eff}} = -m_e\dot{v}/e. \tag{31.1}$$

Hence, the effective electromotive force in the coil due to inertia of free electrons is given by

$$\mathscr{E}_{\text{eff}} = \int_L E_{\text{eff}}\,dl = -\frac{m_e}{e}\dot{v}\int_L dl = -\frac{m_e}{e}\dot{v}L, \tag{31.2}$$

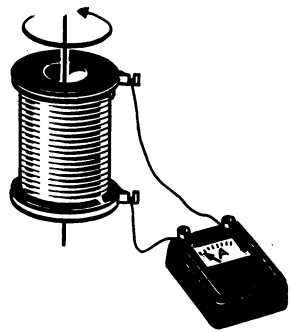


Fig. 128. Tolman and Stewart's experiment

where L is the length of the wire. All points of the wire have the same deceleration rate, and hence $\dot{v}$ in (31.2) is taken out of the integral.

Denoting the current flowing in the closed circuit by I and the resistance of the entire circuit including the resistance of the coil wires and the wires in the external circuit and in the galvanometer by R, we write Ohm's law in the form

$$IR = -m_e\dot{v}L/e. \tag{31.3}$$

The amount of electricity passing through the cross section of a conductor for a current I during time dt is

$$dQ = I\,dt = -\frac{m_e}{e}\frac{L}{R}\dot{v}\,dt = -\frac{m_e}{e}\frac{L}{R}\,dv. \tag{31.4}$$

Thus, the amount of electricity passing through the galvanometer during the braking time required to decrease the coil velocity from the initial linear velocity v_0 to zero is given by

$$Q = \int dQ = -\frac{m_e}{e}\frac{L}{R}\int_{v_0}^{0} dv = \frac{m_e}{e}\frac{L}{R}v_0. \tag{31.5}$$

The magnitude of Q is determined from the readings of the galvanometer, while the values of L, R, and v_0 are known. Hence we can find the sign and the magnitude of e/m_e. Experiments showed that e/m_e corresponds to the ratio of the electron charge to its mass. Thus, it was proved that the current observed with the help of a galvanometer is due to the motion of electrons.

On the band theory. The quantum theory of electrical conductivity is based on the band theory which follows from an analysis of the energy spectrum of electrons see Sec. 2). The electron spectrum is split into bands separated by forbidden gaps. If the upper band of a substance, which still contains electrons, has some free quantum states, i.e. if there is a possibility for rearranging the energy and momentum of electrons, this substance is a conductor. In this case, its up-

www.EngineeringBooksPDF.com

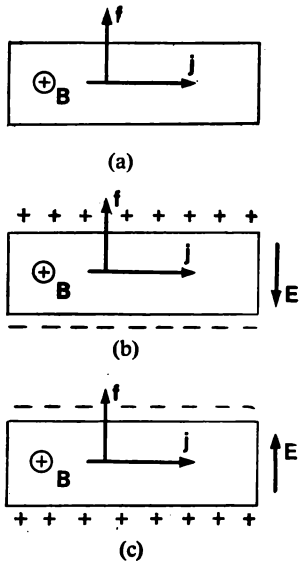


Fig. 129. The Hall effect

per band is called the conduction band, and the substance is called a conductor with electronic-type conductivity. If there are many electrons and free quantum states in the conduction band, the electrical conductivity is sufficiently high. The electrons in the conduction band are the only charge carriers responsible for electric current. Their motion obeys laws of quantum mechanics. The number of these electrons constitutes only a small part of the total number of electrons. This circumstance eliminates drawbacks of the classical theory of electrical conductivity (see Sec. 27).

Temperature dependence of resistance. Motion of electrons is the main cause of electrical conduction not only in metals. For example, in semiconductors with electronic-type conductivity the motion of electrons also contributes significantly to the transport of electric charge. One of the most typical differences in electrical conduction in these two cases is in the nature of temperature dependence of electrical conductivity.

Experiments show that *the resistivity of metallic conductors grows with temperature, i.e. their conductivity decreases.* For a moderate temperature, the temperature dependence of conductivity has the form $\gamma \propto 1/T$.

However, *for some materials* (e.g. *glasses, semiconductors, electrolytes*) *conductivity increases with temperature.* Although the mechanisms of increase in conductivity are different, they all ultimately boil down to a decrease in the number of electric charge carriers responsible for current. The number of carriers in metals, i.e. free electrons, practically does not depend on temperature. Hence, the resistance to the current is determined only by the ability of metals to form an ordered motion under the action of an electric field, viz. by their mobility, which decreases with increasing temperature.

Hall effect. The charges creating an electric current due to their motion are acted upon by Ampère's force (9.23). The density of this force can be written in the form

$$\mathbf{f} = \mathbf{j} \times \mathbf{B} = ne\mathbf{v}_d \times \mathbf{B}, \tag{31.6}$$

where e is the charge whose motion forms the current, and n and $\mathbf{v}_d$ are its concentration and drift velocity respectively.

In the presence of a magnetic field whose induction is perpendicular to the current density $\mathbf{j}$, the charges in the conductor tend to move in the direction of the force with the density $\mathbf{f}$ (Fig. 129*a*). As a result, an excess charge of the same type that generates the current is formed on the corresponding surface of the conductor. Thus, if the current is due to the motion of positive charges, the distribution of the surface charge density will be as shown in Fig. 129*b*, while

www.EngineeringBooksPDF.com

for the motion of negative charges we obtain the distribution depicted in Fig. 129*c*. A potential difference and an electric field E appear between the opposite faces of the conductor. This field neutralizes the action of the forces with density **f** (31.6). The direction of this field depends on the sign of charges forming current, and its magnitude is determined by the factors responsible for the density of force (31.6). The emergence of a potential difference in a current-carrying conductor placed in a magnetic field is called the **Hall effect**. This effect was discovered in 1879.

The induction **B** of the magnetic field and the velocity $\mathbf{v}_d$ of charges are at right angles. The ratio of the force density (31.6) to the charge, like (31.1), can be treated as the effective electric field called the **Hall field:**

$$E_{\text{eff}} = v_d B. \tag{31.7}$$

Consequently, the potential difference emerging between the faces of the conductor (Fig. 129*b*) is given by:

$$U = \int_0^d v_d B \, dx = v_d B d, \tag{31.8}$$

where d is the thickness of the conductor. Considering that $j = nev_d$, we can write this expression in the form

$$U = djB/(ne) = RjBd, \tag{31.9}$$

where

$$R = 1/(ne) \tag{31.10}$$

is the **Hall constant.** The potential difference can be measured. The other quantities, except the concentration n of charges and their sign, are known. *The sign of the potential difference can be used to determine the sign of the charge carriers whose motion creates current, while its magnitude determines the carrier concentration.*

It should be noted that formulas (31.9) and (31.10) coincide with the corresponding formulas in a more complete theory of the Hall effect, where the velocity distribution of electrons, statistical nature of their collisions and other factors are taken into account. In this case, however, calculations are very cumbersome and we shall not consider them in this book.

The results of measurements showed that *current in metals is formed by the motion of negative charges.* The carrier concentration is approximately equal to the atomic concentration. In other words, one charge participating in a current corresponds to about one atom of the metal, although this number varies within certain limits. *Electrons are the charge carriers responsible for current in metals.* This means that in metals, on the average, one free electron corresponds to an atom. For example, 0.7 electron corresponds to a silver atom, 0.8 to copper, 0.9 to gold, and about two electrons to aluminium. It should be recalled that the atomic concentration in metals, and hence the concentration of free electrons, is close to $n \simeq 10^{28}$ m^{-3}.

www.EngineeringBooksPDF.com

The analysis of the Hall effect in other cases revealed that it is not always due to the motion of negative charges. *When the sign of the potential difference in the Hall effect corresponds to the motion of negative charges, the effect is called anomalous.*

The Hall effect is one of the galvanomagnetic phenomena. This term refers to the phenomena observed in a current-carrying conductor placed in a magnetic field. The physical essence of all these phenomena consists in that the electrical conductivity of a conductor in an external magnetic field is a tensor rather than a scalar. The transverse electric field, called the Hall field, is added to the electric field which creates a current in the absence of a magnetic field. As a result, the direction of the resultant electric field forms with the current density a certain angle called the **Hall angle**. This means that the directions of the current density and the electric field do not coincide. These quantities are related through the tensor formula

$$j_i = \sum_k \gamma_{ik} E_k,$$

where γ_{ik} is the electrical conductivity tensor. The conductivity of anisotropic materials is described by the electric conductivity tensor even in the absence of a magnetic field.

Magnetoresistance. Another important galvanomagnetic effect is a *change in the resistance of a conductor placed in a transverse magnetic field (magnetoresistance)*. Experiments show that the relative change in the electrical conductivity $\Delta\gamma/\gamma$ for not very strong fields is expressed by the formula

$$\Delta\gamma/\gamma = -\varkappa_\perp B^2,$$

where $\varkappa_\perp$ is the transverse magnetoresistance coefficient which depends on the properties of the material and B is the magnetic induction.

This phenomenon is a consequence of the tensor nature of electrical conductivity of a conductor placed in a magnetic field. As a result, the electric field component collinear with the current appears, which causes a change in the current manifested as a change in the resistance.

Mobility of electrons. Ohm's law $\mathbf{j} = \gamma\mathbf{E}$ can be written in the form

$$nev_d = \gamma E. \tag{31.11}$$

The mobility b of electrons is defined as the ratio of their drift velocity to the electric field strength:

$$b = v_d/E. \tag{31.12}$$

Taking into account (31.11), we obtain

$$b = \gamma/(ne). \tag{31.13}$$

The electrical conductivity of a metal is known, and ne can be found from the Hall effect. In other words, the change in the Hall effect makes it possible to find the electron mobility in the conductor. The electron mobility in metals is of the order

$$b \sim 10^{-4}\text{-}10^{-3}\ \text{m}^2/(\text{V}\cdot\text{s}). \tag{31.14}$$

www.EngineeringBooksPDF.com

Thus, the drift velocity of electrons in metals is very small as compared to ordinary velocities of motion of microparticles. *A high conductivity of metals is mainly due to a high carrier concentration* ($n \simeq 10^{28}$ m^{-3}) *and not due to their high mobility* [see (31.13)]:

$$\gamma = enb \sim 10^{-19} \cdot 10^{28} \cdot 10^{-3} \text{ S/m} = 10^{6} \text{ S/m}.$$

In dielectrics, most electrons are rigidly connected to the atoms, and the number of free charge carriers is very small. Consequently, *the conductivity of dielectrics is very low although the mobility of charge carriers in them does not differ drastically from the mobility of free electrons in metals. The carrier concentration in semiconductors varies between* 10^{19} *and* 10^{25} m^{-3}, *while their mobilities lie between* 10 *and* 10^{-4} m^2/(V·s), *i.e. are high.* Such a wide range of variation of the concentration and mobility of carriers determines the wide range (over several orders of magnitude) of variation of electrical conductivity of semiconductors. However, it is impossible to attain as high conductivity for semiconductors as for metals, having retained, of course, the temperature dependence of electrical conductivity typical of semiconductors (i.e. an increase in electrical conductivity with temperature).

Superconductivity. In 1911 H. Kamerlingh Onnes discovered that apparently mercury completely loses its resistance to electric current at 4.2 K. The loss of resistance occurs abruptly within an interval of a few hundredths of a degree. The disappearance of resistance was subsequently observed for many other pure materials and alloys. This phenomenon was called superconductivity. Transition temperatures for the superconducting state are different but always very low.

Critical temperature. If an electric current is excited in a superconducting ring with the help of electromagnetic induction the magnitude of this current remains the same for several years. This allows us to determine the upper limit of resistivity of superconductors (which is below 10^{-25} Ω·m). This value is less than the resistivity of copper at low temperature (equal to 10^{-12} Ω·m) by many orders of magnitude. Therefore, *it is assumed that the electric resistance of superconductors is equal to zero.* Before the transition to superconducting state, the resistance may have different values. Many superconducting materials have a rather high resistance at room temperature. The transition to superconducting state always occurs abruptly. For pure single crystals, this temperature interval is less than 10^{-3} degree.

Among pure materials, aluminium, zinc, indium and gallium exhibit superconducting properties. These properties depend on the crystal lattice structure. For example, white tin is a superconductor while grey tin is not. Mercury is a superconductor only in the α-phase.

Critical field. In 1914, Kamerlingh Onnes found that the *superconducting state is destroyed by a magnetic field when the magnetic induction B exceeds a certain critical value.* The critical value of induction depends on the superconductor material and temperature.

The critical field that destroys superconductivity may be created by the supercurrent itself. Therefore, there exists a critical current at which superconductivity vanishes.

www.EngineeringBooksPDF.com

Meissner effect. In 1933, Meissner discovered that *there is no magnetic field inside a superconductor*. When a superconductor placed in a constant external magnetic field is cooled, the magnetic field is completely expelled from its volume at the moment of the transition to the superconducting state. This is the principal difference between a superconductor and an ideal conductor in which the magnetic induction in the volume remains unchanged when its resistivity drops to zero. The property of expulsion of magnetic field from the volume of a superconductor is called the **Meissner effect.** This effect and the absence of electrical resistance are the most important properties of a superconductor.

Surface current. Proceeding from the general laws of magnetic fields (see Chap. 6) and taking into account the absence of a magnetic field in the bulk of a superconductor, we may conclude that *only surface current exists in it. From the physical point of view, it is a real current flowing through a certain thin layer near the surface.* The magnetic field of the current neutralizes the external magnetic field in the superconductor. In this respect, a superconductor behaves formally as an ideal diamagnetic (see Sec. 41). However, it is not a diamagnetic since magnetization inside it is equal to zero.

Soft and hard superconductors. The number of pure materials exhibiting superconducting properties is not large. Most frequently superconductivity is observed in alloys. The Meissner effect is observed to the fullest extent in pure materials, while *in alloys, the magnetic field is not expelled completely from their volume* (partial Meissner effect). *Materials exhibiting complete Meissner effect are called soft superconductors, while those in which the effect is partial are called hard superconductors.*

In the bulk of hard superconductors, circular currents create a magnetic field which, however, does not fill the entire volume of the conductor but is distributed in it in the form of individual filaments. As to the resistance, it is equal to zero in hard as well as soft superconductors.

The theory of superconductivity. In its physical nature, superconductivity is the superfluidity of a liquid consisting of electrons. Superfluidity sets in when the energy exchange between the superfluid component of a liquid and its other parts ceases, and consequently the friction vanishes. An important feature of this process is the possibility of "condensation" of liquid molecules on the lower energy level separated from other levels by a sufficiently large energy gap which cannot be surmounted by the forces of interaction. This is the reason behind the termination of interaction. For the accumulation of many particles on the lower energy level to become possible, it is necesary that they obey the Bose-Einstein statistics, i.e. have an integral spin.

Electrons obey the Fermi-Dirac statistics and hence cannot be "accumulated" on the lower energy level to form a superfluid electron liquid. The repulsive forces between electrons are compensated to a considerable extent by the forces of attraction exerted by the positive ions of the crystal lattice. However, an attractive force may appear between electrons due to thermal fluctuations at the lattice sites, and then the electrons are combined into pairs. *These ele-*

www.EngineeringBooksPDF.com

ctron pairs behave as particles with an integral spin, i.e. obey the Bose-Einstein statistics. They may condense and form a current of superfluid liquid, viz. electron pairs which form a supercurrent. Above the lower energy level there is an energy gap which cannot be surmounted by an electron pair at the expense of the energy of interaction with the remaining charges. In other words, the electron pair cannot change its energy state and hence there is no electrical resistance. The possibility of the formation of electron pairs and their superfluidity is explained by the quantum theory.

The large difference in the electrical conductivities of conductors, semiconductors and dielectrics is due to a large difference in the carrier concentrations rather than to a difference in the mobility of charge carriers.

Example 31.1. *The temperature dependence of resistance is quite essential for the operation of many devices, as can be easily seen from the operation of an ordinary incandescent lamp. The filament of a lamp is made of tungsten, whose conductivity and radiant emittance M (the surface density of the radiant flux from the surface) can be represented by the following formulas in the temperature range between* 300 *and* 3000 K: $\gamma = 0.95 \times 10^{10}\ T^{-1.2}$ S/m, $M = 6.6 \times 10^{-12}\ T^{5}$ W/m², *where T is the thermodynamic temperature. Calculate the diameter d and the length l of a filament if the lamp emits a power P at a voltage U and filament temperature T. Assume that the energy losses due to the thermal conductivity of the filament are negligibly small. Estimate the precision requirements for manufacturing the filament.*

We have

$$R=\frac{U^2}{P},\quad R=\frac{1}{\gamma}\,\frac{4l}{\pi d^2}\quad P=\pi M l d,$$

whence

$$d=\left(\frac{4P^2}{\pi^2\gamma U^2 M}\right)^{1/3}\qquad l=\left(\frac{\gamma P U^2}{4\pi M^2}\right)^{1/3}.$$

Since $\gamma M \propto T^{3/8}$ and $\gamma/M^2 \propto T^{-11.2}$, the temperature dependence of the length and thickness of the filament is rather strong. Therefore, the error in the diameter and length of the filament during manufacturing considerably affects the temperature, and hence the spectral composition of the emitted light. Consequently, the precision requirements are quite stringent.

Sec. 32. Electrical Conductivity of Liquids

The mechanism of electrical conductivity in liquids is described and the dependence of electrical conductivity on various factors is considered.

Dissociation. Pure liquids are basically poor conductors of electricity. This is due to the fact that they consist of neutral atoms and molecules whose motion cannot generate an electric current. However, the solutions of salts, acids and alkalis are good conductors of electricity. It can be explained as follows. The molecules of a dissolved substance dissociate, i.e. are decomposed into positive and negative ions. The ordered motion of ions ensures the transport of

www.EngineeringBooksPDF.com

electric charges, viz. an electric current. *If the process of dissolution does not involve a molecular dissociation, the solution is not a conductor.*

Calculation of electrical conductivity. We denote by $N = N^{(+)} = N^{(-)}$ the concentration of ions of each sign in a solution. For current density, we can write

$$j = q\,(b^{(+)} + b^{(-)})\,NE, \tag{32.1}$$

where q is the value of the ion charge and $b^{(+)}$ and $b^{(-)}$ are the mobilities of positive and negative ions (see (31.12)).

On the basis of (31.12), the drift velocity of ions is proportional to the field strength:

$$v_d^{(\pm)} = b^{(\pm)}E. \tag{32.2}$$

Generally speaking, the positive and negative ions have different mobilities. The mobility of positive ions in liquids is small and usually amounts to about 10^{-7} m^2/(V·s).

The ion concentration depends on the degree of dissociation which is characterized by the dissociation coefficient α defined by the ratio of the ion concentration N to the concentration N_0 of the solute molecules:

$$N = \alpha N_0. \tag{32.3}$$

Consequently, the concentration of undissociated molecules is

$$N' = (1 - \alpha)\,N_0. \tag{32.4}$$

Dissociation and solvation, i.e. the combination of ions into neutral molecules, occur simultaneously and continuously in a solution. At equilibrium, the intensities of these two processes that change the solution composition in opposite directions are equal. The rate of variation of the concentration (dN/dt) of each type of ions as a result of molecular dissociation is proportional to the concentration N' of undissociated molecules:

$$(dN/dt) = \beta\,(1 - \alpha)\,N_0, \tag{32.5}$$

where β is the proportionality factor.

The rate dN/dt of variation of concentration of undissociated molecules as a result of ionization of molecules is proportional to the product of concentrations of positive and negative ions:

$$(dN'/dt) = \eta\alpha^2 N^2{}_0, \tag{32.6}$$

where η is the proportionality factor. At equilibrium,

$$\left(\frac{dN}{dt}\right) = \left(\frac{dN'}{dt}\right). \tag{32.7}$$

Taking into account (32.5) and (32.6), we obtain a formula connecting the dissociation coefficient with the concentration of the solute:

$$\frac{1-\alpha}{\alpha^2} = \frac{\eta}{\beta} N_0. \tag{32.8}$$

www.EngineeringBooksPDF.com

Obviously, the dissociation coefficient depends on the concentration of the solute. For a very low concentration ($N_0 \simeq 0$), formula (32.8) gives

$$\alpha = 1, \tag{32.9}$$

i.e. the dissociation is almost complete. If $\alpha \ll 1$, we obtain from (32 8)

$$\alpha = \sqrt{\frac{\beta}{\gamma}} \frac{1}{\sqrt{N_0}}. \tag{32.10}$$

In other words, α decreases with increasing concentration of the solute.

Taking into consideration Eq. (32.3), we can write (32.1) in the form

$$j = q\,(b^{(+)} + b^{(-)})\alpha N_0 E. \tag{32.11}$$

The mobility of ions does not depend on the electric field over a wide range of the field strength. The *deviation from the linear dependence of the field strength on the drift velocity of carriers is observed only for a very large strength of the order of millions of volts per centimeter*, at which, in accordance with (32.2), the mobility depends on the field strength. The value of α is also independent of E over a very wide range. Consequently, formula (32.11) expresses Ohm's law for fields up to $\simeq 10^6$ V/cm. Thus, the electrical conductivity of a solution is given by

$$\gamma = q\,(b^{(+)} + b^{(-)})\ \alpha N_0. \tag{32.12}$$

Dependence of electrical conductivity on concentration. For a not very high concentration of a solution, the dissociation coefficient is constant. The sum of the mobilities $b^{(+)}$ and $b^{(-)}$ of ions is also nearly constant. Consequently, *for a small concentration of a solution, its electrical conductivity is proportional to the concentration. For large concentrations, the situation becomes much more complicated.* On the one hand, we must take into account the dependence of the dissociation coefficient on concentration [see (32.8) and (32.10)], while on the other hand, the ion mobility also begins to depend noticeably on concentration. In concentrated solutions, the mobility of ions decreases as the electric interaction between ions comes into play. Therefore, *at high solution concentrations, a linear dependence of the electrical conductivity on solution concentration is not observed.*

Temperature dependence of electrical conductivity. As the temperature rises, the dissociation coefficient increases since a more rapid motion of molecules hampers solvation and facilitates dissociation (upon collisions). The viscosity of liquids decreases with heating, and hence the mobility of ions becomes higher. Consequently [see (32.12)], *the conductivity of electrolytes increases with temperature, and may assume quite large values (exceeding the initial values by several orders of magnitude).*

Electrolytes. Since the current through solutions is due to the motion of ions, *the molecules of a dissolved substance are decomposed into components which are liberated at the electrodes. This phenomenon is called electrolysis.* The study of electrolysis has played a significant role in the development of the theory of the structure of matter. The laws of electrolysis discovered by M. Faraday are studied in detail in the course of secondary school physics. The conductors which undergo electrolysis, i.e. are dissociated into ions, upon the passage of electric

www.EngineeringBooksPDF.com

current through them are called electrolytes. Hence it follows that *electrolytes include many solutions of salts, acids and alkalis as well as a number of chemical compounds in liquid and solid states.*

An example of a solid electrolyte is glass which in its physical nature is a supercooled liquid with a very high viscosity. It can be shown experimentally that the Na^+ ions which are responsible for electrical conductivity of glass have a noticeable mobility in it. When glass is heated, its resistance may decrease to several millionths of its initial value. This can be illustrated by a very impressive experiment. A glass rod connected to an electric power source is first heated by the flame of a burner. The Joule heat liberated in the circuit contributes to the heating of the rod. At a certain temperature (which is selected experimentally) the burner is removed, and the further increase in the temperature of the rod is only due to ohmic heating. The rate of change of temperature of the rod is constantly increasing since the conductivity of glass increases with temperature, which, in turn, causes an even sharper increase in the temperature. As a result of such an avalanche increase in temperature, the glass rod vigorously melts and burns with a bright flash.

Sec. 33. Electrical Conductivity of Gases

Various mechanisms of conduction of current in gases are discussed. The characteristic of current and the role of volume charge are outlined

Self-sustained and non-self-sustained currents. A gas containing no charged particles cannot conduct electricity. It becomes a conductor only upon being ionized, when charge carriers appear in the form of free electrons and ions. Positive ions may be singly or multiply charged depending on the number of lost electrons. Negative ions formed as a result of addition of an electron to an atom are usually singly charged.

In order to make a gas conduct, some external ionization factors (a high temperature of the gas, ultraviolet and X-ray radiation, etc.) are required. If the field strength is not high, the current through the gas ceases as soon as the extrinsic ionization factor stops to operate. Such a current is called non-self-sustained.

If the field strength is sufficiently high, the field itself may cause ionization as a result of which the gas becomes a conductor. The current appearing in this case is called self-sustained. *There is no unique universal dependence of current on voltage for self-sustained currents. The situation is determined by specific conditions.* In particular, it may happen that a self-sustained current decreases with increasing voltage.

Non-self-sustained current. Let us consider a non-self-sustained current in greater detail. We denote by N the concentration of charges of each sign and by $(dN/dt)_{cr}$ the rate of variation of the charge concentration due to an external

www.EngineeringBooksPDF.com

source of ionization. The creation of charges is accompanied by their annihilation as a result of recombination, i.e. mutual neutralization. After a sufficiently long time, a dynamic equilibrium sets in, when the rate of charge formation and the rate of recombination become equal. Obviously, in this case

$$N = N^{(+)} = N^{(-)}, \tag{33.1}$$

where we assume for the sake of simplicity that ions are singly charged.

Clearly, the recombination rate must be proportional to the product of charge concentration, i.e. N^2. Hence, at equilibrium we have

$$(\mathrm{d}N/\mathrm{d}t)_{\mathrm{cr}} = -rN^2, \tag{33.2}$$

where r is the recombination coefficient.

By definition, the current density is given by

$$j = j^{(+)} + j^{(-)} = q\,(N^{(+)}v_{\mathrm{d}}^{(+)} + N^{(-)}v_{\mathrm{d}}^{(-)}) = qN\,(v_{\mathrm{d}}^{(+)} + v_{\mathrm{d}}^{(-)}). \tag{33.3}$$

The drift velocity of charge in an electric field is proportional to the field strength:

$$v_{\mathrm{d}} = bE. \tag{33.4}$$

The mobilities $b^{(+)}$ and $b^{(-)}$ of positive and negative charges are generally different. Taking into account (33.4), we can write Eq. (33.2) as

$$j = (b^{(+)} + b^{(-)})\,NE. \tag{33.5}$$

This formula resembles Ohm's law. However, it is equivalent to Ohm's law only when the coefficient of E does not depend on E and j. *Generally speaking, this coefficient for gases depends on the indicated quantities, and hence formula* (33.5) *is not equivalent to Ohm's law.*

If the number of ions recombining in 1 s is much larger than the number of atoms reaching an electrode, we can use expression (33.2) under equilibrium conditions for determining N in (33.5). This gives

$$j = q\,(b^{(+)} + b^{(-)})\,\sqrt{\frac{1}{r}\left|\left(\frac{\mathrm{d}N}{\mathrm{d}t}\right)\right|_{\mathrm{cr}}}\,E \tag{33.6}$$

In order to find the conditions of applicability of this formula, we must bear in mind that the mobility of ions in gases is of the order of 10^{-4} m^2/V·s, while the recombination coefficient $r \simeq 1$ m^3/s. For example, if $\mathrm{d}N/\mathrm{d}t$ is of the order of 10^{16} ions/(m^3·s) and $E = 10^3$ V/m, the number of ions arriving at 1 m^2 of an electrode during 1 s is given by

$$\frac{j}{e} = (b^{(+)} + b^{(-)})\,\sqrt{\frac{1}{r}\left|\left(\frac{\mathrm{d}N}{\mathrm{d}t}\right)\right|_{\mathrm{cr}}}\,E \approx 2\cdot 10^{13}\ \mathrm{m^{-2}\cdot s^{-1}} \tag{33.7}$$

If the distance between plane electrodes is 0.1 m, the number of ions recombining in the space between the electrodes per 1 m^2 of the cross-sectional area is equal to 10^{15}, i.e. the condition of applicability of formula (33.6) is satisfied in this case. The applicability of this formula for other values of parameters is verified in a similar way.

www.EngineeringBooksPDF.com

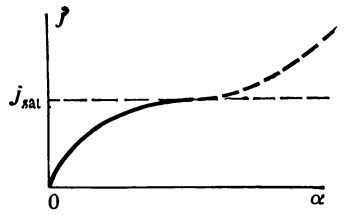


Fig. 130. Characteristics of self-sustained and non-self-sustained currents

Saturation current density. We denote by d the distance between two plane electrodes. If the field strength is sufficiently high, so that all the ions formed by an external source reach the electrodes before they recombine, a saturation current appears with density given by

$$j_{sat} = qd \left(\frac{dN}{dt}\right)_{cr} \tag{33.8}$$

The characteristic of current. In the region of intermediate electric fields, a part of ions has time to recombine before they reach the electrodes. The equation for the balance of loss and creation of ions is written in the form

$$\left(\frac{dN}{dt}\right)_{cr} + \left(\frac{dN}{dt}\right)_{rec} + \left(\frac{dN}{dt}\right)_{j} = 0. \tag{33.9}$$

Taking into account Eqs. (33.2), (33.3), and (33.8), we get

$$j_{sat}/q - rN^2d - N(b^{(+)} + b^{(-)})E = 0. \tag{33.10}$$

Considering that

$$j = qN(b^{(+)} + b^{(-)})E, \tag{33.11}$$

we can write (33.10) in the form of the following equation in j:

$$j^2 + 2\alpha j + 2\alpha j_{sat} = 0, \tag{33.12}$$

where

$$\alpha = |q|(b^{(+)} + b^{(-)})^2E^2/(2rd). \tag{33.13}$$

The positive root of Eq. (33.12) is

$$j = \alpha\left(\sqrt{1 + 2j_{sat}/\alpha} - 1\right). \tag{33.14}$$

The dependence of the current density on α is shown in Fig. 130. In the limiting cases ($\alpha \ll j_{sat}$ and $\alpha \gg j_{sat}$), Eq. (33.14) is transformed into (33.6) and (33.8) respectively.

Expression (33.14) is called the characteristic of non-self-sustained current. It is in good agreement with experiments if ion losses as a result of diffusion are additionally taken into consideration.

Self-sustained current. If we continue to increase the electric field strength at a current density nearly equal to the saturation value, the current density again starts increasing. This is so because before recombining with the ions, the electrons existing in the gas have time to get accelerated by the field up to energies at which they ionize the gas molecules by collisions. Consequently, the ionization rate becomes dependent on the field. The current appearing in this case is called self-sustained current. The initial part of the characteristics for this current is shown in Fig. 130 by the dashed line. It starts at a finite value of α.

The effect of volume charge. As was mentioned above, the mobilities of positive and negative charges are different and $b^{(-)}$ is usually greater than $b^{(+)}$. Con-

www.EngineeringBooksPDF.com

sequently, the density of current resulting from the motion of positive charges is less than that due to the motion of negative charges. Hence, the number of positive charges reaching the cathode during a fixed interval of time is less than the number of negative charges reaching the anode although the number of ions formed is the same as the number of recombining ions during this interval of time. Obviously, such a state cannot be an equilibrium state. The equilibrium state is attained in the following way. As a result of the motion of positive charges towards the cathode and negative charges towards the anode, an excess positive charge is accumulated at the cathode while negative charges are accumulated at the anode. However, in view of a higher mobility of negative charges, the excess negative charge at the anode will be larger than the excess positive charge at the cathode. As a result of such a redistribution of charge concentration and the change in the electric field associated with it, an equilibrium is established, at which the numbers of positive and negative charges reaching the electrodes become equal.

Mobility of charges. An ion having a mass m and charge q moves in a uniform magnetic field E with a constant acceleration

$$a = qE/m \tag{33.15}$$

and passes a distance

$$s = qE\tau^2/(2m), \tag{33.16}$$

during the time τ, the initial velocity being equal to zero.

If l is the mean free path of the ion in a gas for a random thermal motion and v is its mean velocity, we can assume that $\tau = l/v$. The time and the mean free path are defined in such a way that we can assume that during each collision the ion completely loses its energy of ordered motion. Consequently, on the basis of (33.16), we can express the drift velocity as the mean velocity of ordered motion in a direction collinear with the direction of the field:

$$v_d = s/\tau = qE\tau/(2m) = qlE/(2mv). \tag{33.17}$$

The modifications introduced by the statistical distribution of l lead just to a small change in the numerical factor in (33.17). Hence, the mobility of ions is given by

$$b = ql/(2mv). \tag{33.18}$$

This formula shows that the mobilities of positive and negative ions of the same mass must be equal. *However, the mean mobility of negative charges is higher than that of positive charges* since the mobility of negative charges is determined not only by the contribution from negative ions but from electrons as well. The mobility of electrons is significant due to their small mass, and this ultimately determines the high mobility of negative charges.

Comparison of results with experiment. It follows from (33.18) that the mobility is inversely proportional to the gas density since the mean free path is inversely proportional to density. This conclusion is confirmed by experiments.

However, formula (33.18) on the whole does not explain the entire body of experimental facts. In particular, the experimental values for the mobility are

www.EngineeringBooksPDF.com

lower than theoretical values. In order to explain the discrepancy between the theory and experiments Langevin took into account the polarization of ions approaching each other in collisions, owing to which the ions acquire additional dipole moments and the nature of their collisions changes. This circumstance introduces considerable corrections into the formulas. In the framework of this book, however, we shall not discuss this theory.

The presence of some external ionizing factor (high temperature of the gas, ultraviolet radiatoin, X-rays, etc.) is essential to make a gas conduct. For quite strong fields, however, the gas is ionized by the field itself. The current generated in this way is called self-sustained current. If external ionizing factors are present, the current is called non-self-sustained current.

What is a self-sustained and a non-self-sustained current?
Why is a volume charge induced between two electrodes?
What is its effect? Due to what factors is the mobility of negative charges higher than that of positive charges?

Sec. 34. Electric Current in Vacuum

Basic regularities of thermionic emission and their manifestation during the passage of current between electrodes in vacuum are discussed.

Thermionic emission. Electric current cannot exist in vacuum if the latter contains no charge carriers. If, however, electrons exist in vacuum, their motion generates an electric current which is called vacuum current.

Metals contain electron gas. In thermodynamic equilibrium, the distribution of electrons over energy levels is described by the Fermi-Dirac statistics and is given by

$$\frac{n_i}{g_i} = \frac{1}{\exp[\beta(E_i - \mu)] + 1}, \tag{34.1}$$

where $\beta = 1/(kT)$, n_i is the number of electrons having an energy E_i, g_i is the number of quantum states corresponding to this energy, and μ is the Fermi energy at temperature T, which tends to the Fermi energy μ_0 at $T = 0$ as $T \to 0$ K, in accordance with the formula

$$\mu = \mu_0 \left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\mu_0}\right)^2 + \ldots\right]. \tag{34.2}$$

Considering that in all cases of practical interest $\mu \gg kT$, we may assume that the quantity μ in (34.1) is equal to μ_0.

Let E_0 be the energy of an electron at rest near the outer surface of a metal (Fig. 131). Substituting E_0 for E_i in formula (34.1), we can calculate the probability that the electron has the energy E_0. This probability differs from zero,

www.EngineeringBooksPDF.com

the more the higher the temperature (i.e. the smaller the value of β). Thus, *near the surface of the metal there is an electron cloud which is in equilibrium with the electron gas in the metal.* This equilibrium is dynamic: the electrons in the metal, which have a sufficiently high kinetic energy, overcome the forces which confine them within the metal and become free. On the other hand, the electrons which are near the metal surface and have appropriate positions and directions of motion are captured by the forces which confine them to the metal. Thus, in dynamic equilibrium equal and opposite currents flow across the surface of the metal. The total current across the surface is equal to zero. *The formation of an electron cloud near the surface of a metal due to the thermal motion of free electrons is called thermionic emission.* At 0 K, no thermionic emission is observed, i.e. the electron cloud near the surface of a metal does not exist.

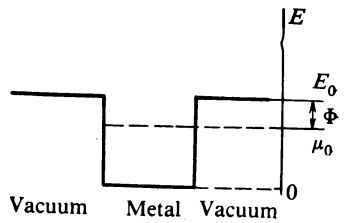


Fig. 131. Energy levels of free electrons in a metal

The total energy of electrons having a kinetic energy W_k near the surface of a metal is $E_i = W_k + E_0$, and formula (34.1) assumes the following form:

$$\left.\frac{n}{g}\right|_{W_k} = \frac{1}{\exp[\beta(W_k+\Phi)]+1}, \tag{34.3}$$

where $\Phi = E_0 - \mu$ is the work function of electrons. It follows from this formula that the electron cloud density near the surface of a metal strongly depends on the work function Φ and sharply decreases when the work function increases.

If an electric field exists near the surface of a metal, the electrons of the cloud start moving and an electric current called the thermionic current is generated. Thus, if two metallic plates to which a potential difference is applied are in vacuum, a thermionic current flows between them. Thus *current should obviously increase with increasing potential difference. There exists a maximum current at which all the electrons which get into the electron cloud across the surface of the cathode are entrained by the electric field towards the anode, and no reverse electron current flows into the cathode through its surface.* This maximum current is called the **saturation current.** Any further increase in the potential difference between the cathode and the anode does not alter the current since all the electrons supplied by the cathode as a result of thermionic emission are involved in the generation of electric current, and there are no other charge carriers for a further increase in the current.

For metals, Φ amounts to several electron-volts. The energy kT is equal to a fraction of an electron-volt even at a temperature of several thousand kelvins. Consequently, $\beta\Phi \gg 1$ and $\exp[\beta(W_k + \Phi)] \gg 1$. Hence we can neglect unity in the denominator of (34.3) in comparison with $\exp[\beta(W_k + \Phi)]$ and write this formula in the form

$$\left.\frac{n}{g}\right|_{W_k} \approx e^{-\Phi/(kT)} e^{-W_k/(kT)}. \tag{34.4}$$

www.EngineeringBooksPDF.com

Thus, the saturation current strongly depends on the work function and temperature since these quantities appear in the exponent. For pure metals, a significant current can be obtained only at a temperature of the order of 2000 K, i.e. *metals with a high melting point should be used for manufacturing cathodes. On the other hand, it is desirable that their work function be as small as possible.* For example, pure tungsten whose work function is 4.5 eV should operate at a temperature of 2500 K. For reducing the operating temperature of the cathode and its work function, oxide-coated cathodes are used, in which a layer of oxides of alkali-earth metals (e.g. BaO, SrO) is deposited on the substrate (base) with the help of special technological processes. The cathode is then activated by passing a thermionic current through it at a temperature of about 1300 K. As a result, the monoatomic layer of the alkali-earth metal is formed, which considerably reduces the work function. For example, barium-strontium oxide-coated cathodes have a work function of about 1.8 eV, due to which considerable currents can be obtained even at temperatures of about 1100 K. The current density attained at such a temperature is of the order of $10^4\ \mathrm{A \cdot m^{-2}}$. The barium-strontium oxide layer is usually deposited on a nickel tube with a tungsten filament inside it, used as a heater. Such a construction has an additional advantage over a heated tungsten filament used as the cathode, since in the latter case a considerable potential drop appears across the filament, and its surface will not be an equipotential surface. The oxide layer in a coated cathode is an equipotential surface, which considerably improves the operating conditions of the cathode as a whole.

The characteristics of an electron cloud. The electron cloud near the surface of a metal is described by formula (34.4). The number of quantum states in the element $dx\ dy\ dz\ dp_x\ dp_y\ dp_z$ of the phase volume is

$$g = \frac{2}{(2\pi\hbar)^3}\,dx\,dy\,dz\,dp_x\,dp_y\,dp_z. \tag{34.5}$$

Hence, the number of electrons in the element $dx\ dy\ dz\ dp_x\ dp_y\ dp_z$ of the phase volume is represented in the form

$$dn = \frac{2}{(2\pi\hbar)^3}\,e^{-\Phi/(kT)}e^{-p^2/(2m_e kT)}\,dx\,dy\,dz\,dp_x\,dp_y\,dp_z, \tag{34.6}$$

where $W_k = p^2/(2m_e)$.

The integration of (34.6) over $dx\ dy\ dz$ yields volume V as a factor. Consequently, the number of electrons in volume V, whose momenta are confined in the element $dp_x\ dp_y\ dp_z$ near the momentum p_x, p_y, p_z, is

$$dn_p = [2V/(2\pi\hbar)^3] \exp[-\Phi/(kT)] \exp[-p^2/(2m_e kT)]\,dp_x\,dp_y\,dp_z \tag{34.7}$$

where $p^2 = p_x^2 + p_y^2 + p_z^2$. Hence, we obtain the following expression for the concentration of the electron cloud near the surface of a metal:

$$n_0' = \frac{1}{V}\int dn_p = \left[\frac{1}{(2\pi\hbar)^3}\right] \exp\left(-\frac{\Phi}{kT}\right) \int_{-\infty}^{\infty}\int\int \exp\left(-\frac{p^2}{2m_e kT}\right) dp_x\,dp_y\,dp_z$$

$$= \frac{1}{4}\left(\frac{2\pi m_e kT}{\hbar^2}\right)^{3/2} \exp\left(-\frac{\Phi}{kT}\right). \tag{34.8}$$

www.EngineeringBooksPDF.com

The average kinetic energy of electrons is given by

$$\langle W_k \rangle = \left\langle \frac{p^2}{2m} \right\rangle = \frac{\int [p^2/(2m_e)]\, dn_p}{\int dn_p} = \frac{3}{2} kT. \tag{34.9}$$

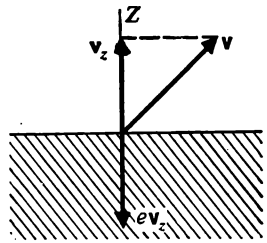


Fig. 132. To the calculation of the saturation current

Saturation current density. Let us direct the Z-axis of a Cartesian system of coordinates normally to the surface of a metal (Fig. 132). The electrons with the velocity component v_z along the Z-axis only contribute to the saturation current density. The contribution to the current density from an electron is equal to $ev_z = ep_z/m_e$. Consequently, the **saturation current density** is given by

$$j_{sat} = \frac{e}{m_e} \int\limits_{p_z>0} p_z\, dn_p = \left[\frac{2e}{m_e (2\pi\hbar)^3} \right] \exp\left(-\frac{\Phi}{kT} \right)$$

$$\times \int\limits_{-\infty}^{\infty} \exp\left(-\frac{p_x^2}{2m_e kT} \right) dp_x \int\limits_{-\infty}^{\infty} \exp\left(-\frac{p_y^2}{2m_e kT} \right) dp_y \int\limits_{0}^{\infty} p_z \exp\left(-\frac{p_z^2}{2m_e kT} \right) dp_z$$

$$= \frac{em_e k^2}{2\pi^2\hbar^3} T^2 \exp\left(-\frac{\Phi}{kT} \right), \tag{34.10}$$

or

$$j_{sat} = AT^2 \exp\left[-\Phi/(kT) \right], \tag{34.11}$$

where the constant

$$A = em_e k^2/(2\pi^2\hbar^3) = 1.2 \cdot 10^6\ \mathrm{A \cdot m^{-2} \cdot K^{-2}}. \tag{34.12}$$

Equation (34.11) is called the Richardson-Dushman equation.

For an experimental verification, it is convenient to represent this formula in the form

$$\ln (j_{sat}/T^2) = \ln A - \Phi/(kT). \tag{34.13}$$

The dependence of $\ln (j_{sat}/T^2)$ on $1/T$ expressed by formula (34.13) is a straight line (Fig. 133). Experiments confirm this form of dependence if we take into account a slight variation of Φ due to a decrease in μ with temperature [see (34.2)]. In accordance with (34.13), the slope of the curve can be used for determining the work function Φ. The value of $\ln A$ is determined by the point of intersection of the straight line with the axis of ordinates. According to formula (34.12), the quantity A should be a universal constant having the same value for all metals. This conclusion is not confirmed by experiments. The value of A slightly differs for different metals. For example, $A = 1.1 \times 10^6\ \mathrm{A \cdot m^{-2} \cdot K^{-2}}$ for copper, $1.2 \times 10^6\ \mathrm{A \cdot m^{-2} \cdot K^{-2}}$ for nickel, and $0.3 \times 10^6\ \mathrm{A \cdot m^{-2} \cdot K^{-2}}$ for platinum. This variation of A is due to surface effects. Besides, the current densities have different values for different faces of a crystal.

www.EngineeringBooksPDF.com

Three-halves power law. Let us consider the dependence of the current flowing in vacuum between two electrodes on the applied potential difference. We shall assume that the electrodes are flat and direct the X-axis along the normal to their surfaces (Fig. 134). The potential of the cathode is assumed equal to zero ($\Phi_c = 0$), while the anode potential is denoted by U.

The main physical factor that influences the motion of electrons between the cathode and the anode is the volume charge: the forces of interaction with this

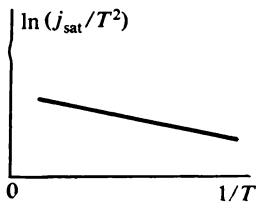


Fig. 133. Temperature dependence of the saturation current

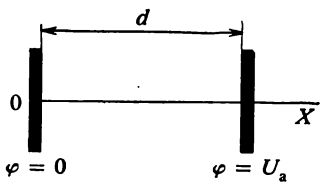


Fig. 134. To the derivation of the three-halves power law

charge hamper the motion of electrons from the cathode to the anode under the action of the applied potential difference.

Suppose that the areas of the cathode and anode plates are sufficiently large so that, while calculating the current density near the line connecting the centres of the electrodes, we can ignore the variation of quantities in the direction perpendicular to this line. In other words, we shall consider a one-dimensional problem, when all the quantities depend only on the x-coordinate. The Poisson equation for the potential has the form

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho_e}{\varepsilon_0} = \frac{n|e|}{\varepsilon_0}, \tag{34.14}$$

where n is the electron concentration. The law of conservation of energy for the electron drift can be written as

$$\frac{1}{2} m_e v_d^2 = |e| \varphi, \tag{34.15}$$

where v_d is the drift velocity at the point with a potential φ. The volume current density at this point is

$$|j| = n|e| v_d. \tag{34.16}$$

All the quantities on the right-hand side of this equation are positive. Calculating the velocity v_d from (34.15) and substituting the result into (34.16), we obtain

$$n|e| = |j| [m_e/(2|e|\varphi)]^{1/2}. \tag{34.17}$$

Taking this equation into account, we can transform Eq. (34.14) into

$$d^2\varphi/dx^2 = \alpha/\sqrt{\varphi}, \tag{34.18}$$

www.EngineeringBooksPDF.com

where $\alpha = (|j|/\varepsilon_0)\sqrt{m_e/(2|e|)}$. Multiplying both sides of (34.18) by $(d\varphi/dx) = \dot{\varphi}$, we obtain

$$\ddot{\varphi}\dot{\varphi} = \alpha\dot{\varphi}/\sqrt{\varphi}, \tag{34.19}$$

where the dots indicate the differentiation with respect to x. Considering that

$$\ddot{\varphi}\dot{\varphi} = (\dot{\varphi}^2)^{\cdot}/2 \quad \text{and} \quad \varphi/\sqrt{\varphi} = 2(\sqrt{\varphi})^{\cdot}, \tag{34.20}$$

we write (34.19) as follows:

$$(\dot{\varphi}^2)^{\cdot} = 4\alpha(\sqrt{\varphi})^{\cdot}. \tag{34.21}$$

Now we can integrate both sides of this equation with respect to x between 0 and the value of x for which the potential is equal to φ. This gives

$$\left(\frac{d\varphi}{dx}\right)^2 - \left(\frac{d\varphi}{dx}\right)_0^2 = 4\alpha\sqrt{\varphi}, \tag{34.22}$$

where we assume that $\varphi(0) = 0$. The derivative $(d\varphi/dx)_0$ characterizes the electric field strength near the cathode, and α is proportional to j. Consequently, the volume current density j attains its maximum at $(d\varphi/dx)_0 = 0$ and then [see (34.22)]

$$\frac{d\varphi}{dx} = 2\sqrt{\alpha}\varphi^{1/4}, \tag{34.23}$$

or

$$\frac{d\varphi}{\varphi^{1/4}} = 2\sqrt{\alpha}\,dx \tag{34.24}$$

Integrating both sides of this equation between $x = 0$, $\varphi = 0$ and $x = d$, $\varphi = U$, we obtain

$$U^{3/4} = \frac{3}{2}d\sqrt{\alpha}. \tag{34.25}$$

Squaring both sides of this equation and considering that

$$\alpha = (|j|/\varepsilon_0)\sqrt{m_e/(2|e|)}, \tag{34.26}$$

we obtain

$$|j| = \beta U^{3/2}, \tag{34.27}$$

where

$$\beta = \frac{4\varepsilon_0}{9d^2}\left(\frac{2|e|}{m_e}\right)^{1/2}. \tag{34.28}$$

The solution of a similar problem for coaxial cylindrical electrodes or concentric spherical electrodes leads to the same form of dependence of the volume current density on the potential difference, viz. j is proportional to U to the power of three halves. By the way, this dependence should be expected without calculation from dimensional analysis. It follows from the Poisson equation written in different coordinate systems that the coefficient β in all cases has the same dimensions.

www.EngineeringBooksPDF.com

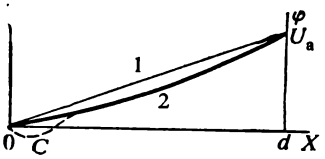


Fig. 135. Effect of volume charge on the potential distribution between the cathode and the anode

In the absence of volume charge between the cathode and the anode, the variation of the potential follows a linear law (Fig. 135, line *1*). The volume charge changes this dependence. Obviously, the volume charge near the cathode reduces the forces acting on electrons in the absence of volume charge, while near the anode these forces are increased. The variation of the potential between the electrodes taking into account the volume charge is shown by curve *2*.

The derivation of formula (34.27) is given under the assumption that electrons leave the cathode at zero velocity. However, they may leave the cathode at a finite velocity of emission. In this case, a current will exist even when there is a small reverse field near the cathode. Consequently, the volume charge density may change to such values at which the potential near the cathode is reduced to negative values. In this case, the variation of the potential near the cathode is described by the dashed curve *C*.

The deviation from the three-halves power law is observed at a sufficiently high potential difference. *This deviation becomes noticeable when the volume charge density decreases so that it becomes impossible to maintain zero electric field near the surface of the cathode, and hence the condition* $(d\varphi/dx)_0 = 0$ *under which this law was obtained is violated.* Upon a further increase in the field, the volume current density becomes independent of the potential difference (saturation current).

The three-halves power law was considered here as an illustration of the nonlinear dependence between current and voltage. It is not of universal nature and even in the case considered above it is valid only for a comparatively narrow range of voltages and currents. The nonlinearity of the current-voltage characteristic is the most important feature of many elements of radio- and electrical engineering circuits, including the elements of solid-state electronics.

What is thermionic emission?
What causes the saturation current? What does it depend on?
Under what conditions are deviations from the three-halves power law observed?

Problems

5.1. The concentration of conduction electrons in copper is $n_0 = 8.5 \times 10^{22}$ cm^{-3}. Find the average drift velocity of conduction electrons for the current density $j = 10$ A/mm^2.

5.2. A charge of $|Q|$ coulombs was passed through an electrolyte. The mobilities of ions were $b^{(+)}$ and $b^{(-)}$. Calculate the amount of electricity carried by positive and negative ions.

5.3. Two electrolytic baths with $AgNO_3$ and $CuSO_4$ solutions are connected in series. Calculate the mass of silver liberated in the time during which 10 mg of copper are liberated.

5.4. Electrolysis of $AgNO_3$ was carried out at the potential difference of 4 V. Find the electric energy spent for the liberation of 100 mg of silver.

5.5. A conducting metallic ribbon of thickness $a = 0.1$ mm and width $d = 5$ cm is placed in a uniform magnetic field with the induction $B = 1$ T perpendicular to the surface of the ribbon. The current I in the ribbon is equal to 1.6 A. Find the Hall voltage.

www.EngineeringBooksPDF.com

5.6. The saturation current in a gas-discharge tube, with electrodes having a surface area equal to 1 cm^2 and separation 3 cm is $I_{sat} = 10^{-7}$ A. The discharge is non-self-sustained. Find the number of elementary charges of each sign induced per second in 1 cm^3 of the tube volume.

Answers

1. $v_d = 0.0736$ cm/s. 5.2. $|Q^{(+)}| = \frac{b^{(-)}|Q|}{b^{(-)}+b^{(+)}}$ $|Q^{(-)}| = \frac{b^{(+)}|Q|}{b^{(-)}+b^{(+)}}$. 5.3. 34 mg.
4. 360 J. 5.5. 10^{-5} V. 5.6. $N \simeq 2 \cdot 10^{10}$ $s^{-1} \cdot cm^{-3}$.

www.EngineeringBooksPDF.com

CHAPTER 6

Stationary Magnetic Field

Stationary magnetic field is due to electric currents. It cannot be caused by the motion of an individual charge since in this case the magnetic field is necessarily varying. Nevertheless, we can use the superposition principle to draw a conclusion about the field created by an individual moving charge.

Sec. 35. Ampère's Circuital Law

The differential form of Ampère's circuital law is derived. The experimental verification of this law is discussed.

Formulation of the problem. As in electrostatics, we have to obtain the laws of magnetic field in differential form. In electrostatics this was done proceeding from Coulomb's law and the superposition principle as experimental facts. Their integral form is given by the Gauss theorem from which differential equation (13.20) follows.

In the case of a magnetic field we may, in principle, proceed in a similar way, viz. from the Biot-Savart law (10.10) or (10.11) and the superposition principle for a magnetic field as the experimental facts. Their integral form is called Ampère's circuital law (in this chapter, we shall consider it for stationary fields), from which the corresponding differential equation can be obtained. However, we may take another route and continue the theoretical derivation of the laws for a magnetic field from the laws for an electric field with the help of the theory of relativity (see Secs. 8, 9). Thus, we proceed from formula (9.28) for the magnetic induction due to the current flowing in an infinite rectilinear conductor, which has been obtained theoretically.

The integral form of Ampère's circuital law. The lines of the magnetic field generated by the current flowing in an infinite thin rectilinear wire are concentric circles with their centres at the line of the current. The value of the induction is given by formula (9.28). Let us calculate the circulation of the vector **B**

$$\oint_L \mathbf{B} \cdot d\mathbf{l} \tag{35.1}$$

www.EngineeringBooksPDF.com

around a certain closed contour L enclosing the current I (Fig. 136). Since the lines of **B** lie in the planes perpendicular to the line of current I, the contour L should be chosen in one of such planes.

Evaluating the integral (35.1) with the notation shown in Fig. 137*a*, we obtain

$$\mathbf{B}\cdot d\mathbf{l} = B\, dl \cos(\widehat{\mathbf{B}, d\mathbf{l}}) = B\, dl_{\perp}. \tag{35.2}$$

By definition, $d\alpha = dl_{\perp}/r$. Taking into account formula (10.3), we write (35.2) in the form

$$\mathbf{B}\cdot d\mathbf{l} = \frac{\mu_0}{2\pi}\frac{I}{r}\, dl_{\perp} = \frac{\mu_0}{2\pi} I\, d\alpha. \tag{35.3}$$

This gives

$$\oint_L \mathbf{B}\cdot d\mathbf{l} = \frac{\mu_0}{2\pi} I \oint_L d\alpha = \mu_0 I, \tag{35.4}$$

where we have considered that the integral of $d\alpha$ over a closed contour around the origin is equal to 2π. Consequently, the circulation of **B** around a closed

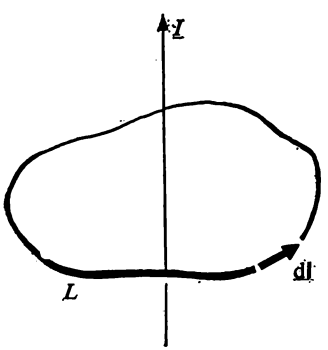


Fig. 136. Calculation of the circulation of vector B around a closed contour

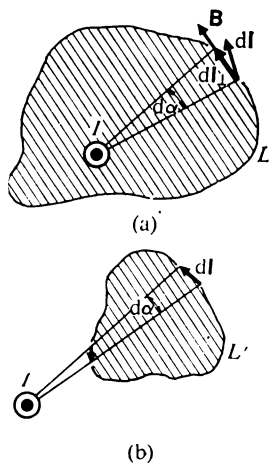


Fig. 137. Current I is perpendicular to the plane of the figure and is directed upwards. The positive direction of circumvention coincides with the counterclockwise direction.

contour surrounding the current does not depend on the shape of the contour and is determined only by the current.

If a closed contour L' does not embrace the current I (Fig. 137*b*), we get

$$\oint_{L'} d\alpha = 0, \tag{35.5}$$

www.EngineeringBooksPDF.com

i.e. the circulation of **B** around a closed contour which does not embrace the current is equal to zero. Consequently, the obtained results can be formulated as follows:

$$\oint_L \mathbf{B}\cdot d\mathbf{l} = \begin{cases} \mu_0 I \text{ (the path of integration embraces the current),} \\ 0 \text{ (the path of integration does not embrace the current).} \end{cases} \tag{35.6}$$

Suppose that we have a large number of currents and the contour embraces a part of them (Fig. 138). In accordance with the superposition principle, the magnetic induction at each point of the contour is equal to the sum of magnetic inductions of the fields created by each current:

$$\mathbf{B} = \sum_i \mathbf{B}_i. \tag{35.7}$$

Substituting **B** into the left-hand side of (35.6), we obtain

$$\int_L \mathbf{B}\cdot d\mathbf{l} = \int_L \left(\sum_i \mathbf{B}_i\right)\cdot d\mathbf{l} = \sum_i \int_L \mathbf{B}_i \cdot d\mathbf{l} = \sum_k \mu_0 I_k = \mu_0 I, \tag{35.8}$$

where the subscript k denotes only the current embraced by contour L. The currents which are not embraced by L make no contribution to the integral. Consequently, the current I in (35.8) is the sum of the currents embraced by the contour. Hence the Ampère's circuital law for the general case can be formulated as follows:

$$\boxed{\oint_L \mathbf{B}\cdot d\mathbf{l} = \mu_0 I,} \tag{35.9}$$

where I is the total current embraced by contour L. If the total current is equal to zero, the circulation is equal to zero as well. Such a situation is realized not only when the contour embraces no currents but also when the embraced currents flow in opposite directions and in total are equal to zero. For example, the circulation of **B** around a contour embracing two currents equal in magnitude and having opposite directions is equal to zero. The sign of the current I in formula (35.9) is determined in accordance with the general rule (see Sec. 14): if the direction of circumvention of contour L and the direction of the current are related through the right-hand screw rule, the current I is positive.

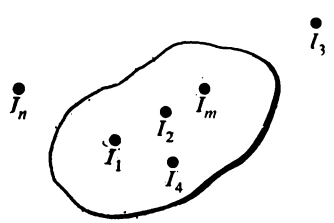


Fig. 138. Generalization of Ampère's circuital law to an arbitrary system of currents.

Otherwise, the current I has negative sign.

Expression (35.9) of Ampère's circuital law for vacuum in stationary case is a direct consequence of relation (9.28) and can be verified experimentally. In the above analysis, this law was verified for the current flowing in a straight infinite conductor. Let us show that it is also valid for an arbitrary current.

Differential form of Ampère's circuital law. Let us

www.EngineeringBooksPDF.com

write formula (35.9) for volume currents. We denote by S the surface enveloped by contour L. As usual, the positive normal to the surface is connected with the direction of circumvention of contour L through the right-hand screw rule.

The total current I flowing through this surface is given by

$$I = \int_S \mathbf{j} \cdot d\mathbf{S}, \tag{35.10}$$

where $\mathbf{j}$ is the volume current density. Consequently, Ampère's circuital law (35.9) assumes the form

$$\int_L \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_S \mathbf{j} \cdot d\mathbf{S}. \tag{35.11}$$

In accordance with Stokes' theorem, the left-hand side of this equation can be transformed to the surface integral:

$$\int_L \mathbf{B} \cdot d\mathbf{l} = \int_S \operatorname{curl} \mathbf{B} \cdot d\mathbf{S}. \tag{35.12}$$

Then (35.11) can be represented in the form

$$\int_S [\operatorname{curl} \mathbf{B} - \mu_0 \mathbf{j}] \cdot d\mathbf{S} = 0. \tag{35.13}$$

Integral (35.13) must be equal to zero for an arbitrary choice of the surface S. Consequently, the integrand is equal to zero and

$$\boxed{\operatorname{curl} \mathbf{B} = \mu_0 \mathbf{j}.} \tag{35.14}$$

This equation is called the differential form of Ampère's circuital law. It is of differential nature and is valid for any point. Hence it follows that *it is valid for an arbitrary field as well, although it has been derived for the field generated by the current flowing in an infinite rectilinear conductor.*

We can now prove that Ampère's circuital law (35.9) is valid for arbitrary currents and not only for rectilinear ones. In order to prove this, let us take arbitrary currents and draw an arbitrary surface S bounded by a closed contour L. Multiplying both sides of (35.14) by the element dS of this surface and integrating over dS, we find

$$\int_S \operatorname{curl} \mathbf{B} \cdot d\mathbf{S} = \mu_0 \int_S \mathbf{j} \cdot d\mathbf{S}. \tag{35.15}$$

We transform the left-hand side of this equality with the help of Stokes' theorem (35.12) to the contour integral and express the right-hand side, with the help of (35.10), in terms of the total current I crossing the surface. As a result, (35.15) assumes the form (35.9). This proves that the law (35.9) is valid for arbitrary currents and arbitrary contours. It should also be noted that while calculating the total current with the help of formula (35.10) we can choose any

www.EngineeringBooksPDF.com

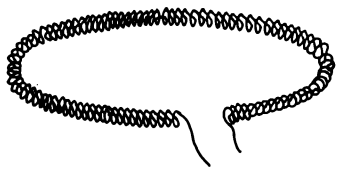


Fig. 139. Rogovskii's belt.

surface S streched over the contour L. Hence it follows that Eq. (35.14) has been obtained from Coulomb's law, the superposition principle for the electric field, invariance of charge, and formulas of the theory of relativity. The Biot-Savart law in the form (10.10) or (10.11) can be obtained from (35.14) as a solution of this equation in the absence of currents at infinity [see (37.11c)].

Experimental verification of Ampère's circuital law. Ampère's circuital law can be demonstrated and verified experimentally to a not very high degree of accuracy with the help of Rogovskii's belt. It consists of a flexible wire spiral made in the form of a belt (Fig. 139) whose ends are connected to a galvanometer. The operation of the belt is based on the Faraday's law of electromagnetic induction (see Chap. 8): an electric current appears in the circuit of the Rogovskii belt spiral upon a change in the magnetic field. The readings of the galvanometer are used to determine

$$\int_L \mathbf{B} \cdot d\mathbf{l}, \tag{35.16}$$

where L is the contour coinciding with the axis of the Rogovskii belt spiral.

In order to demonstrate Ampère's circuital law (35.9), it is sufficient to use a Rogovskii's belt in the form of a closed contour coinciding with the contours L and L' (see Fig. 137). If the current is switched on, as shown in Fig. 137*a*, the pointer of the galvanometer is deflected indicating that the integral is equal to $\mu_0 I$. In the case shown in Fig. 137*b*, the pointer is not deflected, which means that the circulation of $\mathbf{B}$ around the contour L' is equal to zero.

The derivation of the differential form of Ampère's law by direct differentiation of the Biot-Savart law. Formula (35.14) can be immediately obtained if we take the curl of both sides of formula (10.11) expressing the Biot-Savart law. On the right-hand side, this operation is applied only to the integrand since the volume of integration does not depend on the variables with respect to which the curl is taken. In the integrand, $\mathbf{j}$ is independent of these variables and only $\mathbf{r}$ and r depend on them. Taking the curl and integrating, we obtain formula (35.14). These calculations can be made as an exercise.

If the permeability of a body exceeds the permeability of the medium, it behaves as a paramagnetic; if it is lower than that of the medium, the body behaves as a diamagnetic. The circulation of magnetic induction around a closed contour enclosing a current does not depend on the shape of the contour and is determined only by the current.

Example 35.1. *Using Ampère's circuital law, find the magnetic induction in a coaxial cable which is used for transmitting a direct current* (Fig. 140). *The current flows in a central core of radius r_1 and returns along the sheath whose inner and outer radii are equal to r_2 and r_3 respectively. The space between the core and the sheath is filled with a dielectric.*

www.EngineeringBooksPDF.com

Taking into account the axial symmetry of the magnetic field, we get from Ampère's law

$$B=\frac{\mu}{2\pi}\frac{I_r}{r},$$

where I_r is the current embraced by the circular contour of radius r. The current density in the core is $j_1 = I/(\pi r_1^2)$. Hence, for $0 < r < r_1$, we obtain $I_r = j_1\pi r^2 = Ir^2/r_1^2$. Consequently,

$$B = \mu Ir/(2\pi r_1^2).$$

When $r_1 < r < r_2$, we have $I_r = I = \text{const}$, and hence

$$B = \mu I/(2\pi r).$$

If $r_2 < r < r_3$, the contour embraces the reverse current whose density is

$$j_2 = I/[\pi\,(r_3^2 - r_2^2)].$$

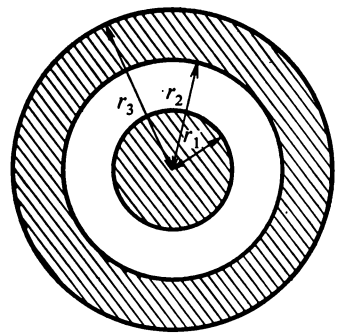


Fig. 140. Coaxial cable.

Then the current embraced by the contour for $r_2 < r < r_3$ and the magnetic induction are given by

$$I_r = I - I\frac{r^2-r_2^2}{r_3^2-r_2^2},\quad B=\frac{\mu I}{2\pi r}\left(1-\frac{r^2-r_2^2}{r_3^2-r_2^2}\right).$$

Outside the cable, the magnetic induction vanishes.

Sec. 36. Maxwell's Equations for a Stationary Magnetic Field

Maxwell's equations for the special case of a stationary magnetic field are formulated. The types of problems involved are discussed.

Equation for div B. Let us calculate div **B** proceeding from formula (9.30):

$$\operatorname{div}\mathbf{B}=\frac{\mu_0 I}{4\pi}\int_{-\infty}^{\infty}\operatorname{div}\left(\mathbf{i}_x\times\frac{\mathbf{r}}{r^3}\right)dx', \tag{36.1}$$

where the operation div is taken under the integral since the integration limits do not depend on the variables with respect to which differentiation is performed while calculating the divergence. It is expedient to write the variables in Eq. (36.1) in the explicit form for convenience of further transformations. Let **B** be the magnetic induction at the point (x, y, z), i.e. $\mathbf{B} = \mathbf{B}(x, y, z)$. The calculation of divergence is reduced to differentiation with respect to x, y, z. We denote the running coordinates of the points of integration in the integrand of (36.1) by x'. Then

$$\mathbf{r}=\mathbf{i}_x(x-x')+\mathbf{i}_y y+\mathbf{i}_z z,\qquad r=\sqrt{(x'-x)^2+y^2+z^2}. \tag{36.2}$$

www.EngineeringBooksPDF.com

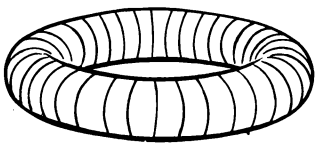


Fig. 141. A line of force is not closed when the ratio of the torus circumference to the spiral pitch is an irrational number.

In accordance with formula (A.15), we have

$$\operatorname{div}\left(\mathbf{i}_x \times \frac{\mathbf{r}}{r^3}\right) = \frac{\mathbf{r}}{r^3} \operatorname{curl} \mathbf{i}_x - \mathbf{i}_x \operatorname{curl} \frac{\mathbf{r}}{r^3} = 0, \tag{36.3}$$

since the first term on the right-hand side is equal to zero as $\mathbf{i}_x$ is independent of the coordinates (x, y, z) with respect to which the differential is performed during the evaluation of the curl. The equality of the second term to zero can be proved by a direct calculation of curl $(\mathbf{r}/r^3)=0$. The fact that curl $(\mathbf{r}/r^3)$ is equal to zero is a direct consequence of the central symmetry of the field of vector $\mathbf{r}/r^3$. It can be easily shown that any centrally symmetric field is a potential field. We leave it for the reader to prove this as an exercise.

Thus, the integrand in (36.1) is identically equal to zero, and hence

$$\boxed{\operatorname{div} \mathbf{B} = 0.} \tag{36.4}$$

From this equation, we conclude (see Sec. 13) that the lines of **B** do not have sources. This means that *there are no magnetic charges which would generate a magnetic field in the same way as electric charges create an electric field.* The lines of **B** have neither beginning nor end. They are either closed or go to infinity. The absence of the beginnings and ends in such lines is obvious. However, there may exist unclosed lines contained in a finite region of space, which nevertheless have no beginning and no end. Let us consider, for example, a torus (Fig. 141), on whose surface a spiral is wound. If the ratio of the large circumference of the torus to the spiral pitch is an irrational number, the field line will never be closed and will be wound around the torus an infinite number of times. Such a line is an example of an open line which does not have a beginning or end and is contained in a finite region of space. The lines of field **B** of this type can be easily obtained in experiments. For this purpose, we must pass the current I_1 along the axis of the torus, perpendicularly to its plane and the current I_2 along the large circle coinciding with the axis of the spiral. At a certain ratio between I_1 and I_2, the above conditions for the existence of unclosed lines of **B** will be realized.

Maxwell's equations. Equations (35.14) and (36.4) form a system of Maxwell's equations for the magnetic field generated by direct currents in vacuum:

$$\operatorname{curl} \mathbf{B} = \mu_0 \mathbf{j}, \tag{35.5}$$

$$\operatorname{div} \mathbf{B} = 0. \tag{35.6}$$

The solution of these equations allows us to determine **B** provided that **j** is known. The number of unknown scalar quantities in these equations is three (B_x, B_y, B_z), while the total number of scalar equations for determining these unknowns is four (three scalar equations obtained from the first vector equation and one scalar equation (36.6)). Thus, the number of equations is larger than

www.EngineeringBooksPDF.com

the number of unknowns. This, however, does not make the system overdetermined (see Sec. 58).

The types of problems involved. Using Eqs. (36.5) and (36.6), two problems can be solved.

1. Knowing the magnetic induction, find the volume density of currents. For this purpose, curl **B** must be calculated by using Eq. (36.5).

2. Knowing the density of currents, calculate the magnetic induction of the field generated by them. For this purpose, the above equations must be solved in unknown **j**'s. The methods of solution of these equations will be considered later. Here, we note that for the case when all currents are concentrated in a finite region of space, the solution is given by formula (10.11) expressing the Biot-Savart law:

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j} \times \mathbf{r}}{r^3} \, dV. \tag{36.7}$$

The complex structure of the integrand and its vectorial nature make the calculations quite cumbersome. For the sake of simplification, it is expedient to introduce the vector potential.

The equation div B = 0 shows that the magnetic field lines have neither beginning nor end: they are either closed or go to infinity. These lines can also be concentrated in a finite region of space, but in this case too they do not have a beginning or end. This means that there are no magnetic charges which would create a magnetic field in the same way as electric charges create an electric field. There are four scalar equations (36.5) and (36.6) for determining the three components of the magnetic induction vector. This, however does not make the system of equations overdetermined (see Sec. 58).

Give an example of a line lying entirely in a finite region of space and having neither beginning nor end.

Sec. 37. Vector Potential

The properties of the vector potential and its gauging are discussed. The magnetic induction of an electric current is calculated.

The possibility of introducing a vector potential. The identity div curl $\equiv 0$ which is known from vector calculus shows that the solution of the equation

$$\operatorname{div} \mathbf{B} = 0 \tag{37.1}$$

can be represented in the form

$$\mathbf{B} = \operatorname{curl} \mathbf{A}, \tag{37.2}$$

where **A** is the vector potential of a magnetic field.

Ambiguity of vector potential. A field with a given magnetic induction **B** can be described by many vector potentials rather than by a single potential. In

www.EngineeringBooksPDF.com

order to verify this, we shall prove that if the potential **A** describes a field with magnetic induction **B**, another potential

$$\mathbf{A}' = \mathbf{A} + \operatorname{grad} \chi, \tag{37.3}$$

where χ is an arbitrary function, also describes the same field **B**. For this purpose, we calculate the magnetic induction **B'** of the field described by the potential **A'**:

$$\mathbf{B}' = \operatorname{curl} \mathbf{A}' = \operatorname{curl} \mathbf{A} + \operatorname{curl} \operatorname{grad} \chi = \operatorname{curl} \mathbf{A} = \mathbf{B}, \tag{37.4}$$

since curl grad $\equiv 0$.

The ambiguity of vector potential is similar to the ambiguity of the scalar potential in the electrostatic field theory, *the only difference being that the scalar potential is defined accurate to an arbitrary constant, while the vector potential is defined to within an arbitrary function of a certain class.*

Potential gauging. Since the potential is chosen ambiguously, we can impose a certain condition on it. In magnetostatics, this condition is most frequently chosen in the form

$$\operatorname{div} \mathbf{A} = 0 \tag{37.5}$$

and is called the **gauging condition for the potential.** Its role is similar to that of the scalar potential in electrostatics. In particular, *the arbitrariness in the choice of the vector potential indicates that the vector potential plays only an auxiliary role and cannot be measured experimentally.*

Equation for vector potential. Substituting (37.2) into (36.5), we obtain

$$\operatorname{curl} \operatorname{curl} \mathbf{A} = \mu_0 \mathbf{j}. \tag{37.6}$$

It is known from vector calculus that

$$\operatorname{curl} \operatorname{curl} \mathbf{A} = \operatorname{grad} \operatorname{div} \mathbf{A} - \nabla^2 \mathbf{A}, \tag{37.7}$$

and hence Eq. (37.6) assumes the form

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{j}, \tag{37.8}$$

where gauge condition (37.5) is taken into account. Let us write Eq. (37.8) in terms of coordinates:

$$\nabla^2 A_x = -\mu_0 j_x, \quad \nabla^2 A_y = -\mu_0 j_y, \quad \nabla^2 A_z = -\mu_0 j_z. \tag{37.9}$$

Thus, each component of the vector potential satisfies Poisson's equation (see Sec. 15). In particular, if all currents are concentrated in a finite region of space, then by analogy with function (14.35) which is the solution of (15.14), we can write the solution of Eqs. (37.9) in the form

$$A_x = \frac{\mu_0}{4\pi} \int \frac{j_x \, dV}{r}, \quad A_y = \frac{\mu_0}{4\pi} \int \frac{j_y \, dV}{r}, \quad A_z = \frac{\mu_0}{4\pi} \int \frac{j_z \, dV}{r} \tag{37.10}$$

or, in vector form

$$\boxed{A = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}}{r} \, dV.} \tag{37.11a}$$

www.EngineeringBooksPDF.com

For a line current, we have

$$\mathbf{A}=\frac{\mu_0}{4\pi}\int\limits_L \frac{I\,\mathrm{d}\mathbf{l}}{r}=\frac{\mu_0}{4\pi}\sum_i I_i\int\limits_{L_i}\frac{\mathrm{d}\mathbf{l}}{r}\,, \tag{37.11b}$$

where L_i are the contours of currents. The currents I_i are generally different in different contours. While integrating over the closed contour L_i around a particular current I_i, this current can be taken out of the integral as in the sum of (37.11b).

Having found the vector potential, we can determine the corresponding magnetic induction with the help of formula (37.2).

Biot-Savart law. Using (37.2), we can obtain the following expression for the magnetic induction from (37.11*a*):

$$\mathbf{B}(x,\ y,\ z)=\frac{\mu_0}{4\pi}\int \operatorname{curl}\left[\frac{\mathbf{j}(x',\ y',\ z')}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}}\right]\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z',$$

where the coordinates of the point of observation at which the curl is calculated and the running coordinates $(x',\ y',\ z')$ of the integration point are written explicitly. The curl operation involves the calculation of partial derivatives with respect to $(x,\ y,\ z)$. Taking into account the formula curl $(\varphi\mathbf{A})$ = φ curl $\mathbf{A}$ + grad $\varphi \times \mathbf{A}$ of vector calculus, we obtain

$$\operatorname{curl}\frac{\mathbf{j}}{r}=\frac{1}{r}\operatorname{curl}\mathbf{j}+\operatorname{grad}\frac{1}{r}\times\mathbf{j}=\frac{\mathbf{j}\times\mathbf{r}}{r^3}\,,$$

where curl $\mathbf{j} = 0$ since $\mathbf{j}$ does not depend on the variables with respect to which the curl is evaluated, and grad $(1/r) = -\mathbf{r}/r^3$. Consequently, we obtain the formula

$$\boxed{\mathbf{B}=\frac{\mu_0}{4\pi}\int\frac{\mathbf{j}\times\mathbf{r}}{r^3}\,\mathrm{d}V,} \tag{37.11c}$$

which expresses the **Biot-Savart law.** This completes the derivation of the basic magnetostatic field laws from the electrostatic field laws with the help of the theory of relativity.

The field of an elementary current. Let us calculate the vector potential and the magnetic induction of the field created by an elementary closed current, viz. the line current flowing around a surface with very small linear dimensions. We shall choose the current loop in the form of a parallelogram with sides l_1, l_2, l_3 and l_4 (Fig. 142). We place the origin of coordinates at a point O of the surface about which the current flows. The choice of the point O is arbitrary since the loop and the surface are infinitely small. The potential is calculated at the point characterized by the radius vector $\mathbf{r}$. Using formula (37.11b), we obtain

$$\mathbf{A}(\mathbf{r})=\frac{\mu_0}{4\pi}I\int\limits_{l_1l_2l_3l_4}\frac{\mathrm{d}\mathbf{l}}{r}\,, \tag{37.12}$$

where a transition to rectilinear currents is made ($\mathbf{j}\,\mathrm{d}V \to I\,\mathrm{d}\mathbf{l}$).

www.EngineeringBooksPDF.com

Since the sides of the parallelogram are very small, we can assume that in the integration of (37.12) over each side, the value of r is constant and equal, for

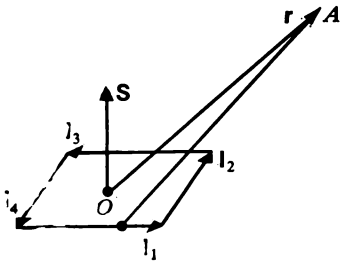


Fig. 142. Elementary current.

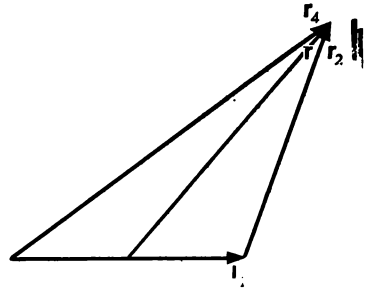


Fig. 143. Calculation of the difference in the distances from two points.

example, to the distance from the point at which the field is determined to the middle of the side. Hence [see (37.12)],

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} I \left(\frac{1}{r_1} \int_{l_1} d\mathbf{l} + \frac{1}{r_2} \int_{l_2} d\mathbf{l} + \frac{1}{r_3} \int_{l_3} d\mathbf{l} + \frac{1}{r_4} \int_{l_4} d\mathbf{l} \right)$$

$$= \frac{\mu_0}{4\pi} I \left(\frac{\mathbf{l}_1}{r_1} + \frac{\mathbf{l}_2}{r_2} + \frac{\mathbf{l}_3}{r_3} + \frac{\mathbf{l}_4}{r_4} \right). \qquad (37.13)$$

Considering that $\mathbf{l}_1 = -\mathbf{l}_3$ and $\mathbf{l}_2 = -\mathbf{l}_4$, we find that

$$\frac{\mathbf{l}_1}{r_1} + \frac{\mathbf{l}_3}{r_3} = \mathbf{l}_1 \left(\frac{1}{r_1} - \frac{1}{r_3} \right) = \mathbf{l}_1 \left(\frac{r_3 - r_1}{r_1 r_2} \right) \approx \frac{\mathbf{l}_1 (-\mathbf{l}_2 \cdot \mathbf{r})}{r^3} = -\frac{\mathbf{l}_1 (\mathbf{l}_2 \cdot \mathbf{r})}{r^3},$$

$$\frac{\mathbf{l}_2}{r_2} + \frac{\mathbf{l}_4}{r_4} = \mathbf{l}_2 \left(\frac{1}{r_2} - \frac{1}{r_4} \right) = \mathbf{l}_2 \left(\frac{r_4 - r_2}{r_2 r_4} \right) = \frac{\mathbf{l}_2 (\mathbf{l}_1 \cdot \mathbf{r})}{r^3}, \qquad (37.14)$$

where we have considered that higher-order infinitesimals can be neglected in calculations. For example, Fig. 143 shows the geometrical constructions used for the calculation of the second group of equalities (37.14):

$$\mathbf{r}_4 = \mathbf{l}_1 + \mathbf{r}_2, \qquad (37.15)$$

whence

$$r_4^2 = l_1^2 + r_2^2 + 2\mathbf{l}_1 \cdot \mathbf{r}_2 \qquad (37.16)$$

and hence

$$r_4^2 - r_2^2 = (r_4 - r_2)(r_4 + r_2) = l_1^2 + 2\mathbf{l}_1 \cdot \mathbf{r}_2. \qquad (37.17)$$

This gives

$$r_4 - r_2 = \frac{2\mathbf{l}_1 \cdot \mathbf{r}_2 + l_1^2}{r_4 + r_2} \approx \mathbf{l}_1 \cdot \mathbf{r}/r. \qquad (37.18)$$

www.EngineeringBooksPDF.com

Here we retained only the first-order terms in l_1. Using the equalities of the form (37.18), we obtain formulas (37.14). Taking into account (37.14), we can write expression (37.13) for the potential in the form

$$\mathbf{A}=\frac{\mu_0}{4\pi}\frac{I}{r_3}[\mathbf{l}_2(\mathbf{l}\cdot\mathbf{r})-\mathbf{l}_1(\mathbf{l}_2\cdot\mathbf{r})]. \tag{37.19}$$

In vector algebra, the following formula for the decomposition of a vector triple product is known:

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=\mathbf{B}(\mathbf{A}\cdot\mathbf{C})-\mathbf{C}(\mathbf{A}\cdot\mathbf{B}), \tag{37.20}$$

which shows that the expression in the brackets in (37.19) can be represented in the form

$$\mathbf{l}_2(\mathbf{l}_1\cdot\mathbf{r})-\mathbf{l}_1(\mathbf{l}_2\cdot\mathbf{r})=\mathbf{r}\times(\mathbf{l}_2\times\mathbf{l}_1)=(\mathbf{l}_1\times\mathbf{l}_2)\times\mathbf{r}. \tag{37.21}$$

Considering that

$$\mathbf{l}_1\times\mathbf{l}_2=\mathbf{S} \tag{37.22}$$

is the vector element of the surface around which the current flows, we can write (37.19) combined with (37.21) and (37.22) in the form

$$\mathbf{A}=\frac{\mu_0}{4\pi}\frac{I\mathbf{S}\times\mathbf{r}}{r^3}. \tag{37.23}$$

The quantity

$$I\mathbf{S}=\mathbf{p}_m \tag{37.24}$$

plays an extremely important role in magnetism and is called the **magnetic moment** of elementary current. Its magnitude is equal to the product of the current in the loop by the area bounded by the loop. Its direction coincides with the direction of the positive normal to the surface.

We represent the vector potential of the elementary current in the form

$$\boxed{\mathbf{A}=\frac{\mu_0}{4\pi}\frac{\mathbf{p}_m\times\mathbf{r}}{r^3},} \tag{37.25}$$

whence

$$\boxed{\mathbf{B}=\operatorname{curl}\mathbf{A}=\frac{\mu_0}{4\pi}\left\{\frac{3(\mathbf{p}_m\cdot\mathbf{r})\mathbf{r}}{r^5}-\frac{\mathbf{p}_m}{r^3}\right\}.} \tag{37.26}$$

This formula shows that the *magnetic induction corresponding to the magnetic moment decreases in inverse proportion to the third power of the distance, while the magnetic induction of the field of the current element decreases in inverse proportion to the square of the distance.* This is due to the fact that the magnetic induction

www.EngineeringBooksPDF.com

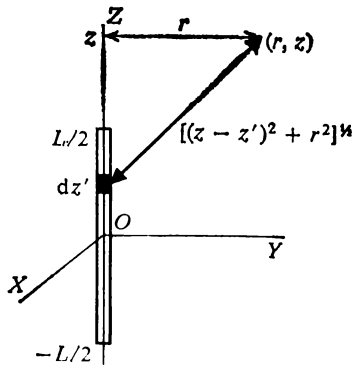


Fig. 144. To the calculation of the potential of a finite region of a rectilinear current.

associated with the magnetic moment is the sum of the magnetic inductions of the fields of current elements having opposite directions and separated by very small distances.

Example 37.1. *Find the vector potential and magnetic induction of the field created by a rectilinear segment of a line conductor of length L carrying current I.*

It is assumed that this segment is a part of a closed circuit. According to the superposition principle, this potential is a summand in the total potential of the current flowing in the closed circuit. Consequently, its calculation has a physical meaning even though an unclosed direct current does not exist.

We place the origin of coordinates at the middle of the segment under consideration and direct the Z-axis along the conductor (Fig. 144). Since the magnetic field of a rectilinear current is axisymmetric, it is sufficient to calculate the induction at points in the ZY plane. We shall characterize the coordinates of a point in this plane by the distance r from the Z-axis and by the z-coordinate. It follows from formula (37.11b) that only the component A_z differs from zero since the current flows in the direction of the Z-axis. Hence

$$A_z = \frac{\mu_0 I}{4\pi} \int_{-L/2}^{L/2} \frac{dz'}{[(z-z')^2+r^2]^{1/2}} = \frac{\mu_0 I}{4\pi} \ln \left[\frac{-z+L/2+[(z-L/2)^2+r^2]^{1/2}}{-(z+L/2)+[(z+L/2)^2+r^2]^{1/2}} \right]. \quad (37.27)$$

The magnetic induction can be calculated by the formula

$$\mathbf{B} = \operatorname{curl} \mathbf{A},$$

which should be written in cylindrical coordinates. The only nonzero component of magnetic induction $\mathbf{B}$ is B_φ where φ is the axial angle of the cylindrical system of coordinates. This component is given by

$$B_\varphi = -\partial A_z / \partial r. \quad (37.28)$$

At the points on the plane ZY in the figure, the component B_φ is perpendicular to this plane and is directed towards negative values on the X-axis. Using formula (37.28) together with (37.27), we get

$$B_\varphi = -\partial A_z / \partial r = \frac{\mu_0 I}{4\pi r} \left[\frac{-z+L/2}{[r^2+(z-L/2)^2]^{1/2}} + \frac{z+L/2}{[r^2+(z+L/2)^2]^{1/2}} \right]. \quad (37.29)$$

For a very long straight conductor, we find from (37.27) and (37.29)

$$A_z (L \to \infty) = -\frac{\mu_0 I}{2\pi} \ln r + \text{const}, \quad (37.30)$$

$$B_\varphi (L \to \infty) = \frac{\mu_0 I}{2\pi r}. \quad (37.31)$$

Example 37.2. *Find the vector potential and the magnetic induction created by the current flowing in a coaxial cable (Fig. 140), assuming that the material of the conductors and the space between them are nonmagnetic.*

www.EngineeringBooksPDF.com

The potential satisfies Eq. (37.8). In view of axial symmetry, it is convenient to use the cylindrical system of coordinates, whose Z-axis coincides with the axis of the cable. Obviously, the potential does not depend on z and the axial angle φ, i.e. $\mathbf{A} = \mathbf{A}(r)$. Besides, if only the component j_z of the current density differs from zero, then the required component A_z of the vector potential will have a nonzero value. Let us denote this component by A and use the subscripts indicating to which region this component corresponds, viz. A_1, A_2, A_3, and A_4 are the vector potentials in the regions $(0, r_1)$, (r_1, r_2), (r_2, r_3), and (r_3, ∞) respectively. Then [see (37.8)]

$$\begin{aligned}
&\nabla^2 A_1 = \frac{1}{r}\frac{d}{dr}\left(r\frac{dA_1}{dr}\right) = \frac{\mu_0 I}{\pi r_1^2} \quad (0 < r < r_1),\\
&\nabla^2 A_2 = 0 \qquad\qquad\qquad\qquad (r_1 < r < r_2),\\
&\nabla^2 A_3 = \frac{1}{r}\frac{d}{dr}\left(r\frac{dA_3}{dr}\right) = -\frac{\mu_0 I}{\pi (r_3^2 - r_2^2)} \quad (r_2 < r < r_3),\\
&\nabla^2 A_4 = 0 \qquad\qquad\qquad\qquad (r_3 < r < \infty),
\end{aligned} \tag{37.32}$$

where

$$j_1 = I/(\pi r_1^2),\ j_2 = 0,\ j_3 = I/[r_3^2 - r_2^2],\ j_4 = 0.$$

The solution of Eqs. (37.32) is given by

$$\begin{aligned}
&A_1 = -\frac{\mu_0 I r^2}{4\pi r_1^2} + C_1 \ln r + C_2 \quad (0 < r < r_1),\\
&A_2 = C_3 \ln r + C_4 \qquad\qquad (r_1 < r < r_2),\\
&A_3 = \frac{\mu_0 I r^2}{4\pi (r_3^2 - r_2^2)} + C_5 \ln r + C_6 \quad (r_2 < r < r_3),\\
&A_4 = C_7 \ln r + C_8 \qquad\qquad (r_3 < r < \infty).
\end{aligned} \tag{37.33}$$

We find the magnetic induction from the formula $\mathbf{B} = \operatorname{curl} \mathbf{A}$, which in the case under consideration is reduced to the expression $B_\varphi = -\partial A/\partial r$.

Since B_φ is the only nonzero component of the magnetic induction, the subscript φ will be omitted in the further analysis. The subscripts will indicate the regions to which the value of B corresponds. Then

$$B_1 = \frac{\mu_0 I r}{2\pi r_1^2} - \frac{C_1}{r}. \tag{37.34}$$

Since B_1 is finite for $r = 0$, we conclude that $C_1 = 0$. We choose for the normalization condition $A_1(0) = 0$. This gives $C_2 = 0$, and hence the equations for A_2 and B_2 assume the form

$$A_1 = -\mu_0 I r^2/(4\pi r_1^2),\ B_1 = \mu_0 I r/(2\pi r_1^2). \tag{37.35}$$

For the region $r_1 < r < r_2$, we obtain

$$B_2 = -C_3/r. \tag{37.36}$$

Using the boundary conditions for $\mathbf{B}$ and considering that $\mu = \mu_0$, we obtain $B_2(r_1) = B_1(r_1) = -C_3/r_1 = \mu_0 I/(2\pi r_1)$. Consequently, $C_3 = -\mu_0 I/(2\pi)$.

Let us write the continuity equation for the vector potential for $r = r_1$ in the form $C_3 \ln r_1 + C_4 = -\mu_0 I/(4\pi)$, which leads to the equation $C_4 = -\mu_0 I/(4\pi) + [\mu_0 I/(2\pi)] \ln r_1$. Hence the equations for the vector potential and the magnetic induction for $r_1 < r < r_2$ become

$$A_2 = -\frac{\mu_0 I}{2\pi}\ln\frac{r}{r_1} - \frac{\mu_0 I}{4\pi},\quad B_2 = \frac{\mu_0 I}{2\pi r}. \tag{37.37}$$

www.EngineeringBooksPDF.com

The magnetic induction in the cable sheath ($r_2 < r < r_3$) is given by

$$B_3 = -\frac{\partial A_3}{\partial r} = \frac{\mu_0 I r}{2\pi (r_3^2 - r_2^2)} - \frac{C_5}{r}.$$

Using the boundary conditions $B_2(r_2) = B_3(r_2)$ and $A_2(r_2) = A_3(r_2)$, we obtain

$$C_5 = -\frac{\mu_0 I r_3^2}{2\pi (r_3^2 - r_2^2)},$$

$$C_6 = -\frac{\mu_0 I r_3^2}{4\pi (r_3^2 - r_2^2)} + \frac{\mu_0 I r_3^2}{2\pi (r_3^2 - r_2^2)} \ln r_2 - \frac{\mu_0 I}{2\pi} \ln \frac{r_2}{r_1},$$

whence

$$\begin{aligned} A_3 &= -\frac{\mu_0 I}{4\pi} \left[\frac{r_3^2 - r^2}{r_3^2 - r_2^2} + \frac{2r_3^2}{r_3^2 - r_2^2} \ln \frac{r}{r_2} + 2 \ln \frac{r_2}{r_1} \right], \\ B_3 &= \frac{\mu_0 I (r_3^2 - r^2)}{2\pi r (r_3^2 - r_2^2)}. \end{aligned} \tag{37.38}$$

Using the boundary conditions for $r = r_3$, we find the following equations for the vector potential and magnetic induction for $r_3 < r < \infty$:

$$\begin{aligned} A_4 &= -\frac{\mu_0 I}{2\pi} \left[\frac{r_3^2}{r_3^2 - r_2^2} \ln \frac{r_3}{r_2} + \ln \frac{r_2}{r_1} \right] = \text{const}, \\ B_4 &= 0. \end{aligned} \tag{37.39}$$

Sec. 38. Magnetic Field in the Presence of Magnetics

The influence of a magnetic on the magnetic field and various mechanisms of magnetization are considered. The relation between the volume and surface densities of molecular currents and the magnetization is derived. The phenomena at the interface between magnetics are discussed and the measurement of magnetic induction in a magnetic is considered. The essence of magnetic screening is elucidated.

Definition. *Magnetics are substances which, upon being introduced into an external magnetic field, change so that they themselves become sources of an additional magnetic field.* The total magnetic induction in this case is the sum of the inductions of the external magnetic field and the magnetic field generated by the magnetic. The change in the state of a magnetic under the action of an external magnetic field, as a result of which the magnetic itself becomes a source of a magnetic field, is called magnetization of the magnetic. This phenomenon was experimentally discovered in 1845 by Faraday for a wide class of materials. He also established the existence of dia- and paramagnetic bodies for which he introduced these terms.

www.EngineeringBooksPDF.com

Mechanisms of magnetization. There are several mechanisms of magnetization. *Accordingly, magnetics are divided into dia-, para-, ferro-, and ferrimagnetics. Antiferromagnetics also belong to magnetics although they do not create a magnetic field in the space surrounding them* (see Chap. 7).

In all cases, the intensity of magnetization is quantitatively characterized in a similar way. In other words, under the action of a magnetic field, all volume elements acquire a magnetic moment. This can be realized through the following mechanisms.

1. When a magnetic is introduced into a magnetic field, the motion of electrons in atoms and molecules varies so that a total circular current oriented in a certain way appears. This current is characterized by a magnetic moment [see (37.24)]. It can be said that *the molecules of the magnetic introduced into the magnetic field acquire an induced magnetic moment.* As a result, they become the sources of an additional field whose induction is defined by formula (37.26), i.e. the substance is magnetized. Such substances are called **diamagnetics.**

2. The motion of electrons in molecules can be such that the molecules will have a magnetic moment even in the absence of a magnetic field, i.e. the *molecules possess a permanent magnetic moment.* Owing to this, each molecule is a source of a magnetic field. In the absence of an external field, the magnetic moments of different molecules are oriented quite randomly so that the total magnetic induction of the field created by them is equal to zero. In other words, infinitely small elements of the body are not the sources of a magnetic field, and the body is not magnetized. *When such a magnetic is introduced into an external field, the permanent magnetic moments of individual molecules are reoriented in the direction of the magnetic induction of the field, as a result of which the preferred direction of orientation of magnetic moments is singled out.* In this case, infinitely small physical volumes acquire a magnetic moment equal to the sum of magnetic moments of molecules contained in the volume and become the sources of a magnetic field. The magnetic is magnetized. Such materials are called **paramagnetics.**

3. Magnetization of ferro- and ferrimagnetics is due to the fact that electrons have a magnetic moment which is in a certain relation with their intrinsic angular momentum, viz., the spin. *Magnetization of this class of magnetics is associated with a certain orientation of spins and is called the spin magnetization.* The explanation of spin magnetization is beyond the scope of the classical theory of electricity and magnetism and is possible only in the framework of the quantum theory. For this reason, in this book we shall only describe the most important properties of this class of magnetics without presenting a quantitative theory. The theory of magnetic field in the presence of magnetics considered below refers only to dia- and paramagnetics unless the opposite is stipulated.

Magnetization. This quantity is defined as the ratio of the magnetic moment of an elementary physical volume to this volume:

$$\mathbf{J}=\frac{1}{\Delta V}\sum_{\Delta V}\mathbf{p}_{mi}, \tag{38.1}$$

www.EngineeringBooksPDF.com

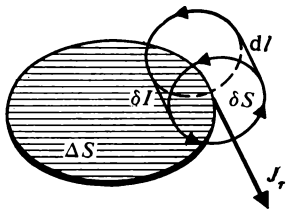


Fig. 145. To the derivation of the expression for the volume density of molecular currents.

where ΔV is the elementary volume and $\mathbf{p}_{mi}$ are the magnetic moments of molecules. The summation is performed over all the molecules in volume ΔV.

In other words, definition (38.1) of magnetization can be formulated as follows: *magnetization is the volume density of the magnetic moment of a magnetic.* It follows from (38.1) that the magnetic moment of a volume element dV is given by

$$d\mathbf{p}_m = \mathbf{J}\, dV. \tag{38.2}$$

Vector potential in the presence of magnetics. It is equal to the sum of the potential $\mathbf{A}_0$ created by the conduction currents and the potential $\mathbf{A}_m$ created by the magnetic as a result of magnetization:

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_m. \tag{38.3}$$

On the basis of (37.11), (37.25) and (38.2), we can write

$$\mathbf{A}_0 = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}}{r}\, dV, \tag{38.4a}$$

$$\mathbf{A}_m = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J} \times \mathbf{r}}{r^3}\, dV. \tag{38.4b}$$

Volume density of molecular currents. It was mentioned above that the appearance of magnetic moments is associated with the presence of circular currents. The currents creating magnetic moments of the required magnitude in elementary volumes are called molecular currents. However, it would be wrong to attach a too literal meaning to this expression. In a strict sense of the term, molecular currents may flow only within molecules. While defining magnetization and other quantities, averaged quantities are meant, owing to which the *magnetic moments of molecules are as if continuously smeared over the entire volume and molecular currents are assumed to flow over the volume of a magnetic as in a continuous medium.* Nevertheless, the term "molecular currents" was retained for them.

Let us consider a very small closed contour L bounding ΔS (Fig. 145) and calculate the circulation of the magnetization around the contour:

$$\int_L \mathbf{J} \cdot d\mathbf{l} = \int_L J_\tau\, dl, \tag{38.5}$$

where J_τ is the tangential component of $\mathbf{J}$ along the path of integration. It is created at the expense of current flowing in closed contours around the line along which the integration in (38.5) is performed (Fig. 145; δS is the area over which the current flows in the plane perpendicular to the line of inte-

www.EngineeringBooksPDF.com

gration). Multiplying the numerator and denominator in (38.5) by δS, we make the following transformations:

$$\int_L J_\tau \, dl = \int_L J_\tau \frac{dl \delta S}{\delta S} = \int_L \frac{J_\tau \, dV}{\delta S} = \int_L \frac{dp_{m\tau}}{\delta S}, \quad (38.6)$$

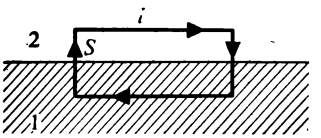


Fig. 146. To the derivation of the formula for the surface density of currents.

where we took into consideration formula (38.2). From the definition of the magnetic moment, we obtain $dp_{m\tau} = \delta I \delta S$ (δI is the current flowing around the area element δS over the length dl, δI intersecting δS along the normal). Hence,

$$\int_L \frac{dp_{m\tau}}{\delta S} = \int_L \frac{\delta I \, \delta S}{\delta S} = \int_L \delta I = \Delta I_n, \quad (38.7)$$

where ΔI_n is the normal component of the current crossing the area element ΔS. Thus, taking into account (38.6) and (38.7), we represent (38.5) in the form

$$\int_L \mathbf{J} \cdot d\mathbf{l} = \Delta I_n. \quad (38.8a)$$

Let us find the component of curl $\mathbf{J}$ in the direction of the normal to the area element ΔS. Using the definition (14.6) for the curl and Eq. (38.8a), we obtain

$$\operatorname{curl}_n \mathbf{J} = \lim_{\Delta S \to 0} \frac{\int_L \mathbf{J} \cdot d\mathbf{l}}{\Delta S} = \lim_{\Delta S \to 0} \frac{\Delta I_n}{\Delta S} = j_{mn}. \quad (38.8b)$$

The quantity

$$j_{mn} = \lim_{\Delta S \to 0} \frac{\Delta I_n}{\Delta S} \quad (38.9)$$

is obviously the normal component of the density of molecular currents since precisely these currents are responsible for the magnetization. Equation (38.8b) is valid for an arbitrary orientation of the area element ΔS, i.e. for any components of curl $\mathbf{J}$ and $\mathbf{J}_m$. Consequently, the following vector equation holds:

$$\mathbf{j}_m = \operatorname{curl} \mathbf{J}. \quad (38.10)$$

This formula gives the expression for the **volume density of molecular currents** which generate magnetization $\mathbf{J}$.

Surface molecular currents. Molecular currents may also flow over the interface between magnetics or over a magnetic-vacuum interface.

Figure 146 shows the interface between magnetics *1* and *2*. All quantities referring to magnetics *1* or *2* are marked by the subscript *1* or *2*. Let us draw a contour L in the plane perpendicular to the interface. The parts of the contour parallel to the interface are equal to l, while the perpendicular parts are very small and tend to zero. This contour bounds the area S of the surface perpendicular to the interface between the magnetics. Suppose that the element $d\mathbf{S}$

www.EngineeringBooksPDF.com

of this area is directed away from us for a chosen direction of circumvention of the contour. Multiplying both sides of (38.10) by dS and integrating over S, we obtain

$$\int_S \operatorname{curl} \mathbf{J} \cdot d\mathbf{S} = \int \mathbf{j}_m \cdot d\mathbf{S}. \tag{38.11}$$

The left-hand side of (38.11) can be transformed in accordance with Stokes' theorem to the integral around the contour L, which gives

$$\int_L \operatorname{curl} \mathbf{J} \cdot d\mathbf{S} = \int_L \mathbf{J} \cdot d\mathbf{l} = (J_{2\tau} - J_{1\tau})\, l + \langle J \rangle_{\text{lat}}\, \Delta l_{\text{lat}}, \tag{38.12}$$

where $J_{1\tau}$ and $J_{2\tau}$ are the magnetization components in the first and second media, which are tangential to the path of integration. The minus sign of $J_{1\tau}$ has appeared because of the reversal of the integration path in the second medium. The quantity $\langle J \rangle_{\text{lat}} \Delta l_{\text{lat}}$ takes into account the integrals over the vertical parts of the path. There is no need in writing these integrals in detail since they vanish as the horizontal regions of the integration path are contracted to the surface. The right-hand side of (38.11) gives the projection of current onto the direction of the normal to surface S. This direction is also tangential to the interface between the magnetics, and hence

$$\int \mathbf{j}_m \cdot d\mathbf{S} = \Delta I_{\text{m.surf}}. \tag{38.13}$$

Taking into account (38.12) and (38.13), and dividing Eq. (38.11) by l, we obtain

$$J_{1\tau} - J_{2\tau} + \langle J \rangle_{\text{lat}} \Delta l_{\text{lat}}/l = \Delta I_{\text{m.surf}}/l = i_{\text{m.surf}}, \tag{38.14}$$

where

$$i_{\text{m.surf}} = \Delta I_{\text{m.surf}}/l \tag{38.15}$$

is the projection of the surface current density onto the direction perpendicular to surface S. Contracting in (38.14) the contour to the surface ($\Delta l_{\text{lat}} \to 0$), we obtain

$$J_{2\tau} - J_{1\tau} = i_{\text{m.surf}}. \tag{38.16}$$

This formula is valid for an arbitrary orientation of the contour relative to different directions along the interface. Consequently, it is more convenient to write it in vector form. Let us denote by $\mathbf{n}$ the unit normal to the interface, which is directed into the second medium (Fig. 147). It is clear from Fig. 147 and from the meaning of the quantities appearing in the above formulas that expression (38.16) can be written in vector form as follows:

$$\mathbf{i}_m = \mathbf{n} \times (\mathbf{J}_2 - \mathbf{J}_1). \tag{38.17}$$

Uniformly magnetized cylinder. By way of an example of calculation with the help of formula (38.17), let us find the surface density of molecular current in a uniformly magnetized cylinder (Fig. 148), which can be realized in the

www.EngineeringBooksPDF.com

form of a permanent magnet. Although the nature of ferromagnetism explaining the existence of permanent magnets cannot be understood in the framework of the classical theory of magnetism, the field created by magnetized ferromagnetics in space can be described with the help of the classical theory. In this case, *magnetization of a ferromagnetic which is assumed to be known is considered to be a source of a magnetic field in the same sense as the magnetization of dia- and paramagnetics is a source of a magnetic field. The magnetization of dia- and paramagnetics exists only in the presence of an external field. The magnetization of ferro-*

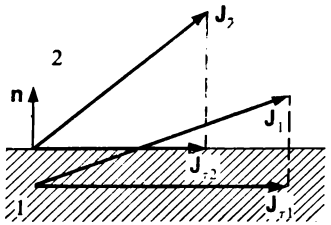


Fig. 147. To the derivation of vector form of the expression for the surface density of molecular currents.

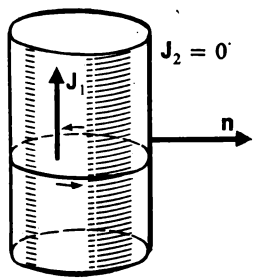


Fig. 148. Surface molecular currents flowing over a uniformly magnetized cylinder.

magnetics is retained in the absence of an external field, and the field generated by this magnetization exists independently. The problem consists in the description of this field.

A uniformly magnetized cylinder can also be imagined as a dia- or paramagnetic placed in an external field which ensures constant magnetization with a sufficient degree of accuracy. In this case, the induction of not the total field but only its part associated with magnetization is determined in the space outside the cylinder.

The magnetization $\mathbf{J}_1$ of the cylinder is shown in Fig. 148 by the arrow. In vacuum $J_2 = 0$, and the normal **n** to the interface is the outward normal to the cylinder. In accordance with formula (38.17), the **surface density of the molecular current** flowing over the cylinder is given by

$$\mathbf{i}_m = -\mathbf{n} \times \mathbf{J}_1 = \mathbf{J}_1 \times \mathbf{n}. \tag{13.18}$$

One of the lines of this current is shown in Fig. 148 by the circle with the arrows. Obviously, magnetization $\mathbf{J}_1$ forms a right-hand screw system with the current flowing over the surface of the cylinder. Formula (38.10) shows that there are no molecular volume currents inside the cylinder since curl $\mathbf{J}_1 = 0$. Consequently, the entire field outside the cylinder is created by the surface currents flowing in circles. Thus, we have proved the equivalence of the field of a permanent cylindrical magnet and the field of circular currents (the field of a solenoid). This statement is valid for all magnetics, including ferromagnetics.

www.EngineeringBooksPDF.com

Magnetic field strength. In the absence of magnetics, the following relation is valid:

$$\operatorname{curl} \mathbf{B} = \mu_0 \mathbf{j}.$$

This relation describes the generation of a magnetic field by conduction currents. In the presence of magnetics, a field is generated by molecular currents $\mathbf{j}_m$ and by conduction currents $\mathbf{j}$ [see (38.10)]. Consequently, in the presence of magnetics Eq. (38.18) can be written in the form

$$\operatorname{curl} \mathbf{B} = \mu_0 (\mathbf{j} + \mathbf{j}_m) = \mu_0 (\mathbf{j} + \operatorname{curl} \mathbf{J}). \tag{38.19}$$

Dividing both sides of this equation by μ_0 and transferring curl $\mathbf{J}$ to the left-hand side, we obtain

$$\operatorname{curl} (\mathbf{B}/\mu_0 - \mathbf{J}) = \mathbf{j}, \tag{38.20}$$

where

$$\mathbf{H} = \mathbf{B}/\mu_0 - \mathbf{J} \tag{38.21}$$

is the **magnetic field strength.** It is not a purely field quantity, since it includes vector $\mathbf{J}$ which characterizes the magnetization of the medium. Hence, *the vector* $\mathbf{H}$ *plays in the magnetic field theory the same role as the vector* $\mathbf{D}$ *in the electric field theory*, and the term field strength applied to it is not quite correct. Nevertheless, this term is historically used for this quantity.

Equation for the magnetic field strength. Taking into account (38.21) we can write Eq. (38.20) in the form

$$\operatorname{curl} \mathbf{H} = \mathbf{j}. \tag{38.22a}$$

This is a very convenient equation for calculating the magnetic field in the presence of magnetics.

In the presence of magnetics, Ampère's circuital law is derived in the same way as in the absence of magnetics, viz. by proceeding from (35.14) with a subsequent transition to (35.15):

$$\int_L \mathbf{H} \cdot d\mathbf{l} = I. \tag{38.22b}$$

Relation between magnetization and magnetic field strength. For the same historical reasons as those concerning the term magnetic field strength for the vector $\mathbf{H}$, this vector and not $\mathbf{B}$ was assumed to be the source of magnetization. Hence the relation between $\mathbf{J}$ and $\mathbf{H}$ is represented in the form

$$\mathbf{J} = \chi \mathbf{H}, \tag{38.23}$$

where χ is the magnetic susceptibility. The relation between $\mathbf{B}$ and $\mathbf{H}$ is usually written in the form

$$\mathbf{B} = \mu \mathbf{H}, \tag{38.24}$$

where μ is the permeability of the medium. These quantities do not depend on $\mathbf{B}$ and $\mathbf{H}$ in the case of dia- and paramagnetics. In order to find the relation between χ and μ, we substitute (38.23) and (38.24) into (38.21) and cancel $\mathbf{H}$ from both sides of the tobained relation:

$$1 = \mu/\mu_0 - \chi, \tag{38.25}$$

www.EngineeringBooksPDF.com

or

$$\chi = (\mu - \mu_0)/\mu_0 = \mu_r - 1, \tag{38.26}$$

where $\mu_r = \mu/\mu_0$ is the relative permeability of the medium. It should be noted that in the Gaussian system of units, the permeability is expressed by a number equal to $1/4\pi$ of the corresponding value in SI.

Different mechanisms of magnetization lead to different dependences of **J** *on* **H** (see Chap. 7). Here, it should only be noted that *magnetization in diamagnetics is directed against* **H**. In diamagnetics $\chi < 0$ [see (38.23)], and hence, in accordance with (38.26), the permeability $\mu < \mu_0$ ($\mu_r < 1$). This means that the field generated by a diamagnetic is directed against the initial field, i.e. *a diamagnetic weakens the external field.* The magnitude $|\chi|$ of their susceptibility is very small (of the order of 10^{-5}). The magnetic susceptibility of diamagnetics is independent of temperature. Diamagnetism is exhibited in all substances.

In paramagnetics, **J** *coincides in direction with* **H**. In this case, $\chi > 0$, $\mu > \mu_0$ and $\mu_r > 1$. The additional field in paramagnetics coincides with the initial one. Consequently, a *paramagnetic strengthens the field.* The susceptibility χ of paramagnetics is temperature-dependent. *At room temperature, the paramagnetic susceptibility of materials in the solid state is of the order of* 10^{-3}, *i.e. is two orders of magnitude higher than the diamagnetic susceptibility.* Consequently, the role of diamagnetic susceptibility in paramagnetic materials is small and can be neglected.

In ferromagnetics, the vector **J** *is directed along* **H** *and has a very large magnitude.* For these materials, $\chi \gg 1$ and $\mu \gg \mu_0$. A typical feature of ferromagnetics is that the values of χ and μ depend on the field and the past history of magnetization. For this reason, they exhibit residual magnetization, i.e. *the magnetization of a sample, on the whole, is preserved even after the external field has become equal to zero. In their formal properties, ferromagnetics are similar to ferroelectrics* (see Sec. 23).

Field in a magnetic. In vacuum, $\mathbf{J} = 0$ and formula (38.21) allows us to define the magnetic field strength in vacuum by the equation $\mathbf{H}_0 = \mathbf{B}/\mu_0$. The conduction currents generate a field **H** in an infinite homogeneous magnetic [see (39.22)]. Equation (35.14) can be written in the form

$$\operatorname{curl} \mathbf{H}_0 = \mathbf{j}. \tag{38.27}$$

A comparison of (38.22) with (38.27) leads to the conclusion that identical conduction currents excite identical magnetic fields in vacuum and in an infinite homogeneous magnetic

$$\mathbf{H} = \mathbf{H}_0. \tag{38.28}$$

Consequently, the magnetic inductions **B** in a magnetic and $\mathbf{B}_0$ in vacuum are connected through the following relation:

$$\mathbf{B} = \mu\mathbf{B}_0/\mu_0 = \mu_r\mathbf{B}_0. \tag{38.29}$$

This relation shows that the magnetic induction in diamagnetics is smaller than its counterpart in vacuum ($\mu_r < 1$) and larger in paramagnetics ($\mu_r > 1$).

www.EngineeringBooksPDF.com

If all magnetics and conduction currents are located in a finite region of space and both the conduction currents and magnetizations of all the magnetics are defined as position functions [$\mathbf{J} = \mathbf{J}(x, y, z)$], the magnetic induction may, in principle, be found in a fairly simple way. The vector potential can be represented in the form of formulas (38.3), (38.4a) and (38.4b) which can be written in a more convenient manner. It can be said that the vector potential $\mathbf{A}$ is the sum of the potentials created by conduction currents (38.4a), molecular currents (38.10) and surface molecular currents (38.17). All these currents create potentials in accordance with the same law (38.4a). Hence, the formula for the potential has the form

$$\mathbf{A}=\frac{\mu_0}{4\pi}\left(\int\limits_V \frac{\mathbf{j}}{r}\,dV+\int\limits_V \frac{\operatorname{curl}\mathbf{J}}{r}\,dV+\int\limits_S \frac{\mathbf{n}\times(\mathbf{J}_1-\mathbf{J}_2)}{r}\,dS\right), \qquad (38.30a)$$

where the last integral takes into account surface molecular currents, and S corresponds to the sum of the interfacial areas between magnetics.

However, the simplicity of determining the potential with the help of (38.30a) is apparent since the potential can be found if only $\mathbf{J}$ is known. This quantity is, however, unknown in many cases and its determination is fraught with considerable difficulties.

Permanent magnets. These materials are either ferro- or ferrimagnetics, and hence the theory described above is inapplicable to them. *Nevertheless, the formulas obtained above can be formally used for calculating the potential of the field created by permanent magnets in the surrounding medium.* The magnetic properties of permanent magnets, as well as of magnetics, are characterized by their magnetization $\mathbf{J}_p$ which generates the field in the same way as if it were the magnetization of a dia- or paramagnetic. Consequently, using formula (38.30a), we can write the following formula for the vector potential generated by a permanent magnet:

$$A_p=\frac{\mu_0}{4\pi}\int\limits_V \frac{\operatorname{curl}\mathbf{J}_p}{r}\,dV+\frac{\mu_0}{4\pi}\int\limits_S \frac{\mathbf{J}_p\times\mathbf{n}}{r}\,dS. \qquad (38.30b)$$

In particular, if the magnetization of a permanent magnet is uniform over the entire volume, the first term in (38.30b) vanishes, and the entire magnetic field is as if created by currents flowing over the surface of the magnet in accordance with the second integral. However, there are no real currents flowing over the surface of a permanent magnet. In the case under consideration, they are just auxiliary quantities for calculating the field strength. The physical content of the auxiliary nature of these quantities can be grasped from the following example. Suppose that we have a permanent magnet in the form of a long cylinder, creating a certain field in the surrounding space. By an appropriate choice of current in a cylindrical solenoid of the same length and diameter with a sufficiently tight winding and a para- or diamagnetic core, we can create a field in the medium surrounding a magnetic whose magnetic induction will practically coincide with the induction of the field of the permanent magnet.

www.EngineeringBooksPDF.com

The current flowing in the solenoid through thin wires can be treated as a surface current flowing over the surface of the permanent magnet. This is the mathematical meaning of the second term on the right-hand side of (38.30b). *The fictitious nature of this current is revealed when the question about the field in a magnet and in a solenoid is considered.* These fields turn out to be different.

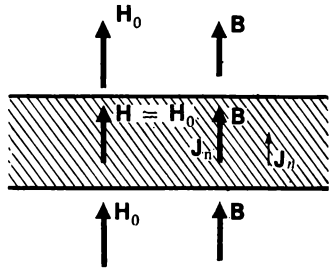


Fig. 149. Magnetic field in the presence of a ferromagnetic.

When permanent magnets are considered, the equation div $\mathbf{B} = 0$ for magnetic induction remains unchanged, while the equation relating the induction with the magnetic field strength somewhat changes. A permanent magnet is an additional source of the magnetic field, and hence we must write the equation

$$\mathbf{B} = \mu_0\mathbf{H} + \mu_0\mathbf{J} + \mu_0\mathbf{J}_p, \tag{38.31a}$$

instead of (38.21), where $\mathbf{J}_p$ is the magnetization of the permanent magnet. Considering that $\mu_0\mathbf{H} + \mu_0\mathbf{J} = \mu\mathbf{H}$, we obtain

$$\mathbf{B} = \mu\mathbf{H} + \mu_0\mathbf{J}_p. \tag{38.31b}$$

It should be noted that μ in this formula is only the diamagnetic and paramagnetic permeability of a substance rather than the ferromagnetic permeability that has already been taken into account by the term $\mu_0\mathbf{J}_p$. Consequently, if we treat $\mathbf{J}_{tot}$ as the total magnetization ($\mathbf{J}_{tot} = \mathbf{J} + \mathbf{J}_p$), it is more convenient to represent (38.31a) in the form

$$\mathbf{B} = \mu_0\mathbf{H} + \mu_0\mathbf{J}_{tot}. \tag{38.31c}$$

Let us consider, for example, a permanent magnet in the form of a plane plate of a finite thickness and infinite area (Fig. 149). The permanent magnetization $\mathbf{J}_p$ is normal to the surface of the permanent magnet. Dia- and paramagnetic properties of the permanent magnet are disregarded.

Suppose that outside the permanent magnet, a magnetic field of strength $\mathbf{H}_0$ is perpendicular to its surface. The magnetic induction of this field inside and outside the magnet is the same: $B = \mu_0 H_0$. Then [see (38.31c)], $\mu_0 H_0 = \mu_0 H + \mu_0 J_p$. Hence the magnetic field strength inside the permanent magnet is (see Fig. 149)

$$H = H_0 - J_p.$$

Boundary conditions for the field vectors. Vectors **B** and **H** undergo abrupt changes at the boundary between magnetics with different values of μ, which are characterized by the boundary conditions. In order to derive these conditions, we proceed from Eqs. (36.4) and (38.22) which are valid for vacuum as well as for a medium filled with a magnetic. The procedure for deriving the boundary conditions is precisely the same as in the case of an electrostatic field [see Sec. 17, (17.21) and (17.30)].

The boundary condition for the normal component of vector B. This condition is derived in the same way as (17.21) is derived by proceeding from (17.17),

www.EngineeringBooksPDF.com

but instead of (17.17) we now use the equation

$$\operatorname{div} \mathbf{B} = 0. \tag{38.32}$$

This gives

$$\boxed{B_{1n} = B_{2n}.} \tag{38.33}$$

The boundary condition for the tangential component of vector H. It is derived in the same way as (17.30) is derived from (17.29), but instead of (17.29), the following equation should be used:

$$\oint_{ABCD} \mathbf{H} \cdot d\mathbf{l} = \int_S \mathbf{j} \cdot d\mathbf{S}, \tag{38.34}$$

which is obtained from (38.22) by multiplying both sides by dS and integrating over the area bounded by the contour $ABCDA$ (see Fig. 83), having transformed

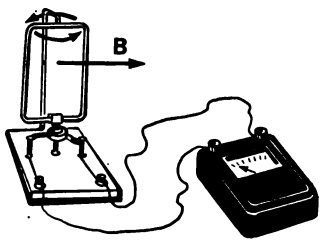


Fig. 150. Measurement of magnetic induction with the help of Faraday's law.

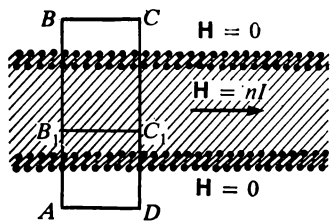


Fig. 151. Field of a very long solenoid.

its left-hand side in accordance with Stokes' theorem. As a result, we obtain

$$\boxed{H_{2\tau} - H_{1\tau} = i_{\text{surf}},} \tag{38.35}$$

where i_{surf} is the surface current density in the direction perpendicular to that of the tangential components of the magnetic field strength. It should also be kept in mind that these are surface conduction currents rather than surface molecular currents i_m [see (38.16)].

Refraction of magnetic field lines. At the interface between magnetics, the lines of force experience refraction which is determined with the help of the boundary conditions in the same way as it was done in the analysis of formula (17.31).

The measurement of magnetic induction. The most visual and simple method for measuring magnetic induction is based on the application of Faraday's law of electromagnetic induction. If a small conducting loop (Fig. 150) connected to a galvanometer is oriented in a plane perpendicular to **B** and then rotated through 90° about an axis lying in this plane, the galvanometer will register a current pulse which can be used for determining **B** in the region of

www.EngineeringBooksPDF.com

the loop (see Chap. 8). This method is used for measuring the average magnetic induction on the area bounded by the loop. Instead of rotating the loop, we can switch off the field.

The fields of a very long solenoid and a uniformly magnetized very long cylinder. Suppose that a field is created by a current flowing in the winding of a very long solenoid (Fig. 151). The number of turns per metre of the solenoid, the current and the permeability of the core are denoted by n, I, and μ respectively. The magnetic field is axially symmetric and can only have a component parallel to the axis of the solenoid (the turns are wound very tightly).

In order to find the magnetic field strength, we shall use Eq. (38.22a). Integrating along the contour $ABCDA$, we get

$$\int\limits_{ABCDA} \mathbf{H} \cdot d\mathbf{l} = 0, \tag{38.36}$$

since the currents at the opposite ends of the solenoid flow in opposite directions. Hence the total current through the surface stretched over the contour $ABCDA$ is equal to zero. The contribution to the integral from integration paths BC and DA is equal to zero since the vector $\mathbf{H}$ can be only normal to segments AB and CD:

$$H_{BC}l - H_{AD}l = 0, \tag{38.37}$$

where H_{BC} and H_{AD} are the field strengths on BC and AD, l being the length of these segments. The minus sign is due to the fact that the paths of integration are opposite on these segments. Stretching the contour along AB and CD by moving, for example, AD away from the cylinder, we notice that for (38.37) to be identically equal to zero, H must be independent of the distance, i.e. it must be constant outside the solenoid. At an infinitely long distance from the solenoid, the field is absent, and hence it is absent in the entire space outside the solenoid.

In order to determine the field strength inside the solenoid, we apply law (38.22a) to the contour AB_1C_1DA (Fig. 151). The integral differs from zero only on the segment B_1C_1, and hence

$$H_{B_1C_1}l = nlI, \tag{38.38}$$

since the surface bounded by the contour AB_1C_1DA is intersected by nl turns carrying current I. Formula (38.38) shows that the field inside the solenoid is uniform and its strength is equal to

$$H = nI. \tag{39.39}$$

This formula allows us to measure the magnetic field strength in ampere-turns, which is widely used in electrical engineering. According to this formula, the magnetic field strength inside the solenoid does not depend on its material and, other conditions being equal, is the same for all materials. Taking into account (38.24) and (38.39), we obtain the following expression for the magnetic induction of the field inside the solenoid:

$$B = \mu H = \mu nI \tag{38.40}$$

www.EngineeringBooksPDF.com

which shows that it depends on the core material. The magnetic induction in a hollow solenoid is larger than for diamagnetics and smaller than for paramagnetics.

The magnetic induction of the field of a very long uniformly magnetized cylinder can be found in a similar way, the only difference being that in this case there are no surface currents. Relation (38.37) remains unchanged, and the magnetic field strength outside the cylinder is equal to zero as in the case of a very long cylinder. Instead of formula (38.38), we obtain $Hl = 0$ or $H = 0$. This means that the magnetic field inside a very long uniformly magnetized

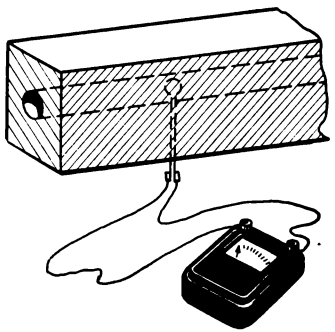


Fig. 152. Measurement of the magnetic field strength in a magnetic.

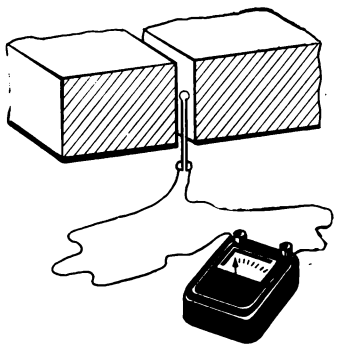


Fig. 153. Measurement of the magnetic induction in a magnetic.

cylinder is equal to zero, while for the solenoid it differs from zero. However, the magnetic induction inside the cylinder is not equal to zero ($B = \mu_0 \mathbf{J}$). If the length of the cylinder is finite, the magnetic field strength has nonzero values inside as well as outside the cylinder.

The measurement of permeability, magnetic induction and the field strength inside magnetics. Suppose that we have a very long solenoid whose core has a very narrow channel along its axis (Fig. 152). The field inside the solenoid is created by passing a current in the winding. A measuring coil connected to a galvanometer is introduced into the channel. Boundary condition (38.35) shows that the magnetic field strength in the channel is equal to that in the magnetic. The magnetic induction in the channel is given by $\mathbf{B}_{\parallel} = \mu_0 \mathbf{H}$. It can be measured by rotating the coil through 90° or by switching on the field. The field strength inside the magnetic is calculated by the formula

$$\mathbf{H} = \mathbf{B}_{\parallel}/\mu_0. \tag{38.41}$$

In order to measure the magnetic induction in the magnetic, we make a very small transverse cut in a very long solenoid (Fig. 153). Boundary condition (38.33) shows that the magnetic induction $\mathbf{B}_{\perp}$ in this cut is equal to the induction $\mathbf{B}$ in the magnetic. Consequently, it is sufficient to measure the magnetic induction in the transverse cut.

www.EngineeringBooksPDF.com

If we know the magnetic induction and field strength in a magnetic, we can determine the magnetic permeability:

$$\mu = B/H = \mu_0 B_\perp / B_{||}. \qquad (38.42)$$

A magnetic sphere in a uniform field. Suppose that a sphere of radius R made of a magnetic having permeability μ_1 is placed in an infinite medium having permeability μ_2, in which a uniform magnetic field of strength $\mathbf{H}_0$ is created (Fig. 154*a*, *b*). We shall determine the magnetic field strength in the sphere and outside it. It is assumed that conduction currents are absent.

In this case, Eq. (38.22) has the form

$$\text{curl } \mathbf{H} = 0, \qquad (38.43)$$

i.e. in the space without conduction currents, the magnetostatic field is a potential field. No conduction currents flow inside the sphere and outside it, and hence we have a potential field over the entire space. Let us denote the potential of this field by φ_m. Then

$$\mathbf{H} = -\text{grad } \varphi_m. \qquad (38.44)$$

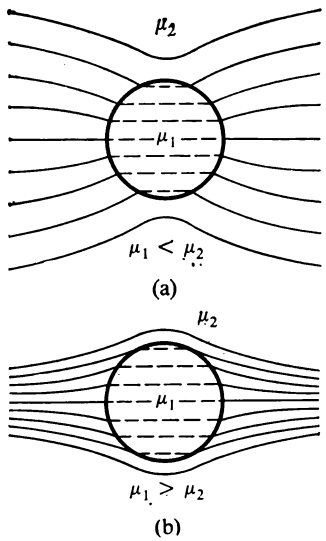


Fig. 154. A magnetic sphere in a uniform magnetic field.

For a homogeneous medium (μ = const), the equation div $\mathbf{B} = 0$ is equivalent to the equation

$$\text{div } \mathbf{H} = 0. \qquad (38.45)$$

Substituting (38.44) into (38.45), we obtain the following equation for all points outside the sphere (μ_2 = const) and for all points inside it (μ_1 = const):

$$\nabla^2 \varphi_m = 0. \qquad (38.46)$$

Thus, the magnetic field potential satisfies the Laplace equation.

It should be noted that if the permeability is not constant, we obtain another equation instead of (38.46). For its derivation, we take into account Eq. (38.21) which can be written in the form

$$\mathbf{B} = \mu_0 \mathbf{H} + \mu_0 \mathbf{J}. \qquad (38.47)$$

Taking the divergence of both sides of this equation, we obtain

$$\text{div } \mathbf{B} = \mu_0 \text{ div } \mathbf{H} + \mu_0 \text{ div } \mathbf{J} = -\mu_0 \text{ div grad } \varphi_m + \mu_0 \text{ div } \mathbf{J} = 0, \qquad (38.48a)$$

where we took into account relation (38.44) and the equation div $\mathbf{B} = 0$. Consequently, the equation for the potential φ_m has the form

$$\nabla^2 \varphi_m = \text{div } \mathbf{J}, \qquad (38.48b)$$

which considerably complicates the solution of the problem for a magnetic with a varying permeability.

Let us place the origin of coordinates at the centre of the sphere and direct the polar axis of the spherical system of coordinates along $\mathbf{H}_0$. In view of the

www.EngineeringBooksPDF.com

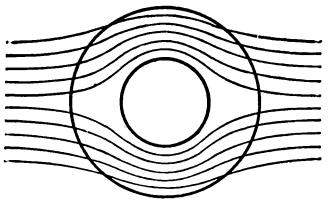


Fig. 155. Magnetic screening.

axial symmetry, the Laplace equation (38.46) assumes the form (17.42). This equation should be solved under the boundary conditions (38.33) and (38.25) on the surface of the sphere, which are completely identical to the boundary conditions for D_n and E_τ [see (17.42)]. Since there are no surface conduction currents, we can put $i_{\text{surf}} = 0$ in (38.35). Consequently, the solution of this problem is similar to the solution of the problem about a dielectric sphere in a uniform magnetic field. We must just make in the solution of Eq. (17.42) the following substitution: $\varphi \to \varphi_m$, $\mathbf{E} \to \mathbf{H}$, $\mathbf{D} \to \mathbf{B}$, and $\varepsilon \to \mu$.

The magnetic field strength inside the sphere is constant and in analogy with (17.51) is given by

$$H_{1z} = \frac{3\mu_2}{\mu_1 + 2\mu_2} H_0. \tag{38.49}$$

This is the sum of the external magnetic field strength H_0 and the magnetic field strength created by the sphere as a result of its magnetization. The field created inside the sphere due to its magnetization is called the "demagnetizing field $\mathbf{H}_{\text{dem}}$". This term is conditional since there is no demagnetization at all. Rather, the magnetic is magnetized in an external field, thus leading to the creation of an additional field which is added to the initial field. But since the term for the field H_{dem} has already been established, we have to use it. Then

$$H_{\text{dem}} = H_{1z} - H_0 = \frac{\mu_2 - \mu_1}{\mu_1 + 2\mu_2} H_0. \tag{38.50}$$

This expression can be written in a different form. Combining (38.26) and (38.23), we obtain

$$J_1 = (\mu_1/\mu_0 - 1) H_{1z}, \quad J_2 = (\mu_2/\mu_0 - 1) H_0, \tag{38.51}$$

whence

$$J_2 - J_1 = \frac{(\mu_2 - \mu_1)(\mu_0 + 2\mu_2)}{\mu_0 (\mu_1 + 2\mu_2)} H_0. \tag{38.52}$$

Consequently, formula (38.50) can be written as follows:

$$H_{\text{dem}} = [\mu_0/(\mu_0 + 2\mu_2)] (J_2 - J_1). \tag{38.53}$$

In particular, if a sphere is in vacuum, $\mu_2 = \mu_0$ and $J_2 = 0$. Hence

$$H_{\text{dem}} = -J_1/3.$$

Magnetic shielding. It follows from (38.50) that for $\mu_1 > \mu_2$ the magnetic field inside the sphere is weakened, i.e. the sphere as if screens its interior from the external magnetic field. If we calculate the induction of the field inside the cavity surrounded by the shell made of a magnetic with a sufficiently high permeability μ_1, it turns out that the magnetic lines are concentrated mainly

www.EngineeringBooksPDF.com

in the shell (Fig. 155), without penetrating into the cavity. This means that the *shell made of a magnetic with a high* μ *operates as a screen which does not allow the magnetic field to penetrate into the space bounded by the shell.*

In the literal sense, molecular currents may flow only within molecules. However, in the continuous medium model we consider the quantities averaged over infinitely small volumes. For this reason molecular currents are visualized as flowing in the volume of a magnetic as a continuous medium.
In its content, the magnetic field strength plays the same role in the magnetic field theory as the displacement in the electric field theory.
In diamagnetics, the magnetization is directed against the magnetic field strength, and the magnetic induction of an external field decreases.
In paramagnetics, the magnetization is directed along the magnetic field strength, and the magnetic induction of an external field is enhanced.
The classical theory is unable to explain ferromagnetism, but it can explain the magnetic field outside a ferromagnetic if its magnetization is considered to be known.

Which quantity in the electric field theory corresponds to permeability μ in the magnetic field theory?
Why cannot molecular current be represented as flowing only inside the molecular volume?

Why is the diamagnetism of paramagnetics insignificant in comparison with their paramagnetism? Give quantitative estimates.
How can the magnetic induction of the magnetic field inside a magnetic and the magnetic field strength be measured?

Explain why **H** in the theory of magnetic field plays the same role as **D** in the electric field theory.

Example 38.1. *A linear current I flows along the axis of a very long right circular cylinder of radius a. The permeability of the material of the cylinder is* μ. *The cylinder is surrounded by a free space. Find the magnetic field strength, magnetic induction and magnetization at all points in space.*

We direct the Z-axis of the Cartesian system of coordinates along the cylinder axis in the direction of current I (Fig. 156). Let us take for the integration contour L a circle of radius r concentric with the cylinder and lying in a plane perpendicular to the current. Then the magnetic field strength at all points can be determined from Ampère's circuital law:

$$\int_L \mathbf{H}\cdot d\mathbf{l} = H_\varphi 2\pi r = I,$$

whence

$$H_\varphi = I/(2\pi r) \qquad (38.54)$$

is the magnetic field strength directed along the tangent to the circle. The lines of force are circles concentric with the current.

The magnetic induction is given by

$$B_\varphi = \begin{cases} \mu H_\varphi = \dfrac{\mu I}{2\pi r} & (0 < r < a), \\ \mu_0 H_\varphi = \dfrac{\mu_0 I}{2\pi r} & (a < r). \end{cases} \qquad (38.55)$$

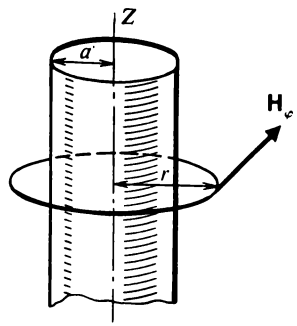


Fig. 156. To the calculation of the magnetic field created by the current flowing in a circular cylinder.

www.EngineeringBooksPDF.com

It is convenient to find the magnetization from relation (38.21):

$$J_\varphi = \begin{cases} \dfrac{\mu-\mu_0}{\mu_0} H_\varphi = \dfrac{\mu-\mu_0}{\mu_0} \dfrac{I}{2\pi r} & (0<r<a) \\ 0 & (a<r). \end{cases} \tag{38.56}$$

The volume density of molecular currents can be found with the help of (38.10). Considering that magnetization is given in cylindrical coordinates, it is convenient to calculate the curl in (38.10) also in cylindrical coordinates. We have

$$\mathbf{j}_m = \operatorname{curl} \mathbf{J} = -\mathbf{i}_r \frac{\partial J_\varphi}{\partial z} + \mathbf{i}_z \frac{1}{r} \frac{\partial}{\partial r} (rJ_\varphi) = 0. \tag{38.57}$$

Thus, there are no volume molecular currents. There is, however, a surface molecular current whose density, on the basis of (38.17) and taking into account (38.56), is given by

$$i_{mz} = -\frac{(\mu-\mu_0) I}{\mu_0 2\pi a}. \tag{38.58}$$

Sec. 39. Forces in a Magnetic Field

The forces acting on currents and body forces acting on incompressible magnetics are considered.

Forces acting on a current.

$$\boxed{d\mathbf{F} = \mathbf{j} \times \mathbf{B}\, dV = I\, d\mathbf{l} \times \mathbf{B},} \tag{39.1a}$$

$$\boxed{\mathbf{F} = \int_V \mathbf{j} \times \mathbf{B}\, dV = \int_L I\, d\mathbf{l} \times \mathbf{B}.} \tag{39.1b}$$

Lorentz's force. The force acting on a point charge q moving with a velocity $\mathbf{v}$ is

$$\boxed{\mathbf{F} = q\mathbf{v} \times \mathbf{B},} \tag{39.2}$$

where q includes the sign of the charge, i.e. can be a positive or negative quantity. Formula (39.2) can be obtained from (39.1b) if we take into account that $\mathbf{j} = nq\mathbf{v}\, dV = \rho \mathbf{v}\, dV$, where ρ is the volume charge density, and hence $\rho\, dV$ is the charge contained in the volume dV, while $\int_V \rho\, dV = q$.

The force and the torque acting on magnetic dipole. Suppose that an elementary circular current creating a magnetic moment flows along a square loop with

www.EngineeringBooksPDF.com

side l. We place the origin of coordinates at the centre of the square and direct the Z-axis perpendicularly to the plane of the loop (Fig. 157). The direction of current I in the loop is shown by arrows. The magnetic field is arbitrary, and there are no extraneous currents or ferromagnetics in the region of the loop (div $\mathbf{B} = 0$, curl $\mathbf{B} = 0$). Let us determine the force and the moment of force (torque) acting on the current-carrying loop in terms of its magnetic moment. The loop is small in size, and we have to take into account the variation of the magnetic induction within the loop only up to the first-order terms in l.

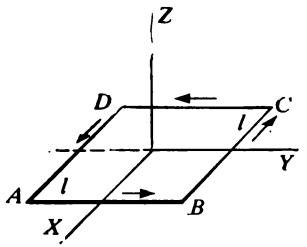


Fig. 157. To the calculation of the force acting on a magnetic dipole.

In accordance with formula (39.1a), the forces exerted on the sides AB, BC, CD, DA of the loop by the magnetic field are

$$\mathbf{F}_{AB} = Il\mathbf{i}_y \times \mathbf{B}(\mathbf{i}_x l/2), \quad \mathbf{F}_{BC} = Il\,[-\mathbf{i}_x \times \mathbf{B}(\mathbf{i}_y l/2)],$$
$$\mathbf{F}_{CD} = Il\,[-\mathbf{i}_y \times \mathbf{B}(-\mathbf{i}_x l/2)], \quad \mathbf{F}_{DA} = Il\,[\mathbf{i}_x \times \mathbf{B}(-\mathbf{i}_y l/2)],$$

where $\mathbf{i}_x$ and $\mathbf{i}_y$ are the unit vectors in the direction of the X- and Y-axes. The arguments of $\mathbf{B}$ contain distances from the centre of the loop to the corresponding side considering the direction. The total force acting on the loop is

$$\mathbf{F} = \mathbf{F}_{AB} + \mathbf{F}_{BC} + \mathbf{F}_{CD} + \mathbf{F}_{DA} = Il\mathbf{i}_y \times [\mathbf{B}(\mathbf{i}_x l/2 - \mathbf{B}(-\mathbf{i}_x l/2)] + Il\mathbf{i}_x \times [\mathbf{B}(-\mathbf{i}_y l/2) - \mathbf{B}(\mathbf{i}_y l/2)]. \tag{39.3}$$

Considering that when only the first-order terms are retained, we have

$$\mathbf{B}\left(\pm \frac{\mathbf{i}_x l}{2}\right) = \mathbf{B}(0) \pm \frac{l}{2}\frac{\partial \mathbf{B}(0)}{\partial x} \quad \text{and} \quad \mathbf{B}(\pm \mathbf{i}_y l/2) = \mathbf{B}(0) \pm \frac{l}{2}\frac{\partial \mathbf{B}(0)}{\partial y}.$$

Equation (39.3) can be transformed as follows:

$$\mathbf{F} = Il^2\left(\mathbf{i}_y \times \frac{\partial \mathbf{B}}{\partial x} - \mathbf{i}_x \times \frac{\partial \mathbf{B}}{\partial y}\right). \tag{39.4}$$

Since $Il^2 = p_m$ is the magnitude of the magnetic moment of a current-carrying loop, we transform (39.4) by taking into account the well-known relations between the unit vectors ($\mathbf{i}_x \times \mathbf{i}_y = \mathbf{i}_z$, $\mathbf{i}_y \times \mathbf{i}_z = \mathbf{i}_x$, $\mathbf{i}_z \times \mathbf{i}_x = \mathbf{i}_y$) as follows:

$$\mathbf{F} = (\mathbf{p}_m \times \mathbf{i}_x) \times \frac{\partial \mathbf{B}}{\partial x} + (\mathbf{p}_m \times \mathbf{i}_y) \times \frac{\partial \mathbf{B}}{\partial y},$$

where $\mathbf{p}_m = \mathbf{i}_z p_m$ is the magnetic moment of the loop. Representing the double vector product with the help of the formula $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ of vector algebra, we obtain

$$\begin{aligned}\mathbf{F} &= \mathbf{i}_x\left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial x}\right) - \mathbf{p}_m\left(\mathbf{i}_x \cdot \frac{\partial \mathbf{B}}{\partial x} + \mathbf{i}_y\left(\mathbf{p} \cdot \frac{\partial \mathbf{B}}{\partial y}\right) - \mathbf{p}_m\left(\mathbf{i}_y \cdot \frac{\partial \mathbf{B}}{\partial y}\right)\right. \\ &= \mathbf{i}_x\left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial x}\right) + \mathbf{i}_y\left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial y}\right) - \mathbf{p}_m\left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y}\right),\end{aligned} \tag{39.5}$$

www.EngineeringBooksPDF.com

where $\mathbf{i}_x \cdot (\partial \mathbf{B}/\partial x) = \partial B_x/\partial x$, $\mathbf{i}_y \cdot (\partial \mathbf{B}/\partial y) = \partial B_y/\partial y$. Since $\operatorname{div} \mathbf{B} = \partial B_x/\partial x + \partial B_y/\partial y + \partial B_z/\partial z = 0$, we get

$$-\mathbf{p}_m \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} \right) = \mathbf{p}_m \frac{\partial B_z}{\partial z} = \mathbf{p}_m \left(\mathbf{i}_z \cdot \frac{\partial \mathbf{B}}{\partial z} \right) = \mathbf{i}_z \left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial z} \right),$$

whence

$$\boxed{\mathbf{F} = \mathbf{i}_x \left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial x} \right) + \mathbf{i}_y \left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial y} \right) + \mathbf{i}_z \left(\mathbf{p}_m \cdot \frac{\partial \mathbf{B}}{\partial z} \right).} \tag{39.6}$$

This formula shows that *an elementary circular current (magnetic dipole) is acted upon by a force only in a nonuniform field.* Since formula (39.6) expresses this force in terms of the magnetic moment $\mathbf{p}_m$, the special shape of the loop chosen above is insignificant, and formula (39.6) *is valid for an arbitrary magnetic dipole whose spatial dimensions are sufficiently small.*

In order to calculate the moment acting on a magnetic dipole, we proceed in a similar way. We place the origin of coordinates at the centre of the loop and calculate the moment by the formula

$$\mathbf{M} = I \int_L \mathbf{r} \times (d\mathbf{l} \times \mathbf{B}). \tag{39.7}$$

In this case, however, calculations are simplified since the distance r is of the order of magnitude of sides l of the loop, and the quantity $\mathbf{B}$ should be taken into account only in the zero order of smallness in the side l of the loop, i.e. considered to be constant. As a result, we have

$$\boxed{\mathbf{M} = \mathbf{p}_m \times \mathbf{B}.} \tag{39.8}$$

This formula shows that *the torque tends to rotate the magnetic dipole until its magnetic moment coincides with the vector of the magnetic induction of the field.*

Body forces acting on incompressible magnetics. Since the magnetic moment of a volume element dV of a magnetic having magnetization $\mathbf{J}$ is given by

$$d\mathbf{p}_m = \mathbf{J}\, dV, \tag{39.9}$$

the force acting on this volume element is [see (39.6)]

$$dF_x = \mathbf{J} \cdot \frac{\partial \mathbf{B}}{\partial x} dV, \quad dF_y = \mathbf{J} \cdot \frac{\partial \mathbf{B}}{\partial y} dV, \quad dF_z = \mathbf{J} \cdot \frac{\partial \mathbf{B}}{\partial z} dV. \tag{39.10}$$

Obviously, these expressions are always valid for rigid magnetics since formula (39.6) was obtained as a result of differentiation at $\mathbf{p}_m = \text{const}$.

Let us represent (39.10) in vector form. Considering that

$$\mathbf{J} = \frac{\mu - \mu_0}{\mu \mu_0} \mathbf{B}, \tag{39.11}$$

we obtain the following expression for the volume density of force:

$$f_x = \frac{dF_x}{dV} = \frac{\mu - \mu_0}{\mu \mu_0} \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial x} = \frac{1}{2} \frac{\mu - \mu_0}{\mu \mu_0} \frac{\partial B^2}{\partial x} \tag{39.12}$$

www.EngineeringBooksPDF.com

and so on. Thus, the volume density of the force acting on a magnetic is

$$\mathbf{f} = \frac{1}{2} \frac{\mu - \mu_0}{\mu \mu_0} \operatorname{grad} B^2. \tag{39.13}$$

This means that

(a) *in paramagnetics* $\mu > \mu_0$ *and hence the volume density of force is directed towards increasing magnetic induction.*

(b) *in diamagnetics* $\mu < \mu_0$ *and hence the volume density of force is directed towards decreasing magnetic induction of the field.*

Different behaviour of para- and diamagnetics in the same field is visually demonstrated in many experiments. Suppose that a magnetic field is created

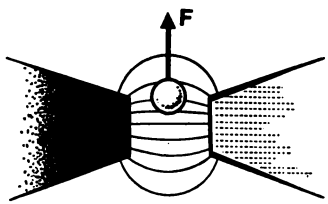


Fig. 158. A diamagnetic body pushed out of the region of the maximum field

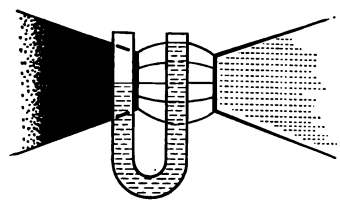


Fig. 159. A paramagnetic body drawn towards the region of the maximum field

in vacuum between the poles of a permanent magnet (Fig. 158). Obviously, the magnetic induction of the field between the poles of the magnet decreases as we move from the central line connecting the poles to the periphery. A light bismuth ball which is a diamagnetic body is pushed out of the region with the maximum field induction (see Fig. 158). On the other hand, a paramagnetic liquid, say, aqueous solution of ferric chloride is drawn towards the regions of the field having the maximum magnetic induction (Fig. 159).

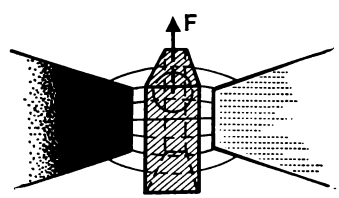


Fig. 160. A paramagnetic body behaves as a diamagnetic in a paramagnetic medium whose permeability is higher than that of the body

If the space between the poles is filled with a material medium, the direction of forces depends on the ratio of permeabilities of the medium and the body. *If the permeability of the body is higher than that of the medium, the body behaves as a paramagnetic, if it is lower, then the body behaves as a diamagnetic.* For example, if we place a paramagnetic liquid with a sufficiently high permeability between the poles of a magnet (Fig. 160), the force acting on a paramagnetic ball whose permeability is lower than that of the liquid is similar to that acting on a diamagnetic ball in vacuum.

A force acts on an elementary circular current (magnetic dipole) only in a nonuniform magnetic field.

The torque appearing as a result of the action of a magnetic field on a magnetic dipole tends to rotate it so that its magnetic moment coincides in direction with the magnetic induction vector.

www.EngineeringBooksPDF.com

Body forces acting on a paramagnetic are directed towards the increasing magnetic induction, while in diamagnetics they are directed towards decreasing magnetic induction.

Describe the variation of forces acting on a magnetic placed in a medium whose permeability differs from the magnetic constant and becomes higher or lower than the permeability of the magnetic.

Example 39.1. *A current I flows in a ring of radius r_0 made of a very thin wire. The tensile strength of the wire is equal to f_0. The ring is placed in a magnetic field whose magnetic induction is perpendicular to the plane of the ring so that the forces tend to break the ring. Find the magnetic induction at which the ring will be broken, assuming that $f_0 = 1.5$ N, $r_0 = 15$ cm and $I = 10$ A.*

The forces act on the ring in the radial directions. Denoting by dl an element of length of the ring, we find that the element of force acting on the element dl in the radial direction is $d\mathbf{F} = I\, d\mathbf{l} \times \mathbf{B}$. Let us draw the X-axis through the centre of the ring so that it lies in the plane of the ring. The projection of the force dF onto the X-axis is $dF_x = dF \cos\alpha = IB\, dl \cos\alpha$, where α is the angle between the X-axis and the radius drawn to the element dl.

Since $dl = r_0\, d\alpha$, the expression for the force acting on a half-ring in the direction of positive values of the X-axis has the form $F_x = IBr_0 \int_{-\pi/2}^{\pi/2} \cos\alpha\, d\alpha = 2IBr_0$. This force is distributed between two wire cross sections at the points of its intersection with the Y-axis. Hence the breaking condition has the form $2IBr_0 = 2f_0$, and consequently, $B = f_0/(Ir_0) = 1$ T.

Problems

6.1. A copper spiral of radius a has n turns per metre. The turns are wound so that there are small gaps between them. The upper end of the spiral is fixed while its lower end is connected to a conducting load of mass m, lying on a metallic table. In this position, elastic forces exerted on the load by the spiral are equal to zero. Assuming that the gaps between the spiral turns decrease uniformly, find the current that should be passed through the spiral in order to lift the load from the table. The mass of the spiral can be ignored.

6.2. Two small magnets having the same magnetic moment p_m and mass m are suspended from light threads of the same length. The distance d between the points of suspension is very large. Prove that the magnets get oriented so that they will attract each other. Find the angle of deviation of the threads from the vertical, assuming that the effect of the magnetic field of the Earth can be neglected.

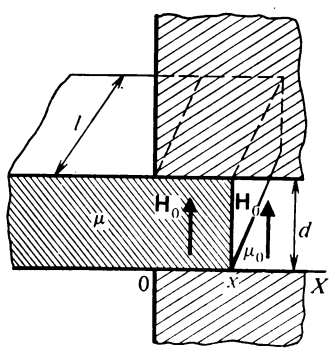


Fig. 161. To the calculation of forces of interaction between magnets.

6.3. A sphere of radius a, uniformly charged with the surface charge density σ, rotates around the axis passing through its centre at an angular velocity ω. Find the magnetic induction at the centre of the rotating sphere.

6.4. Find the magnetic moment created by a point charge q moving in a circle of radius r_0 at a constant angular velocity ω.

6.5. A plate made of a magnetic with a permeability μ is placed in the space between the poles of a permanent magnet where the magnetic field is H_0 (Fig. 161). Find the force acting on the magnetic plate.

6.6. Find the force in Problem 6.5 assuming that the plate is a permanent magnet whose magnetization J_p coincides in direction with H_0.

6.7. Find the force with which a uniform surface current of density i_{sur} flowing over an infinite plane acts over length l of an infinite linear conductor

www.EngineeringBooksPDF.com

carrying current I parallel to the plane at a distance d from it. Denote by $\mathbf{n}$ the normal to the plane in the direction towards the linear conductor.

6.8. A current I_1 flows in a circular conductor of radius a, lying in the plane (x, y) with the centre at the origin of coordinates and forms a right-hand screw system with the positive direction of the Z-axis. A current I_2 flows in very long straight conductor parallel to the X-axis in the direction of its positive values, intersecting the Z-axis at a point $z = d$. Find the force acting on the rectilinear current.

6.9. Find the magnetic induction at the centre of a solenoid of length L with n turns, having a square cross section with side a. The current in the solenoid winding is equal to I.

6.10. A disc of radius r rotates at an angular velocity ω about the axis perpendicular to its surface and passing through its centre. Find the magnetic induction of the field on the axis of rotation of the disc at a distance h from its surface. The surface charge density is equal to σ.

6.11. A polarized dielectric sphere of radius a rotates at an angular velocity ω about an axis passing through its centre. The polarization $\mathbf{P}$ is constant and coincides in direction with ω. Find the magnetic induction at the points of intersection of the spherical surface with the axis of rotation.

6.12. A very long rectilinear cylindrical beam of circular cross section of radius a, having a constant volume charge density ρ, moves along its axis at a velocity $\mathbf{v}$. Find the magnetic induction.

6.13. A current I flows in the positive direction of the Z-axis along a very long rectilinear cylindrical conductor of radius a whose axis coincides with the Z-axis of the Cartesian system of coordinates. Find the vector potential of the field created by this current.

6.14. Find the axial component of the vector potential at the centre of a spiral carrying current I. The data on the spiral are given in Problem 1.7.

6.15. A dielectric sphere of radius a rotates at an angular velocity ω about an axis passing through its centre. A constant volume charge density of the sphere is equal to ρ. Find the magnetic induction of the sphere on its axis of rotation.

6.16. A uniformly charged circular cylinder of radius a and length l, whose charge is equal to Q, rotates at an angular velocity ω about its axis. Find its magnetic dipole moment.

6.17. Find the mutual inductance of two circular currents of radii a_1 and a_2, lying in the same plane, in the dipole approximation. The distance between the turns is r.

6.18. The axis of a right circular cylinder coincides with the Z-axis of the Cartesian system of coordinates whose origin is at the centre of the cylinder. The cylinder is uniformly magnetized and the magnetization vector coincides with the positive direction of the Z-axis: $\mathbf{J} = J\mathbf{i}_z$. Find the magnetic induction on the axis of the cylinder if the radius of its cross section is a and the length is l.

6.19. A spherical layer of a magnetic, whose inner and outer concentric surfaces have radii r_1 and r_2 respectively, is uniformly magnetized. The magnetization vector is parallel to the Z-axis of the Cartesian system of coordinates whose origin coincides with the centre of the surfaces and is equal to $J\mathbf{i}_z$. Find the magnetic field strength on the Z-axis for positive values of z.

6.20. A right circular cylinder of length l and radius a is uniformly magnetized. The magnetization vector is parallel to the cylinder axis and is equal to $\mathbf{J}$. Find the magnetic induction at the centre of the cylinder, assuming that $l \gg a$.

6.21. A sphere with the surface charge density σ rotates about its diameter at an angular velocity ω. Find its magnetic dipole moment.

6.22. A current I flows along a very long rectilinear conductor parallel to the plane interface between a medium having permeability μ_0, in which the current-carrying conductor lies, and a medium having permeability μ. Find the force acting on a segment l of the conductor. The distance from the conductor to the interface is equal to d.

6.23. A thin wire is wound very tightly in one layer on the surface of a wooden sphere. The planes of all the turns can be assumed to be perpendicular to the same diameter of the sphere. The turns cover the entire surface of the sphere. The radius of the sphere is a and the total number of turns is n. The current in the winding is I. Find the magnetic induction at the centre of the sphere.

www.EngineeringBooksPDF.com

6.24. A cylindrical conductor of radius a has a cylindrical cavity of radius b, whose axis is parallel to the axis of the conductor and is at a distance d from it. The current in the conductor has the volume density j. Find the magnetic induction at the points on the diameter of the cavity, coinciding with the diameter of the conductor.

Answers

6.1. $I=\frac{1}{na}\sqrt{\frac{2mg}{\pi\mu_0}}$. 6.2. $\theta=\frac{3}{2}\frac{p_m^2}{\pi\mu_0 d^4}\frac{1}{mg}$. 6.3. $\mathbf{B}=2/3\ \mu_0\sigma a\omega$. 6.4. $p_m=q\omega r_0^2/2$. 6.5. $F_x=1/2\ (\mu-\mu_0)\ H_0^2 ld$. 6.6. $F_x=\mu_0 J_n\ (H_0+J_n)\ ld$. 6.7. $\mathbf{F}=-1/2\ \mu_0 i_{surf} I\mathbf{n}l$. 6.8. $\mathbf{F}=i_y\mu_0 I_1 I_2\ (1-d/\sqrt{d^2+a^2})$. 6.9. $B=\mu_0 nI\left(1-\frac{2}{\pi}\arcsin\frac{a^2}{L^2+a^2}\right)$. 6.10. $B_h=\sigma\omega\left(\frac{h^2+a^2/2}{\sqrt{h^2+a^2}}-h\right)$. 6.11. $\mathbf{B}_1=2/5\ \mu_0 Pa\omega$, $\mathbf{B}_2=-2/5\ \mu_0 Pa\omega$. 6.12. $\mathbf{B}=1/2\ \mu_0\rho\mathbf{v}\times\mathbf{r}$ for $0<r<a$, $\mathbf{B}=1/2\ \mu_0\rho a^2\mathbf{v}\times\mathbf{r}/r^2$ for $a<r<\infty$. 6.13. $A_z=-\frac{\mu_0 I}{4\pi}\frac{r^2}{a^2}+\text{const}$ for $r<a$, $A_z=\frac{\mu_0 I}{2\pi}\ln r+\text{const}$ for $a<r<\infty$, where $r=\sqrt{x^2+y^2}$. 6.14. $\frac{\mu_0 I}{2\pi}\ln(n\pi\tan\alpha+\sqrt{1+\pi^2 n^2\tan^2\alpha})$. 6.15. 0. 6.16. $Qa^2\omega/4$. 6.17. $L_{12}=\pi\mu_0 a_1^2 a_2^2/(4r^3)$. 6.18. $B_z=\frac{\mu_0}{2}J\left(\frac{z+l/2}{\sqrt{a^2+(z+l/2)^2}}-\frac{z-l/2}{\sqrt{a^2+(z-l/2)^2}}\right)$. 6.19. $H_z=0$ for $0<z<r_1$, $H_z=\frac{J(z^3+2r_1^3)}{3z^3}$ for $r_1<z<r_2$, $H_z=2J(r_2^3-r_1^3)/(3z^3)$ for $r_3<z<\infty$. 6.20. $\mathbf{B}=\mu_0\mathbf{J}(1-a^2/l^2)$. 6.21. $p_m=4/3\pi\sigma a^4\omega$. 6.22. $F=-\frac{\mu_0 l}{4\pi d}\frac{\mu-\mu_0}{\mu+\mu_0}I^2$. 6.23. $\mu_0 nI/(4a)$. 6.24. $\mu_0 jd/2$.

www.EngineeringBooksPDF.com

CHAPTER 7

Magnetics

The phenomenological properties of a magnetic in a magnetic field are taken into account through its permeability μ. The dependences of μ on various parameters are very diversified as well as magnetics themselves. These dependences are interpreted by constructing models of magnetics which allow for the peculiarities of their behaviour in the magnetic field.

Sec. 40. Diamagnetics

The physical nature of diamagnetic susceptibility and its properties are considered.

Larmor precession. The frequency of rotation of electrons in an atom placed in a magnetic field differs from their frequency in the absence of the field. In order to prove this, let us consider the simplest case when an electron rotates around a nucleus in an orbit of radius r in the absence of the field at a frequency ω_0 (Fig. 162). Newton's equation for the motion of an electron has the form

$$m\omega_0^2 r = F_c, \tag{40.1}$$

where F_c is the centripetal force emerging as a result of attraction of the electron by the nucleus. This force is quite strong in comparison with the forces which may be exerted on the electron by external fields, and hence *the radii of electron orbits do not change when an atom is introduced into external fields. To a high degree of accuracy, an atom can be assumed to be rigid with respect to the action of external fields.*

Suppose now that an atom is in an external field whose magnetic induction vector **B** is perpendicular to the plane of an electron orbit. The Lorentz force is acting along the radius, and its direction either coincides with that of the centripetal force or is opposite to it, depending on the relative orientation of the angular velocity vectors of electron motion in the orbit and of the magnetic induction. The magnitude of this force is

$$F = |e|\,\omega r B, \tag{40.2}$$

where e is the electron charge and ω is the frequency of the electron circulating in the orbit in the presence of a magnetic field, which differs from ω_0.

www.EngineeringBooksPDF.com

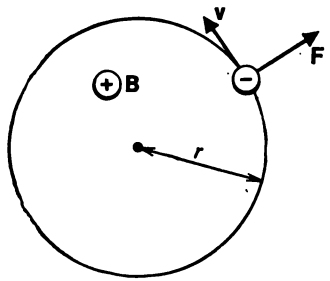


Fig. 162. Emergence of an additional angular velocity of an ron rotating in a magnetic field

The equation of the electron motion in a magnetic field has the form

$$m\omega^2 r = F_c \pm |e| \omega r B, \qquad (40.3)$$

where the radius r of the electron orbit is the same as in (40.1) while the plus or minus sign is chosen in accordance with relative orientation of the angular velocity vector of the electron orbital motion and the magnetic induction vector. Naturally, the centripetal force F_e in (40.3) is the same as in (40.1) since it is the force of attraction by the nucleus, and the distance r has remained unchanged. Cancelling F_c from (40.1) and (40.3), we obtain

$$m\omega^2 r - m\omega_0^2 r = \pm |e| \omega r B. \qquad (40.4)$$

Considering that $\omega^2 - \omega_0^2 = (\omega - \omega_0)(\omega + \omega_0) \approx 2\Delta\omega\omega$, where $|\Delta\omega| = |\omega - \omega_0| \ll \omega_0$, we get from (40.4)

$$\Delta\omega = \pm |e| B/(2m). \qquad (40.5)$$

Thus, an electron in a magnetic field acquires an additional angular velocity characterized by the frequency

$$\omega_L = |e| B/(2m). \qquad (40.6)$$

which is called the **Larmor frequency.** The direction of the angular velocity vector can be easily determined. For example, if the magnetic induction **B** (see Fig. 162) is directed against the angular velocity of the electron motion around the nucleus, the force F is directed against F_c, and hence the electron velocity and frequency should decrease. This means that ω_L coincides with **B** in direction. If **B** is directed against the initial orientation of the magnetic induction, we shall arrive at the same conclusion. Hence we can write

$$\boxed{\omega_L = -e\mathbf{B}/(2m),} \qquad (40.7)$$

where we took into consideration that the electron charge e is negative. *The appearance of this additional angular velocity without a change in the radius of the orbit can be represented as an additional rotation of the atom as a whole at a frequency* ω_L *in the magnetic field.* The total frequency of electron rotation is equal to the sum of its frequency ω_0 of rotation in the atom and the frequency ω_L of rotation of the atom. This is valid only for the case when the angular velocity and magnetic induction vectors are collinear.

Since the electron velocity in an atom placed in a magnetic field varies, its kinetic energy varies as well. On the other hand, since r remains unchanged, the potential energy also does not change. The question arises: what is the cause of the change in the energy of an electron in an atom if it is known that the magnetic field is always acting perpendicularly to the velocity and does not

www.EngineeringBooksPDF.com

perform any work? The answer to this question can be obtained only in the framework of the theory of electromagnetic induction (see Chap. 8): *the appearance of a magnetic field gives rise to an electric field which changes the velocity of the electron motion in an atom.*

In order to imagine the motion of an atom for an arbitrary mutual orientation of the angular velocity of rotation of an electron around the nucleus and the magnetic induction of the external field, let us extend the results obtained above to the general case. An atom with an electron rotating in it can be visualized

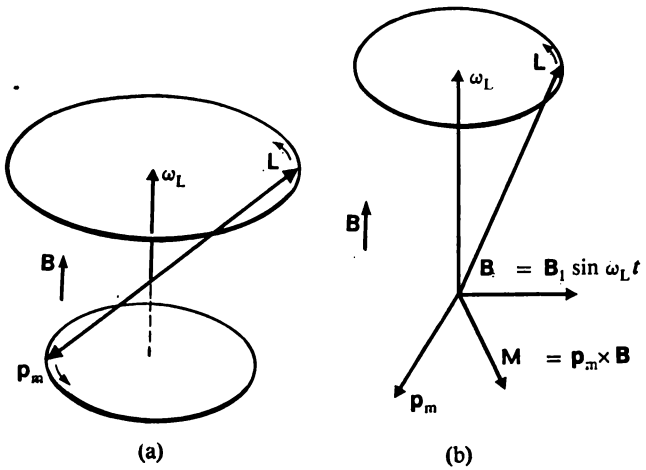


Fig. 163. Larmor precession (a) and the emergence of paramagnetic resonance (b)

as a gyroscope having a certain magnetic moment. The angular momentum of the electron is equal to $m\omega r^2$. The electron moving in the orbit is equivalent to a circular current $e/T = e\omega/(2\pi)$, and hence the magnetic moment of the atom is equal to $\pi r^2 e\omega/(2\pi)$. Taking into account the directions of the angular momentum and the magnetic moment of the atom due to the motion of the electron, we can write

$$\mathbf{L} = mr^2\boldsymbol{\omega}, \quad \mathbf{p}_m = (er^2/2)\,\boldsymbol{\omega}. \tag{40.8}$$

Here we assumed that the electron charge e is negative, and the angular momentum $\mathbf{L}$ and the magnetic moment $\mathbf{p}_m$ have opposite directions (Fig. 163*a*).

The equation of motion of an atom treated as a gyroscope has the form

$$\frac{d\mathbf{L}}{dt} = \mathbf{M}, \tag{40.9}$$

where $\mathbf{M}$ is the torque [see (39.8)]. It follows from (40.8) that

$$\mathbf{p}_m = e\mathbf{L}/(2m) \tag{40.10}$$

www.EngineeringBooksPDF.com

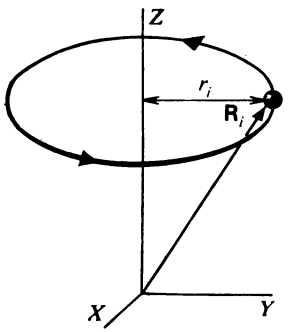


Fig. 164. To the calculation of diamagnetic susceptibility

and consequently, Eq. (41.9) assumes the form

$$\frac{d\mathbf{L}}{dt} = \frac{e}{2m}\mathbf{L}\times\mathbf{B} = -\frac{e}{2m}\mathbf{B}\times\mathbf{L}. \qquad (40.11)$$

A comparison of this equation with the equation of motion for the points of a perfectly rigid body rotating at an angular velocity

$$\mathbf{v} = d\mathbf{r}/dt = \boldsymbol{\omega}\times\mathbf{r} \qquad (40.12)$$

shows that the tip of the vector **L** circulates around the direction of the induction vector at a frequency

$$\omega_L = -e\mathbf{B}/(2m). \qquad (40.13)$$

Consequently, an atom precesses in a magnetic field like a gyroscope (Fig. 163*b*). This motion is called the **Larmor precession.**

Diamagnetism. A circular current appearing as a result of Larmor precession of each electron in an atom forms a left-handed system with the direction of the magnetic induction vector. Consequently, an additional magnetic induction due to this circular current is directed against the magnetic induction vector of the external field. *The magnetic moment of the atom, appearing as a result of precession, as well as the magnetization are directed against the magnetic induction of the external field.* This mode of origination of the Larmor precession and the magnetic moment and the additional magnetic field associated with it form the basis of **diamagnetism**. Obviously, diamagnetism is inherent in any substance. The problem consists in estimating its magnitude.

Diamagnetic susceptibility. Each electron in an atom performs Larmor rotation about the axis coinciding with the direction of the magnetic field (Fig. 164). The resulting magnetic moment is

$$p_{mi} = S_i I_i = \pi r_i^2 e/T = e r_i^2 \omega_L/2, \qquad (40.14)$$

whence

$$J = \frac{1}{\Delta V}\sum_{\Delta V} p_{mi} = -\frac{e^2}{4m} BN\left\langle \sum_i r_i^2 \right\rangle, \qquad (40.15)$$

where N is the atomic concentration. Formula (40.15) uses the expression for the Larmor frequency, and the quantity in the angle brackets which denote averaging is the sum of the squares of the distances from the electrons in the atom to the Larmor precession axis. Figure 164 shows that

$$R_i^2 = x_i^2 + y_i^2 + z_i^2, \qquad (40.16)$$

where R_i is the distance between an electron and the nucleus. Taking into account random orientation of atoms in space, we have

$$\langle x_i^2\rangle = \langle y_i^2\rangle = \langle z_i^2\rangle = \langle R_i^2\rangle/3 \qquad (40.17)$$

www.EngineeringBooksPDF.com

and hence

$$\langle r_i^2 \rangle = \langle x_i^2 + y_i^2 \rangle = 2 \langle R_i^2 \rangle / 3 = 2 \langle R^2 \rangle / 3, \tag{40.18}$$

from which it follows that

$$\langle \sum_i r_i^2 \rangle = 2Z \langle R^2 \rangle / 3, \tag{40.19}$$

where Z is the number of electrons in the atom. Hence, we finally obtain the following formula for magnetization:

$$J = -\frac{e^2}{6m} NZ \langle R^2 \rangle \mu H. \tag{40.20}$$

Comparing this formula with

$$J = \chi_d H, \tag{40.21}$$

we obtain the following expression for the diamagnetic susceptibility:

$$\chi_d = -\frac{e^2}{6m} NZ \langle R^2 \rangle \mu_0, \tag{40.22}$$

where we took into account that $\mu \approx \mu_0$ since the permeability of diamagnetics differs from the permeability of vacuum only insignificantly. This formula is in good agreement with experiment if we treat $\langle R^2 \rangle$ as the mean square distance between the electrons and the nucleus in an atom, calculated with the help of quantum theory. The diamagnetic susceptibility for solids and liquids is of the order of 10^{-5}, while for gases it is considerably lower due to a smaller atomic concentration (i.e. smaller values of N in formula (40.22)).

Temperature independence of diamagnetic susceptibility. Formula (40.22) shows that χ_d *is independent of temperature* since none of the quantities appearing in this formula depend on temperature. This is due to the fact that the Larmor motion of electrons stabilizes very quickly, viz. during a time typical of atomic processes. Consequently, *thermal motion as well as atomic collisions do not bring the atoms out of the Larmor precession for any appreciable periods of time.* This was brilliantly confirmed by experiments. The independence of diamagnetic susceptibility from temperature was experimentally discovered in 1895 by P. Curie (1859-1906).

The variatoin of the frequency of an electron circulating in an atom, which is responsible for diamagnetism, occurs upon a change in the magnetic induction due to the introduction of the atom into a magnetic field or during the emergence of a magnetic field. The magnetic field itself does not perform work and cannot alter the velocity of motion of electrons in an atom.

The diamagnetic susceptibility is independent of temperature since the thermal motion and collisions of atoms are unable to draw them from the state of Larmor precession for any appreciable time.

www.EngineeringBooksPDF.com

Sec. 41. Paramagnetics

The physical nature and properties of paramagnetic susceptibility are discussed. Magnetism due to free electrons and the paramagnetic resonance are also discussed.

Mechanism of magnetization. Paramagnetic materials are substances whose molecules have a constant magnetic moment. The energy corresponding to the magnetic moment in an external magnetic field is equal to

$$W = -\mathbf{p}_m \cdot \mathbf{B}. \tag{41.1}$$

The minimum value of energy is attained when $\mathbf{p}_m$ coincides with the direction of the magnetic induction vector. In this case, when a paramagnetic is introduced into a magnetic field, a preferred orientation of magnetic moments of paramagnetic atoms takes place in the direction of the magnetic induction in accordance with the Boltzmann distribution, and the body is accordingly magnetized. The induction of the additional field created due to magnetization coincides in direction with the external magnetic induction and enhances it. However, *the angle between the direction of the magnetic moment of an atom and the magnetic induction does not change with field: the magnetic moment just precesses around the magnetic induction vector, and the angle between the two remains the same* [see (40.11)]. *A reorientation of magnetic moments in accordance with the Boltzmann distribution takes place only as a result of collisions and interactions between atoms.*

Temperature dependence of paramagnetic susceptibility. The mechanism of magnetization of paramagnetics is analogous to that of electrostatic charging in polar dielectrics (see Sec. 22). The only difference is that we now use formula (41.1) instead of (22.1). Thus, the formulas for paramagnetic susceptibility are obtained by substituting $\mathbf{p}_m$ for $\mathbf{p}$ and $\mathbf{B}$ for $\mathbf{E}$ in the formulas of Sec. 22 pertaining to dielectric susceptibility.

Instead of (22.10), we now get

$$\langle p_{mz} \rangle = p_m L(\beta), \tag{41.2}$$

where $L(\beta)$ is the Langevin function (see Sec. 22) for $\beta = p_m B/(kT)$. At comparatively high temperatures and weak fields, when $p_m B \ll kT$, i.e. $\beta \ll 1$, we obtain the following relation instead of (22.13):

$$\langle p_{mz} \rangle = p_m^2 B/(3kT) \approx p_m^2 \mu_0 H/(3kT), \tag{41.3}$$

where $\mu \simeq \mu_0$, since the difference between the permeability of paramagnetics and μ_0 is not significant. For magnetization, we obtain the formula

$$J = N \langle p_{mz} \rangle = [p_m^2 N \mu_0/(3kT)]\, H. \tag{41.4}$$

www.EngineeringBooksPDF.com

Comparing this formula with the relation

$$J = \chi_p H, \tag{41.5}$$

we get the following expression for paramagnetic susceptibility:

$$\chi_p = p_m^2 N \mu_0 / (3kT) = C/T, \tag{41.6}$$

where C is the Curie constant.

The dependence $\chi_p \propto 1/T$ is called **Curie's law** which was discovered experimentally by P. Curie in 1896.

The atomic magnetic moments are of the order $p_m \sim 10^{-23}$ A·m². Hence, at room temperature, $\chi_p \sim 10^{-3}$, a value which is two orders of magnitude higher than the diamagnetic susceptibility. This means that *the diamagnetic susceptibility in paramagnetics can be generally neglected.*

Langevin's theory provides a fairly accurate description only for gases, where the interaction between molecules is negligibly small in view of large distances that separated them. This interaction can be significant in liquids and solids. In most cases, a consideration of this interaction modifies the dependence (41.6) of susceptibility on temperature. The new dependence is called the **Curie-Weiss law:**

$$\chi_p = \text{const}/(T - T_0), \tag{41.7}$$

where the temperature T_0 is a characteristic of the substance and is determined by its properties.

Magnetic moments of free atoms. Two factors are responsible for the origin of magnetic moments:

(1) the orbital motion of electrons. The total orbital magnetic moment of an atom is the sum of orbital magnetic moments of individual electrons;

(2) the existence of an intrinsic magnetic moment in electrons, which is associated with their spin, i.e. with the intrinsic angular momenta of electrons.

The magnetic moments of individual electrons are mutually related and form the so-called total spin magnetic moment of an atom. Owing to the spin magnetic moment, each electron moving in a magnetic field created by the orbital motion of all the remaining electrons interacts with this field. This interaction is called **spin-orbital interaction.** *Because of this interaction, the total orbital angular momentum of electrons is associated with their total spin magnetic moment, thus forming the total magnetic moment of an atom. This mode of formation of the total magnetic moment of an atom is called an LS-bond.* In principle, the total magnetic moment of an atom can emerge in another way also: *the spin magnetic moment of each electron is first linked with the orbital magnetic moment of the same electron, thus forming the total magnetic moment of the electron. After this, the total magnetic moments of electrons are linked with one another to give the total magnetic moment of the atom.* With the exception of the heaviest elements, however, this mode of formation of total magnetic moment is generally not realized since the intensity of interaction of the spin magnetic moment of an electron with its intrinsic orbital angular momentum turns out to be weaker than its interaction with the spin magnetic moments of other electrons. Thus

www.EngineeringBooksPDF.com

the total magnetic moment for an individual electron is not realized. In most cases, *the LS-bond is responsible for the total magnetic moment of atoms.*

While adding the total orbital magnetic moment and the total spin magnetic moment, we must bear in mind that the coefficient of proportionality in the linear relation between the total orbital magnetic moment and the total orbital angular momentum differs from the coefficient of proportionality in the linear relation between the total spin magnetic moment and the total spin. The total angular momenta in an atom are added according to the rule of summation of vectors, while the addition of magnetic moments is obtained as a result of summation of angular momenta. Consequently, *the total magnetic moment of an atom may not be collinear with its total intrinsic angular momentum.*

The problem of magnetic moments of free atoms is simplified on account of the fact that it is advantageous from the energy point of view to fill the electron orbits so that the total magnetic moment is minimum. Consequently, the total orbital and spin moments of completely filled closed atomic shells and the total moment of the completely filled shells are equal to zero. Hence, *the magnetic moment of an atom is determined only by the electrons which occupy partially filled shells.* In most cases, these are the outer orbits of electrons. The situation is further simplified in view of the fact that electron spins and orbital angular momenta in an outer shell tend to orient themselves in opposite directions so as to compensate each other to the maximum possible extent. Thus, *the magnetic moment of a free atom is mainly determined by uncompensated spins of the outer electrons.*

Magnetic moments of molecules. The magnetic moment of a molecule is not equal to the sum of the magnetic moments of the atoms constituting it, since a chemical bond formed between the atoms requires a certain rearrangement of the outer electron shells. For example, nitrogen molecules N_2 form a covalent bond, and two electrons participating in a covalent bond have antiparallel spins. The orbital angular momenta also compensate each other and their sum is equal to zero. As a result, we see that the N_2 molecules do not have a constant magnetic moment or, in other words, nitrogen is not a paramagnetic. A similar tendency towards mutual compensation of magnetic moments is also observed in molecules having an ionic bond. For example, a NaCl molecule has an ionic bond between Na^+ and Cl^-. Both ions have closed electron shells and hence their total magnetic moment is equal to zero. It can be stated that *the general tendency in the formation of molecules is to ensure that the total magnetic moment is equal to zero.* The only gases with paramagnetic properties among all commonly encountered gases are oxygen (O_2) where the spins of collective electrons are not compensated, and NO and NO_2 in which the total number of electrons is odd and hence the spin of one electron is left uncompensated.

Most solids are composed of ions with closed shells, and hence they do not possess paramagnetic properties. Solids are generally diamagnetics. The main exception to this rule are compounds of transition elements. The electron shell of these elements is only partially filled. Consequently, they are multivalent and their ions have permanent magnetic moments. Thus, *the paramagnetism of transition element compounds is due to the magnetic moments of their ions.* Ions

www.EngineeringBooksPDF.com

whose outer electron orbits have similar configurations form compounds with properties close to those of paramagnetics.

Magnetism due to free electrons. Although free electrons in a magnetic field move in circular orbits under the action of the Lorentz force, the classical theory predicts the absence of the diamagnetic effect due to the reflection of electrons at the boundaries. Quantum theory, on the other hand, predicts the existence of this effect. The diamagnetic susceptibility is found to be equal to

$$\chi_p = -\frac{\mu_0 e^2}{6\pi m^*}\left(\frac{3n}{8\pi}\right)^{1/3}, \tag{41.8}$$

where m^* is the effective mass of free electrons and n is their concentration. If the magnetic induction is not very large, the diamagnetic susceptibility is constant and does not depend on temperature.

Another magnetic effect associated with the conduction electrons is due to the interaction of the spin magnetic moment of an electron with the magnetic field. This results in the appearance of an excess of electrons whose spin magnetic moments are aligned with the magnetic induction in contrast to the electrons with opposite spin magnetic moments. This phenomenon is called the **paramagnetism of conduction electrons.** Calculations show that the paramagnetic susceptibility of conduction electrons under laboratory conditions is practically independent of temperature. Paramagnetism of conduction electrons is most pronounced in transition metals. Under laboratory conditions, the diamagnetic susceptibility of conduction electrons is almost always lower than their paramagnetic susceptibility (nearly by a factor of three), and hence their total susceptibility is found to be positive (paramagnetic).

Paramagnetic resonance. Suppose that in a paramagnetic placed in a magnetic field the induction vector of an additional periodic magnetic field is perpendicular to that of the constant magnetic field. Under the action of the constant magnetic field (Fig. 163*b*), the magnetic moments of atoms perform Larmor precession. A torque $\mathbf{M}$ created as a result of interaction between the magnetic moment $\mathbf{p}_m$ of an atom and the induction $\mathbf{B}$ of the additional periodic magnetic field tends to change the angle between $\mathbf{p}_m$ and $\mathbf{B}$. If the frequency of the periodic magnetic field is different from the Larmor precession frequency, the torque will strive to increase the angle between $\mathbf{p}_m$ and $\mathbf{B}$ for a certain duration of time, and then strive to decrease this angle for some time. On the average, no effect of the periodic field is observed. If, however, the frequency of the periodic magnetic field coincides with the Larmor precession frequency, the torque caused by the periodic magnetic field will strive either to increase the angle between the magnetic moment of the atom and the induction of the constant field or to decrease it all the time, depending on the phase difference between the Larmor precession and the induction of the periodic magnetic field. *As a result of the prolonged action of such a torque, the magnetic moment of the atom is reoriented, and its angle with the induction vector of the permanent magnetic field changes.* This phenomenon is called **paramagnetic resonance.** The reorientation of the magnetic moment in accordance with formula (41.1) is associated with a change in the energy of magnetic moment in a permanent magnetic field. In accordance

www.EngineeringBooksPDF.com

with the law of conservation of energy, this is accompanied by an exchange of energy with the periodic magnetic field. This field is realized in the form of standing electromagnetic waves whose magnetic induction vector is perpendicular to the induction vector of the constant magnetic field. Thus, *energy is exchanged with an electromagnetic wave.*

This results in the formation of groups of atoms whose magnetic moments are oriented parallel and antiparallel to the magnetic induction, i.e. which have different energies of interaction with the magnetic field in accordance with (41.1). The energy of atoms with antiparallel orientation is higher than that of atoms with parallel orientation.

Besides being reoriented by a periodic magnetic field, the magnetic moments of atoms are continuously subjected to thermal fluctuations and interactions between atoms. The thermal motion and the interaction between atoms predominantly orient the magnetic moments in a direction antiparallel to the magnetic induction vector. The energy liberated in this case is converted into heat. The reorientation of magnetic moments parallel to the magnetic induction takes place on account of absorption of energy of an electromagnetic wave. Hence, *the observation of paramagnetic resonance involves the measurement of the intensity of an electromagnetic wave passing through a paramagnetic placed in a magnetic field.* From the experimental point of view, it is more convenient to use an electromagnetic wave of constant frequency and to attain resonance condition by varying the magnetic induction. When the Larmor frequency corresponding to the magnetic induction will be equal to the frequency of the electromagnetic wave, a sharp attenuation in the wave intensity will be observed, indicating the onset of paramagnetic resonance.

Paramagnetic resonance can provide a great deal of diverse information on the properties of a paramagnetic and is widely used in scientific research.

The classical picture presented here and concerning the emergence of paramagnetic resonance is only of a qualitative nature. A more rigorous approach is possible in the framework of quantum theory, which is based on the concept of absorption and emission of quanta of electromagnetic radiation by atomic systems characterized by an abrupt reorientation of magnetic moments which ensure that the law of conservation of energy is obeyed. Within the framework of these concepts, it is possible to obtain quantitative relations characterizing paramagnetic resonance.

It follows from formula (40.13) that if the magnetic induction is equal to 1 T, the paramagnetic resonance frequency is of the order of 10^{10} Hz. As the magnetic induction decreases, this frequency also decreases accordingly and one can expect to observe paramagnetic resonance at relatively low frequencies. The resonance, however, cannot be observed at frequencies lower than 10^8 Hz, i.e. when the induction of the constant magnetic field is of the order of 0.01 T.

This is in accord with the quantum theory of paramagnetic resonance which predicts a significant decrease in the absorption of electromagnetic waves upon a decrease in their frequency. Consequently, the resonance at comparatively low frequencies is manifested very weakly. The most commonly used frequencies in actual practice are of the order of 10^{10} Hz (at a wavelength of 3 cm).

www.EngineeringBooksPDF.com

Sec. 42. Ferromagnetics

Basic experimental facts concerning ferromagnetism are discussed and an elementary theoretical interpretation is provided. General concepts about ferromagnetism, antiferromagnetism, ferrimagnetism and ferromagnetic resonance are introduced.

Definition. *Magnetics whose permeability attains large values and depends on an external magnetic field as well as the past history are called ferromagnetics.* They possess a residual magnetization, i.e. their magnetization may differ from zero in the absence of an external magnetic field. In this case, such materials are

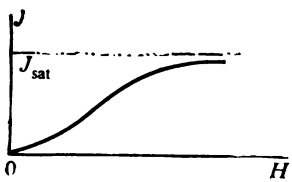


Fig. 165. Saturation magnetization

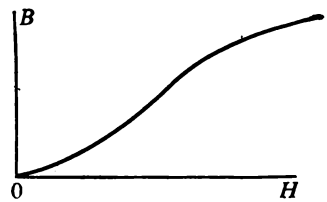


Fig. 166. Magnetization curve

permanent magnets. Thus, the formal manifestation of ferromagnetics is analogous to that of ferroelectrics (see Sec. 23). It should be noted that ferromagnetism was discovered and investigated a long time before ferroelectricity. Magnetization of ferromagnetics was investigated in 1878 by A.G. Stoletov (1839-1896). He constructed the permeability curve (see Fig. 168) which was later called the Stoletov curve. Hysteresis was discovered in 1880 by Wahrburh (1846-1931).

Magnetization curve and hysteresis loop. Magnetic susceptibility of ferromagnetics is a function of the external field, and the J vs H dependence has the form shown in Fig. 165. *The magnetization does not increase indefinitely with magnetic field, but has a limit called the saturation magnetization.* By analogy with paramagnetism, its existence indicates that the magnetization of ferromagnetics is also associated with the reorientation of certain elementary magnetic moments.

Since

$$B = \mu_0 H + \mu_0 J, \tag{42.1}$$

the B vs H curve does not show a saturation although J experiences saturation. This dependence is called the **magnetization curve** (Fig. 166).

www.EngineeringBooksPDF.com

If we carry out the alternating magnetization of a sample in a periodic magnetic field, the B vs H curve has the form of a loop called the hysteresis loop (Fig. 167) by analogy with ferroelectrics. The segment OA is the magnetization curve, since the field is switched on at zero induction, i.e. in the absence of permanent magnetization. The closed curve $ACDFGKA$ is the hysteresis loop.

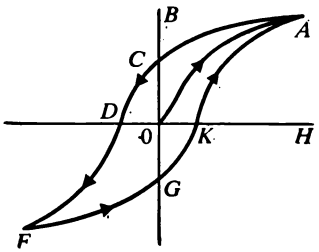


Fig. 167. Hysteresis loop

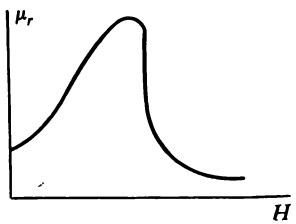


Fig. 168. Curve of relative permeability (Stoletov's curve)

This loop can be demonstrated with the help of a diagram similar to the one used for demonstrating the hysteresis loop of ferroelectrics when the capacitor is replaced by coils (see Sec. 23).

As the magnetic field H is reduced from a certain value (corresponding to point A) to zero, the magnetic induction decreases only slightly to a value characterized by the segment OC. This magnetic induction is called the **residual induction**. A ferromagnetic in this state is called a permanent magnet.

In order to neutralize the residual field, it is necessary to apply a reverse field whose strength is given by OD. This magnetic field strength is called **the coercive force of a ferromagnetic.** The shape of a hysteresis loop, residual induction and coercive force depend on the material of the ferromagnetic and differ for different types of materials over quite a wide range.

Permeability curve. The relative permeability $\mu_r = \mu/\mu_0 = B/(\mu_0 H)$ can be plotted as a function of H from the data of the magnetization curve (see Fig. 166) and has the form shown in Fig. 168. With increasing H, μ_r attains its maximum value after which it rapidly falls as the magnetic saturation is attained. Values of μ_r of the order of 10^4 at the maximum are not a rarity for ferromagnetics.

Classification of ferromagnetic materials. Ferromagnetics can be divided into two groups:

(1) magnetically soft materials, which have a high permeability, can be easily magnetized and demagnetized, and have a weak coercive force;

(2) magnetically hard materials which have a relatively low permeability, are difficult to magnetize or demagnetize, and have a strong coercive force.

The materials belonging to the first group are mainly used in electrical technology of alternating fields, especially in transformers, while the materials from the second group are used for making permanent magnets.

Interaction of electrons. Ferromagnetism can be analyzed only in the framework of quantum theory. *Classical theory of magnetism only describes the properties of*

www.EngineeringBooksPDF.com

ferromagnetics and qualitatively analyze the mechanism of emergence of ferromagnetism.

Einstein and de Haas were the first to establish experimentally that ferromagnetism is caused by the electron spins. Ferromagnetics have the property of spontaneous magnetization, when the electron spins tend to orient themselves inthe same direction in the absence of an external magnetic field. This orientation is due only to internal reasons. However, from the energy point of view, it is not advantageous for the sample to be magnetized as a whole. Hence the material is split into small magnetized regions, or domains. Each domain is magnetized in a certain direction, but the directions of the magnetization vector in neighbouring domains are different and hence the magnetic moment of small physical volumes is found to be equal to zero. In other words, the magnetic material is not magnetized on the whole.

This means that the *main question in the theory of ferromagnetism is to explain the tendency of electron spins to orient themselves in the same direction.* Since states with the lowest energy are encountered in a system, our task is to find the interaction under which a parallel orientation of the spin magnetic moments of different atoms is found to be advantageous from the point of view of energy. For this purpose, the total energy must be minimum for a parallel orientation of moments.

The emergence of such a situation is associated with the exchange interaction. Since the electrons obey the Fermi-Dirac statistics and hence two particles cannot exist in the same state, electrons with parallel spins as if move apart in space. Consequently, their Coulomb interaction energy is lower than that for electrons with antiparallel spins, in which case they can be arranged more closely in space. The **exchange interaction energy** is the difference in the energies corresponding to parallel and antiparallel spins.

Such a situation, however, does not ensure the emergence of ferromagnetism, since a decrease in the Coulomb interaction in the case of parallel spins is accompanied by an increase in the kinetic energy. In most cases, this increase compensates the decrease in the potential energy and the total energy corresponding to the state with parallel spins is not found to be advantageous. Only in rare cases, when the decrease in the potential energy for parallel spins is more than the increase in the kinetic energy, the total energy is reduced. In this case the configuration with parallel spins turns out to be more advantageous from the energy point of view and ferromagnetism is observed. An investigation of conditions under which such a situation is possible is the main task of the theory of ferromagnetism. The choice of the expression for the interaction energy plays a very significant role in this case.

Basic theory of ferromagnetism. In the theory of ferromagnetism, the exchange interaction energy is expressed through the formula

$$W_{ex} = -2I_{ex}\mathbf{S}_1 \cdot \mathbf{S}_2 \tag{42.2}$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the spins of the interacting electrons, and I_{ex} is the exchange interaction integral. It can be seen from this expression that for $I_{ex} > 0$, the potential energy attains its minimum value for parallel spins. This energy is

www.EngineeringBooksPDF.com

due to the interaction of the magnetic moment of an electron with the magnetic field and is expressed by a formula of type (41.1), where by **B** we mean the induction $\mathbf{B}_{ex}$ of the exchange field. The intrinsic magnetic moment $\mathbf{p}_m^{(0)}$ of an electron is connected with its intrinsic angular momentum, or spin, **S** through a relation of type (40.10) but the value of the proportionality factor in this case is twice as large:

$$\mathbf{p}_m^{(0)} = (e/m)\,\mathbf{S}. \tag{42.3}$$

Hence, expressing the interaction energy (42.2) as the energy corresponding to the magnetic moment of the second electron which is in the magnetic field created by the first electron due to the exchange interaction, we obtain

$$W_{ex} = -\frac{2I_{ex}S_1 m}{e}\cdot\frac{e}{m}\mathbf{S}_2 = -\mathbf{p}_{m2}^{(0)}\cdot\mathbf{B}_{ex} \tag{42.4}$$

where

$$\mathbf{B}_{ex} = (2I_{ex}m/e)\,\mathbf{S}_1. \tag{42.5}$$

The total magnetic induction is the sum of the induction **B** of the field in the absence of an exchange interaction and the induction $\mathbf{B}_{ex}$ of the exchange field. Taking (38.23) into account, we can write (38.21) in the form

$$\mu_0(1+\chi)\,\mathbf{J} = \chi\mathbf{B}, \quad \text{or} \quad \chi_0\mathbf{J} = [\chi/(1+\chi)]\,\mathbf{B}. \tag{42.6}$$

When the exchange interaction takes place, this relation is generalized to the following formula:

$$\mu_0\mathbf{J} = [\chi/(1+\chi)]\,(\mathbf{B} + \mathbf{B}_{ex}). \tag{42.7}$$

The magnetic susceptibility χ in this formula is assumed to be equal to its value in (42.6) for a paramagnetic in the absence of exchange interaction.

Subsequent analysis is carried out in the mean field approximation, whose basic assumption is that the exchange magnetic induction is proportional to magnetization:

$$\mathbf{B}_{ex} = \lambda\mu_0\mathbf{J}, \tag{42.8}$$

where λ is the exchange interaction constant. Substituting (42.8) into (42.7), we obtain the relation

$$\mu_0\mathbf{J} = [\chi/(1+\chi-\lambda\chi)]\,\mathbf{B}, \tag{42.9}$$

which can be written in a form similar to (42.7):

$$\mu_0\mathbf{J} = [\chi'/(1+\chi')]\,\mathbf{B}, \tag{42.10}$$

where

$$\chi'/(1+\chi') = \chi/(1+\chi-\lambda\chi) \tag{42.11}$$

characterizes the susceptibility taking into account the exchange interaction. From (42.11) we obtain

$$\chi' = \frac{\chi}{1-\chi\lambda} = \frac{C}{T-\lambda C}, \tag{42.12}$$

where $\chi = C/T$.

In the temperature range $T > \lambda C$, a body behaves as a paramagnetic with a characteristic decrease in the magnetic susceptibility with increasing tem-

www.EngineeringBooksPDF.com

perature. As we approach the temperature $T = \lambda C$, the susceptibility χ' tends to infinity. This means that indefinitely weak fields cause a finite magnetization. In other words, *a spontaneous magnetization or a transition to the ferromagnetic state is observed at* $T = \lambda C$. The elementary theory described above does not permit a quantitative analysis of the change in spontaneous magnetization upon a further decrease in temperature in the range $T < \lambda C$. A more exact theory shows that at $T = \lambda C$, spontaneous magnetization increases abruptly to a finite value and continues to increase further with decreasing T, although the rate of decrease gradually drops. Thus, a magnetic is in the ferromagnetic phase at $T < \lambda C$.

Curie-Weiss law. For every ferromagnetic there exists a temperature above which it undergoes a (second-order) phase transition and is transformed into

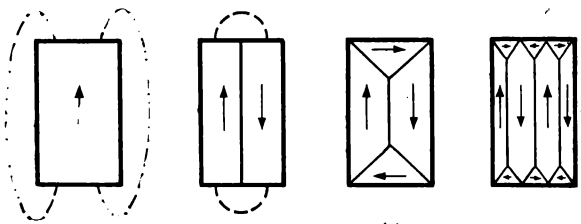


Fig. 169. Idealized domain structures in a single crystal

a paramagnetic. The magnetic susceptibility in the paramagnetic region in the vicinity of the transition temperature (called the Curie temperature) is described by a relation of the type (42.12), called the **Curie-Weiss law.** The quantity $\lambda C = \Theta$ is called the **Curie-Weiss temperature.** It has been shown theoretically that the phase transition takes place not at the Curie-Weiss temperature but at a temperature close to it.

Hence, the distinction between the Curie temperature at which the phase transition takes place and the Curie-Weiss temperature is sometimes not emphasized.

Magnetization anisotropy. It was shown during an investigation of magnetization curves for ferromagnetic monocrystals that different magnetization curves are obtained for different orientations of the magnetizing field with respect to the crystal axes. In other words, the ferromagnetic properties of crystals depend on the direction of magnetization. *The direction in which the magnetization is the strongest for a given value of the field is called the direction or axis of easy magnetization, while the direction corresponding to the lowest magnetization for a given field is called the direction or axis of difficult magnetization.*

Domains. Figure 169 shows idealized structures of domains in a single crystal (the arrows indicate the direction of magnetization):

(a) the external magnetic induction is strong;

(b) the external field is mainly concentrated near the upper and lower walls, and has a much lower energy than in case (a);

www.EngineeringBooksPDF.com

(c) there are no free poles and the field does not leave the domain,

(d) the same situation as (c) is encountered, but the structure is now split into smaller domains.

Domain boundaries. To minimize the magnetic field energy it is advantageous to reduce the domain size to the maximum possible extent. *This, however, is prevented by the need to spend energy for the formation of domain boundaries,* since magnetization on both sides of the boundary has different directions. The domain boundary has a finite thickness d, within which the magnetization slowly changes its direction from the orientation in one domain to the orientation in the other domain. In other words, the domain boundaries are the walls of finite thickness. The walls are classified in accordance with the peculiarities

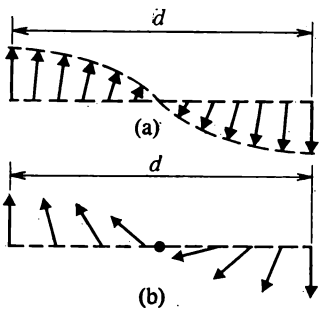


Fig. 170. A change in the direction of magnetization in the Bloch wall (a) and the Néel wall (b)

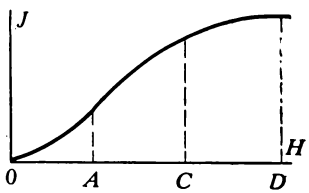


Fig. 171. Regions of different mechanisms of magnetic reversal

of rotation of the magnetization vector in them. If the magnetization component perpendicular to a wall does not change during its rotation, the wall is called a Bloch wall. In other words, the rotation of magnetization in a Bloch wall takes place in a plane parallel to the wall (Fig. 170*a*). If a change in the direction of magnetization is accompanied by a change in its component perpendicular to the wall, the wall is called a Néel wall (Fig. 170*b*).

Magnetic reversal. An increase in magnetization of a sample with increasing field strength occurs at first due to a reversible displacement of domain boundaries and rotations of the boundary walls (Fig. 171, segment OA). An irreversible displacement of boundaries and the disappearance of some domains takes place in the region AC. Finally, in the region CD preceding saturation, a change in the direction of magnetization inside the domains is observed.

Antiferromagnetism. Under certain conditions, exchange interaction leads to a situation in which the antiparallel orientation of the spin moments of neighbouring atoms is more favourable from the point of view of energy. For this purpose, it is necessary to create conditions which are identical to those under which ferromagnetism occurs, but they now correspond to antiparallel spin configurations. As a result, the spin magnetic moments of neighbouring atoms are found to be oriented in opposite directions (Fig. 172).

www.EngineeringBooksPDF.com

Such a situation can be interpreted as a simultaneous existence of two sublattices which are spontaneously magnetized in opposite directions and have the same intensity. The total magnetization in this case is equal to zero. This phenomenon is called **antiferromagnetism**, and the materials in which it is observed are called **antiferromagnetics.**

The exchange interaction vector in antiferromagnetics is directed against the magnetization vector **J**. Hence the following relation is valid for them instead of (42.8):

$$\mathbf{B}_{ex} = -\lambda_a \mu_0 \mathbf{J}. \qquad (42.13)$$

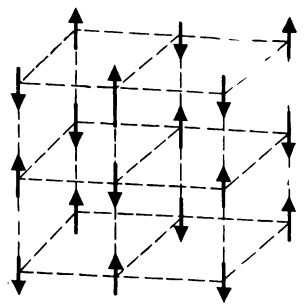


Fig. 172. Antiferromagnetism

Carrying out calculations identical to those which led from (42.8) to (42.12), we obtain formula (42.12) for the susceptibility of a ferromagnetic, but in this case λ is replaced by $-\lambda_a$:

$$\chi_a = C/(T + \lambda_a C) = C/(T + \Theta), \qquad (42.14)$$

where $\Theta = \lambda_a C$ is the Curie-Weiss temperature. As in ferromagnetics, the transition to the antiferromagnetic state takes place at a temperature other than the Curie-Weiss temperature. The temperature of transition to the antiferromagnetic state is called the **Néel temperature** T_N.

In the absence of a field below the Néel temperature, the total spontaneous magnetization of an antiferromagnetic is equal to zero since opposite magnetizations of the sublattices completely compensate each other. The application of an external magnetic field leads to the appearance of a small magnetization associated with positive susceptibility.

The two-sublattice model is good enough to explain the emergence of antiferromagnetism in most cases. However, in some cases when the situation is not confined to collinear magnetic moments and it is necessary to ensure that the vector sum of several magnetic moments is equal to zero (which is characteristic of antiferromagnetism), we have to use a model of more than two sublattices.

Ferrimagnetism. It may so happen that sublattices have spontaneous magnetizations in opposite directions but with different intensities. Thus, the magnetization is not eliminated completely as in the case of antiferromagnetism. Such materials have *a spontaneous magnetization which is, however, weaker than in materials whose magnetic moments are all oriented in the same direction.* Such materials have properties similar to those of ferromagnetics. In particular, they have a residual magnetization, are characterized by a coercive force, and so on. These materials are called **ferrimagnetics** or **ferrites.** *Ferrimagnetism is sometimes referred to as uncompensated antiferromagnetism.*

The significant advantages that ferrites have over ferromagnetics are due to their extremely low electrical conductivity. The ferromagnetics are good conductors of electric current, and this is a serious drawback when these materials are used in radio engineering.

www.EngineeringBooksPDF.com

By a **sublattice** we mean an aggregate of all ions in a crystal which are similar not only in their crystallographic structure, but also in their electrostatic and magnetic interactions with the surrounding ions. Consequently, the necessary condition for the existence of ferrimagnetism is the existence of at least two nonequivalent sublattices. The simplest situation in which ferrimagnetism can emerge are shown in Fig. 173*a-c*.

Ferromagnetic resonance. This phenomenon is associated with the interaction of the spin magnetic moments of electrons with a varying electromagnetic field. In ferromagnetics, however, this interaction is much more complicated

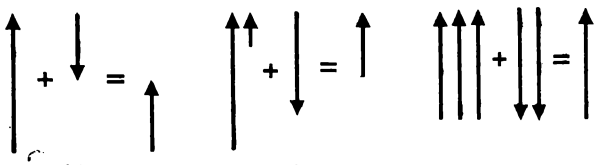


Fig. 173. Simple cases of emergence of ferrimagnetism

than in paramagnetics. This is due to the fact that ferromagnetics have a spontaneous magnetization and a domain structure while the spins of electrons are strongly connected through exchange interaction. Hence *the phenomenon of resonance in a ferromagnetic has a cooperative nature from the very beginning. The precession of spins, however, is caused not only by the external field, but also by an effective field which depends on the external field as well as on the internal field of the ferromagnetic, for example, the anisotropy field.*

Ferromagnetic resonance is observed at frequencies of several thousand megahertz. If a microwave field has a uniform amplitude, a uniform precession of spins is observed in the entire ferromagnetic sample, which leads to the appearance of a resonance peak. Simultaneously, other resonance peaks due to domain boundaries (domain-boundary resonance) are also observed. The inhomogene ty of the microwave field leads to the appearance of additional resonance peaks which depend on the shape and size of the sample. An analysis of this fairly complicated ferromagnetic resonance pattern provides valuable information on the properties of the ferromagnetic and allows us to measure several characteristics of its magnitude, such as saturation magnetization, gyromagnetic ratio, anisotropy constant, etc.

As in the case of ferromagnetism, only quantum theory can explain the phenomenon of ferromagnetic resonance.

A typical feature of the magnetization curve for ferromagnetics is the presence of saturation region. The curve describing the reversal of magnetization is characterized by hysteresis.

www.EngineeringBooksPDF.com

Sec. 43. Gyromagnetic Effects

Gyromagnetic effects and their experimental manifestations are described.

Relation between angular momentum and magnetic moment. The magnetization of a magnetic material is always associated with a reorientation of magnetic moments in a certain direction. New magnetic moments which are oriented in the same direction right from their origin are formed only in diamagnetic phenomena. The magnetic moment due to the orbital motion of an electron is connected with the angular momentum of this motion through relation (40.10). The intrinsic magnetic moment of an electron is also connected with its intrinsic angular momentum through a linear relation. Hence it is clear that the magnetic moment of an atom is also connected with its angular momentum through a certain relationship. This means that the *reorientation of magnetic moments takes place simultaneously with the reorientation of the corresponding angular momenta.*

The total magnetic moment of an atom is the sum of the magnetic moments due to the orbital motion of electrons and their spin magnetic moments. The angular momenta are summed up in the same way. However, since the proportionality factors between magnetic moments and angular momenta corresponding to the orbital motion differ from those corresponding to the spin, the total magnetic moment of the atom, generally speaking, is not collinear with its angular momentum and forms a certain angle with it (Fig. 174). The angular momentum of an isolated system is conserved. Consequently, the direction of $\mathbf{L}_t$ for a free atom remains unchanged in space. Hence, as a result of the motion of electrons in the atom, $\mathbf{p}_{mt}$ precesses round the direction of the total magnetic moment; the angular velocity of this precession is determined by the duration of intratomic processes, i.e. has very large values. Hence, *when the magnetic moment interacts with external fields, only the component* $\mathbf{p}_{m\,eff}$ *has an effective value in the direction of total angular momentum of the atom.* During interaction with an external field, the effective magnetic moment of the atom is $\mathbf{p}_{m\,eff}$, which is collinear with $\mathbf{L}_t$. Thus, the relation between the moments can be represented in all cases in the following form:

$$\mathbf{p}_m = ge\mathbf{L}/(2m), \tag{43.1}$$

where e and m are the charge and mass of an electron, and g is the **gyromagnetic ratio.** For the orbital motion of an electron, $g = 1$, while for spin $g = 2$. The value of g for atoms lies between 1 and 2 depending on the proportion and the manner in which the orbital motion of electrons and their spins contribute to the total magnetic moments. It should be recalled that in (43.1) by $\mathbf{p}_m$ we do not mean the real total magnetic moment of the atom, but its projection

www.EngineeringBooksPDF.com

onto the direction of the total angular momentum. In Fig. 174, it is denoted by $p_{m\ eff}$.

Einstein-de Haas experiment. Let us consider a cylinder made of a magnetic material and suspended from an elastic string (Fig. 175). Relation (43.1) between angular momenta and magnetic moments shows that during magnetization of the cylinder along the axis, the atoms acquire not only a magnetic moment along the cylinder axis, but also a corresponding angular momentum directed

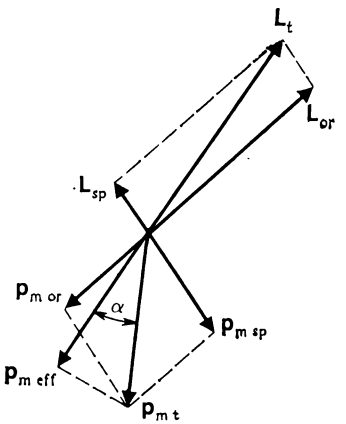


Fig. 174. Summation of magnetic moments and angular momenta in an atom

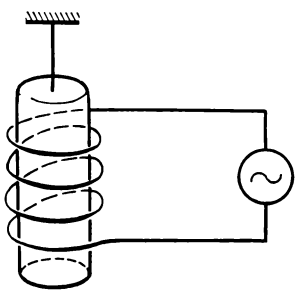


Fig. 175. Einstein-de Haas experiment

along the cylinder axis. The total angular momentum of a rod is the sum of the angular momenta of individual atoms and the angular momentum of the rod as a whole. Before magnetization, the total angular momentum of the rod is equal to zero. The total angular momentum is conserved in an isolated system. In the present case, the isolated system consists of a rod and the magnetizing field created by the solenoid currents.

Without going into the proof (see Chap. 9), it can be mentioned that the angular momentum of an electromagnetic field with respect to the cylinder axis is equal to zero and hence does not affect the law of conservation of angular momentum in this case. This means that the sum of the angular momenta of all atoms and of the rod as a whole must be constant, i.e. it should be equal to zero even after magnetization. But since the angular momentum of atoms changes as a result of magnetization, the angular momentum of the rod as a whole also changes. It follows from (43.1) that the following relation is satisfied during magnetization:

$$\Delta p_{mz} = g\,[e/(2m)]\,\Delta L_z, \tag{43.2}$$

where ΔL_z and Δp_z are the angular momentum and the magnetic moment respectively, acquired by each atom upon magnetization along the Z-axis. Sum-

www.EngineeringBooksPDF.com

ming both parts of Eq. (43.2) over all the atoms, we get

$$VJ = \Sigma\, \Delta p_{mz} = g\,[e/(2m)]\,\Sigma\, \Delta L_z, \tag{43.3}$$

where J is the magnetization of the rod and V is its volume. It follows from the law of conservation of angular momentum that the angular momentum acquired by the rod as a whole due to magnetization is equal to

$$L_z = -\,\Sigma\, \Delta L_z = -\,[2m/(eg)]\, VJ. \tag{43.4}$$

The angular velocity ω of rotation of the rod is connected with the angular momentum L_z along the axis of rotation and the moment of inertia I_z through the relation

$$L_z = I_z \omega. \tag{43.5}$$

The kinetic energy of rotation is equal to

$$W = \frac{1}{2} I_z \omega^2. \tag{43.6}$$

On the other hand, the torsion modulus D of the string is connected with frequency ω_0 of torsional vibrations of the rod through the relation

$$I_z \omega_0^2 = D. \tag{43.7}$$

Upon acquiring the kinetic energy (43.6), the rod twists the string by an angle θ, which can be determined from the law of conservation of energy as follows:

$$\frac{1}{2} I_z \omega^2 = \frac{1}{2} D\theta^2. \tag{43.8}$$

Taking into account (43.3), (43.4) and (43.7), we obtain from (43.8)

$$I_z \omega = D\theta^2/\omega = -\,2mVJ/(eg), \tag{43.9}$$

whence

$$g = -\,2mVJ\omega/(2\theta^2 D). \tag{43.10}$$

All quantities on the right-hand side are either known, or can be measured in principle. Thus the value of g can be determined with the help of this equation.

The twisting of the string upon magnetization is not large. Hence the experiment was actually carried out not by a single magnetization as before, but by multiple magnetic reversal of the sample at a frequency ω_0. This leads to an increase in the torsional vibrations of the sample, the amplitude of forced vibrations at resonance can be measured easily and reliably at a fairly high Q-factor. In principle, a transition to resonance build-up does not introduce any changes in the situation described above.

Einstein and de Haas carried out their experiments on ferromagnetic rods in which the magnetization effect is especially noticeable. It was found experimentally that

$$g = 2. \tag{43.11}$$

www.EngineeringBooksPDF.com

This value is twice that would be expected if the orbital motion of electrons in an atom were responsible for the magnetism. When these experiments were first carried out in 1915, nothing was known about the spin of an electron and the results obtained were considered to be enigmatic. When spin was discovered later, it was shown that $g = 2$ for it. It then became clear that the result of the Einstein-de Haas experiments directly indicates the fact that ferromagnetism is due to the intrinsic magnetic moment of electrons rather than to their orbital motion.

For other magnetic materials, the gyromagnetic ratio obtained from similar experiments is found to vary between 1 and 2. In all cases, the sign indicates that magnetism is due to motion of electrons.

Barnett effect. Any magnetic material possesses a diamagnetism. If it is a paramagnetic, its diamagnetism is caused by precession of the magnetic moments of atoms around the induction vector of the magnetic field created in a system of coordinates where the magnetic as a whole is at rest. In other words, its diamagnetism is a consequence of the precession of atoms with respect to the lattice. Let us rotate the magnetic as a whole. The individual atoms behave as small gyroscopes which tend to preserve the direction of their rotation axis in space. Hence the direction of the magnetic moments of individual atoms in space remains unchanged. Consequently, these magnetic moments will precess with respect to the crystal lattice of the magnetic with the frequency of rotation of the latter. Such an ordered precession of atoms with respect to the magnetic as a whole, however, leads to magnetization. Hence, *a magnetic is magnetized as a result of rotation.* This effect was first discovered by Barnet in 1909.

It is clear from the above discussion that when a magnetic rotates at a frequency ω, its magnetization is the same as for a diamagnetic introduced into a magnetic field whose induction is

$$B = 2m_e\omega/(|\, e \,|\, g). \tag{43.12}$$

It should be emphasized that *the rotation of a paramagnetic only creates a diamagnetic magnetization in it.* This magnetization is about two orders of magnitude smaller than the magnetization resulting from the paramagnetic effect (reorientation of magnetic moments).

Why are the total angular momentum and the total magnetic moment of an atom noncollinear?
Which quantity plays the role of the effective total magnetic moment of an atom during its interaction with external magnetic fields?
Why is the reversal of magnetization in a periodic external field used in the Einstein-de Haas experiment? What requirements are imposed on the frequency of an external field?
What is the nature of magnetization in the Barnet effect?

Problems

7.1. The diamagnetic susceptibility of copper (in solid state) is equal to $\chi_d = -8.8 \times 10^{-6}$. Find the mean distance of electrons from the nucleus in a copper atom.

www.EngineeringBooksPDF.com

7.2. The magnetic moment of an oxygen molecule is equal to $p_m = 2.6 \times 10^{-23}$ A·m². Find the paramagnetic susceptibility of oxygen under normal conditions.

7.3. The magnetic dipole moment of molecules is of the order of one Bohr magneton $\mu = e\hbar/(2m_e) = 9.27 \times 10^{-24}$ A·m². Considering that the molecules of an ideal gas have a permanent magnetic moment μ, find the maximum possible magnetization at $t = 100$ °C and $p = 101.3$ kPa.

Answers

7.1. $\sqrt{\langle R^2 \rangle} = \sqrt{-6m\chi_d/(e^2 Z \mu_0 N)} = 0.9 \times 10^{-10}$ m. 7.2. $\chi_p = p_m^2 N \mu_0/(3kT) = 18 \times 10^{-7}$.

7.3. $J_{max} = 182$ A/m.

www.EngineeringBooksPDF.com

CHAPTER 8

Electromagnetic Induction and Quasistationary Alternating Currents

The quasistationary approximation is valid for describing electromagnetic fields in the regions whose linear dimensions are much smaller than the wavelength and when displacement currents can be neglected. The electric field generated due to a variation of the magnetic field is taken into account, while the magnetic field generated by the variation of the electric field is discarded. The lines of the conduction current density are closed since displacement currents are ignored. The magnetic field is determined by the instantaneous values of density of conduction currents at the same moment of time. The conduction current densities depend on the variation of the magnetic field, and hence on the variation of the conduction current density.

Sec. 44. Currents Induced in Moving Conductors

A quantitative description of currents induced in moving conductors is given. Physical processes occurring in a.c. generators are described.

Emergence of an e.m.f. in a moving conductor. ***Free electrons*** **in a conductor** moving in a magnetic field *are set in motion relative to the conductor under the action of Lorentz' force, i.e. an electric current appears in the conductor. This phenomenon is called the induction of currents in moving conductors.*

Let us consider a rectilinear segment DG of a conductor (Fig. 176) which, moving at a velocity $\mathbf{v}$, slides along conductors CK and AL, as along the guides so that the loop $AGDCA$ remains closed. The magnetic induction of an external uniform magnetic field is normal to the plane containing the loop. The Lorentz force acting on moving charges is

$$\mathbf{F} = e\mathbf{v} \times \mathbf{B}, \tag{44.1}$$

This force is directed along DG. The forces acting on positive and negative charges of the conductor are shown by vectors $\mathbf{F}_{(+)}$ and $\mathbf{F}_{(-)}$ respectively. Free electrons are set in motion and form an electric current. Its direction is taken as the positive direction of circumvention of the loop. Consequently, the vector $\mathbf{n}$ shown in the figure is the positive normal to the surface containing the loop.

www.EngineeringBooksPDF.com

The force **F** [see (44.1)] is equivalent to an effective electric field acting on charges in the conductor:

$$\mathbf{E}_{\text{eff}} = \mathbf{F}/e = \mathbf{v} \times \mathbf{B} \qquad (44.2)$$

and hence the e.m.f. induced between certain points *1* and *2* of the conductor is given by

$$(\Delta\mathscr{E}^{\text{ind}})_{21} = \int_{(1)}^{(2)} \mathbf{E}_{\text{eff}} \cdot \mathrm{d}\mathbf{l} = \int_{(1)}^{(2)} \mathbf{v} \times \mathbf{B} \cdot \mathrm{d}\mathbf{l}. \qquad (44.3)$$

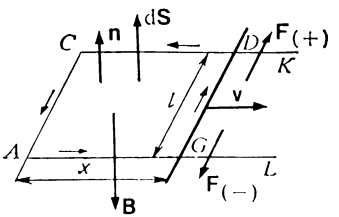


Fig. 176. Currents induced in moving conductors

In the case under consideration, this e.m.f. appears between points D and G:

$$(\Delta\mathscr{E}^{\text{ind}})_{DG} = \int_{(G)}^{(D)} vB\,\mathrm{d}l = vBl. \qquad (44.4)$$

On the fixed sections of the closed loop no e.m.f. is induced. Therefore, the electromotive force induced in the closed loop $AGDCA$ due to motion of its section DG in the external field is

$$\mathscr{E}^{\text{ind}} = \int_{AGDCA} \mathbf{E}_{\text{eff}} \cdot \mathrm{d}\mathbf{l} = vBl. \qquad (44.5)$$

Expressing the velocity of the conductor DG in the form

$$v = dx/dt \qquad (44.6)$$

where x is the coordinate of its contacts with the guiding conductors at points D and G, we present (44.5) in the form

$$\mathscr{E}^{\text{ind}} = \mathrm{d}x\, lB/\mathrm{d}t \qquad (44.7)$$

Recall that

$$\Phi = -xlB \qquad (44.8)$$

is the magnetic flux through the surface bounded by the loop $AGDCA$. The minus sign in this formula indicates that the directions of **B** and dS are opposite. Consequently, formula (44.5) can now be written in the final form:

$$\boxed{\mathscr{E}^{\text{ind}} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t},} \qquad (44.9)$$

i.e. *the electromotive force induced in a closed conductor moving in an external magnetic field is equal to the rate of variation of the magnetic flux of the external magnetic field through the surface stretched over the closed loop.*

Formula (44.9) was derived for a particular case when only a part of the conductor moves in the plane perpendicular to the magnetic induction vector. If several segments of a conductor move simultaneously, the electromotive force induced in the closed conductor is equal to the algebraic sum of e.m.f.s induced

www.EngineeringBooksPDF.com

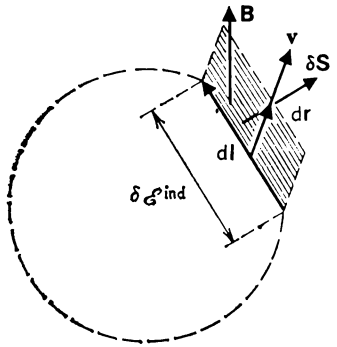


Fig. 177. Generalization of the formula for currents induced in moving conductors for an arbitrary case

in these sections. Therefore, without further calculations, formula (44.9) is generalized for the case of an arbitrary motion of a conductor in the plane perpendicular to the magnetic induction vector. Naturally, during this motion the loop formed by the conductor may be arbitrarily deformed.

Generalization to an arbitrary case. Let us consider an element of length dl of a conductor, moving at a velocity $v = dr/dt$ (Fig. 177). In accordance with formula (44.3), the electromotive force induced over this length is given by

$$d\mathcal{E}^{ind} = \mathbf{v} \times \mathbf{B} \cdot d\mathbf{l} = \frac{d}{dt}(d\mathbf{r} \times \mathbf{B} \cdot d\mathbf{l}). \qquad (44.10)$$

The triple scalar product in this formula is transformed as follows:

$$d\mathbf{r} \times \mathbf{B} \cdot d\mathbf{l} = d\mathbf{l} \times d\mathbf{r} \cdot \mathbf{B} = -d\mathbf{r} \times d\mathbf{l} \cdot \mathbf{B} = -d\mathbf{S} \cdot \mathbf{B} = -\delta\Phi, \qquad (44.11)$$

where $\delta\Phi$ is the magnetic flux through the surface element $d\mathbf{S} = d\mathbf{r} \times d\mathbf{l}$ formed by the element of length dl during its motion. The positive direction of the normal to this surface element is chosen so that it coincides with the positive direction to the surface bounded by the closed loop.

Substituting (44.11) into (44.10), we obtain

$$d\mathcal{E}^{ind} = -\delta\Phi/dt. \qquad (44.12)$$

In order to find the total electromotive force induced in the closed loop, we must sum up the e.m.f.s induced in all elements of length dl of this loop:

$$\mathcal{E}^{ind} = \oint d\mathcal{E}^{ind} = -\frac{1}{dt}\oint \delta\Phi = -\frac{d\Phi}{dt}, \qquad (44.13)$$

where

$$\oint \delta\Phi = d\Phi \qquad (44.14)$$

is the variation of the magnetic flux through the surface bounded by the closed loop.

Formula (44.13) coincides with (44.9). Thus we have proved that the *latter is valid for arbitrary motions and deformations of a closed loop.*

A.c. generators. If a closed conductor moves in a magnetic field so that the magnetic flux enveloped by it constantly varies, an electromotive force and the corresponding alternating current are continuously induced in it. In other words, such a closed loop is an **a.c. generator.** A simple diagram of an a.c. generator is shown in Fig. 178*a*. If the magnetic field is uniform and the loop rotates in it with a constant angular velocity, the e.m.f. $\mathcal{E}^{ind}$ induced in the loop is a harmonic electromotive force whose frequency is equal to the frequency

www.EngineeringBooksPDF.com

of the loop rotation in the magnetic field. An alternating current of the corresponding frequency appears in the closed loop (Fig. 178*b*).

If instead of one loop, two parallel series-connected loops move in a magnetic field, the induced e.m.f. becomes twice as large as in the previous case. For this reason, in real generators the windings containing a large number of turns are used. The problems of optimization of the shape of windings, creation of a magnetic field, removal of current from rotating windings, etc. are considered in detail in electrical engineering. It should only be noted here that the removal

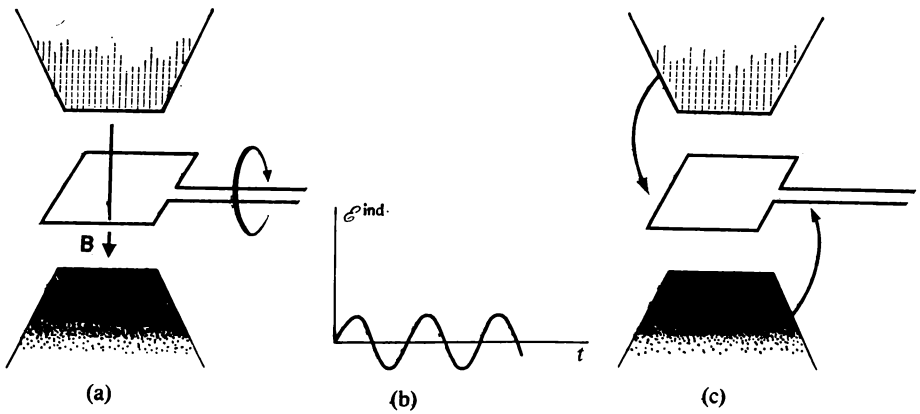


Fig. 178. Schematic diagram of an a.c. generator

of strong currents from moving windings is quite a complicated problem. For this reason, instead of current-carrying conductors, the sources of magnetic field are moved, and the conductors remain fixed. In the simple diagram (Fig. 178*c*) this corresponds to the motion of permanent magnets around a fixed current loop. The e.m.f. induced in the fixed loop is quantitatively the same for identical relative velocities of the magnets and the loop. However, the physical nature of phenomena occurring in these two cases is different.

First generators employed permanent magnets, but even in 1866 a generator was designed in which the magnetic field was created by an electromagnet. After this the design of generators was perfected further.

The law of conservation of energy. When a current flows in a circuit containing an ohmic resistance, Joule's heat is liberated. The energy liberated in the form of heat is obtained at the expense of the work of mechanical forces in a current generator.

Of course, when the energy is transformed from one form to another, the law of conservation of energy is observed. Let us illustrate this by using a simple example (see Fig. 176).

Let R be the resistance in the circuit $AGDCA$ and I be the current in it. Consequently, the energy liberated in the circuit in the form of heat is characterized

www.EngineeringBooksPDF.com

by the power

$$P_1 = I^2R. \tag{44.15}$$

On the other hand, when the section DG of the conductor with current I moves, the Lorentz force

$$F = IlB. \tag{44.16}$$

should be overcome. Consequently, the forces responsible for the motion of the conductor must develop the power

$$P_2 = Fv = IlB\ \mathrm{d}x/\mathrm{d}t = -I\mathscr{E}^{\mathrm{ind}} = -I^2R, \tag{44.17}$$

where we take into account formula (44.9) and the fact that $\mathscr{E}^{\mathrm{ind}} = IR$. The minus sign in (44.17) indicates that the work is done on the system. A comparison of (44.15) and (44.17) shows that $P_1 + P_2 = 0$. This means that the energy liberated in the form of heat in the circuit is equal to the work of forces moving the conductor. In other words, the extraneous electromotive forces in this case are ultimately mechanical forces causing the motion of the conductor.

When a closed loop moves and is deformed in an external magnetic field, the e.m.f. induced in it is numerically equal to the rate of variation of the magnetic flux of the external field through the surface stretched over the closed loop.
The entire work performed by the current induced in a moving conductor is done at the expense of the work of the forces that set the conductor in motion.

What physical phenomena form the basis of a.c. generators? Describe the basic types of generators.

Sec. 45. Faraday's Law of Electromagnetic Induction

The physical essence and mathematical formulation of Faraday's law are discussed. The relationship between Faraday's induction and the induction of current in moving conductors is analyzed.

Definition. In 1831, Faraday experimentally discovered electromagnetic induction, manifested in the appearance of an electric current in a closed conductor upon a change in a magnetic flux through the contour formed by the conductor. The direction of induced e.m.f. is determined by the law formulated in 1833 by H.F.E. Lenz (1804-1865): *induced current is in such a direction as to oppose the magnetic flux variation causing it.* In other words, the direction of current induced in a loop forms with the direction of magnetic flux variation a left-handed system (Fig. 179). In 1845, F. E. Neumann (1798-1895) formulated the law of electromagnetic induction in its present form for a fixed loop:

$$\mathscr{E}^{\mathrm{ind}} = -\mathrm{d}\Phi/\mathrm{d}t, \tag{45.1}$$

www.EngineeringBooksPDF.com

Physical essence of the phenomenon. Formula (45.1) is completely identical to (44.9) in its appearance, but has a quite different physical content. *The induced e.m.f. described by formula* (44.9) *is associated with the Lorentz force acting on moving charges. On the other hand, no Lorentz force participates in creating the e.m.f. described by formula* (45.1) *since the conductors are fixed.* An electric current, however, appears in the conductor, and hence we may conclude that an electric field is present in it. Consequently, *Faraday's law* (45.1) *expresses a new physical phenomenon: a varying magnetic field generates an electric field. Thus, electric field is created not only by electric charges but by a varying magnetic field as well.*

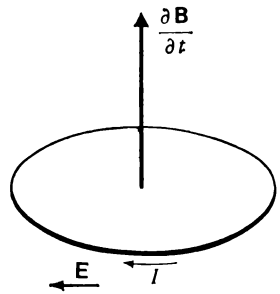


Fig. 179. Faraday's law of electromagnetic induction

Strictly speaking, the presence of current in a closed conductor indicates that an electric field exists only inside the conductor. However, in the case under consideration, the conductor detects an electric field. In the absence of the conductor, a varying magnetic field also generates an electric field. This can be confirmed, for example, by the fact that an electric force is acting on a charge in a varying magnetic field (see Sec. 56). This proves that the law of electromagnetic induction is a fundamental law establishing the relation between electric and magnetic fields. The difference in the physical content of phenomena described by formulas (44.9) and (45.1) can be demonstrated by using the following example. Suppose that the conductor DG shown in Fig. 176 moves at a velocity $\mathbf{v}$ and the magnetic induction $\mathbf{B}$ decreases simultaneously. As a result of motion of the conductor, an e.m.f. is induced in the loop, which generates a current (Fig. 176). In accordance with Faraday's law, the variation of $\mathbf{B}$ also induces an e.m.f. in the loop, which is directed against the e.m.f. created due to motion of the conductor DG. We can choose such a rate of variation of $\mathbf{B}$ ($\partial\mathbf{B}/\partial t$) that these two e.m.f.s compensate each other. As a result, no current will flow in the loop since the total induced e.m.f. is equal to zero. However, the mutual compensation of induced e.m.f.s occurs in the loop as a whole and not at each point of the loop. The e.m.f. induced due to motion of the conductor appears only in the segment DG, while the e.m.f. induced according to Faraday's law appears on the segment DG as well as on the remaining segments DC, CA and AG. The e.m.f. induced in the element of length $d\mathbf{l}$ of the conductor as a result of motion depends only on $\mathbf{B}$ and on the velocity $\mathbf{v}$ of motion of this element but does not depend on $\partial\mathbf{B}/\partial t$. On the other hand, the e.m.f. induced in the element $d\mathbf{l}$ of the conductor as a result of the variation of the magnetic induction does not depend on the induction $\mathbf{B}$ and on the velocity $\mathbf{v}$ of motion of this element and is determined only by $\partial\mathbf{B}/\partial t$. This shows that the physical nature of e.m.f.s induced in these two cases is different.

A conductor moving in a varying magnetic field. If a closed conductor moving in a varying magnetic field experiences arbitrary deformations, the e.m.f. induced in it is due to the motion and deformations, taken into account by for-

www.EngineeringBooksPDF.com

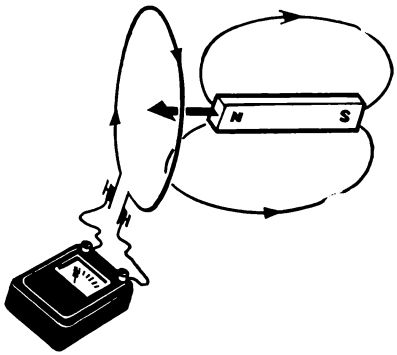


Fig. 180. Demonstration of Faraday's electromagnetic induction

mula (44.9), and as a result of the variation of the magnetic induction, which is taken into account by a similar formula (45.1). Therefore, it can be shown that the e.m.f. induced in the conductor is determined by formula (45.1), where $d\Phi/dt$ is treated as the total rate of variation of the magnetic flux embraced by the conductor, which occurs due to its motion and deformation and as a result of variation of the magnetic field.

Application of the law of electromagnetic induction to a.c. generators. It now becomes clear why an electric current can be generated not only by the motion of conductors in a magnetic field but also by the motion of magnets relative to fixed conductors. Figure 180 presents a schematic arrangement for demonstrating electromagnetic induction.

An electric field can be generated not only by electric charges but also by a varying magnetic field.
Induced e.m.f. is expressed by formula (45.1), where $d\Phi/dt$ stands for the total rate of variation of the magnetic flux through the loop due to motion and deformation of the conductor as well as due to variation of the magnetic field.

Sec. 46. Differential Form of the Law of Electromagnetic Induction

Differential form of Faraday's law is given and the properties of vector and scalar potentials of a varying electromagnetic field are discussed.

Differential form of Faraday's law. We shall write **Faraday's law of electromagnetic induction** [see (45.1)] in the form

$$\oint_L \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S}, \tag{46.1}$$

where L is the contour and S the surface stretched over this contour. In this formula, the following definitions are taken into account:

$$\mathscr{E}^{ind} = \int_L \mathbf{E}\cdot d\mathbf{l}, \quad \Phi = \int_S \mathbf{B}\cdot d\mathbf{S}. \tag{46.2}$$

It should be noted that the direction of circumvention of the contour L and the vector $d\mathbf{S}$ form a right-handed system. It should also be emphasized that

www.EngineeringBooksPDF.com

the surface S appearing in the definition of the magnetic flux Φ [see (46.2)], through which the flux to be determined passes, is an arbitrary surface stretched over the contour L. Such a definition presumes that the surface integral does not depend on the shape of the surface. It is only required that the surface be bounded by the contour L or, so to say, stretched over contour L. Let us prove this. We choose any two surfaces stretching over the contour L. Their combination forms a closed surface $S = S_1 + S_2$, bounding a certain volume V between them. The flux of $\mathbf{B}$ through the closed surface S is equal to zero since, according to the Gauss theorem, it is equal to the integral of div $\mathbf{B} = 0$ over the volume V bounded by the surface S. Hence we can state that the fluxes through S_1 and S_2 are equal (the fluxes have the same signs for the same orientation of positive normals to these surfaces relative to the direction of circumvention of the contour).

Let us transform the left-hand side of (46.1) according to the Stokes formula

$$\int_L \mathbf{E}\cdot d\mathbf{l} = \int_S \operatorname{curl}\mathbf{E}\cdot d\mathbf{S}. \tag{46.3}$$

This gives

$$\int_L \operatorname{curl}\mathbf{E}\cdot d\mathbf{S} = -\int_S \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{S}, \tag{46.4}$$

where the time derivative is introduced into the integral since the integration surface does not depend on time. Since S is an arbitrary surface, it follows from (46.4) that

$$\boxed{\operatorname{curl}\mathbf{E} = -\partial\mathbf{B}/\partial t.} \tag{46.5}$$

This equation is the **differential form of Faraday's law of electromagnetic induction**. It describes the law of generation of an electric field at a certain point due to a variation of the magnetic induction at the same point. The field $\mathbf{E}$ is often called the induced field.

Nonpotential nature of induced electric field. In an alternating magnetic field, $\partial\mathbf{B}/\partial t \neq 0$, and hence, in accordance with (46.5),

$$\operatorname{curl}\mathbf{E} \neq 0. \tag{46.6}$$

This means that *unlike the electrostatic field created by fixed charges, the induced electric field is not a potential field.* The work done in this field on moving a charge q over a closed contour generally is not equal to zero:

$$A = q\mathscr{E}^{\text{ind}} = q\int_L \mathbf{E}\cdot d\mathbf{l} \neq 0. \tag{46.7}$$

Hence it follows, in particular, that this field cannot be represented as the gradient of a certain function, i.e. cannot be represented in the form (14.27). For this field, some different representation should be used.

www.EngineeringBooksPDF.com

Vector and scalar potentials in a varying electromagnetic field. Since the law of electromagnetic induction is not related to the laws of generation of a magnetic field, Eq. (36.4) for the divergence of magnetic induction remains unchanged, i.e. div $\mathbf{B} = 0$. Consequently, formula (37.2) relating the vector potential to the magnetic induction remains in force:

$$\mathbf{B} = \operatorname{curl} \mathbf{A}. \tag{46 8}$$

The relation between the scalar potential with the electric field strength is, however, altered. Expressing $\mathbf{B}$ in (46.5) with the help of (46.8), we obtain

$$\operatorname{curl} \mathbf{E} = -\frac{\partial}{\partial t} \operatorname{curl} \mathbf{A} = -\operatorname{curl} \frac{\partial \mathbf{A}}{\partial t}, \tag{46.9}$$

where the order of differentiation with respect to time and with coordinates is changed in view of their independence. Equation (46.9) written in the form

$$\operatorname{curl} \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0 \tag{46 10}$$

shows that vector $\mathbf{E} + \partial \mathbf{A}/\partial t$ is a potential vector and hence can be represented in the form of the gradient of a certain function

$$\mathbf{E} + \partial \mathbf{A}/\partial t = -\operatorname{grad} \varphi, \tag{46.11}$$

where φ is the scalar potential. Thus, for varying fields the electric field strength is expressed not only in terms of the scalar potential but through the vector potential as well:

$$\boxed{\mathbf{E} = -\operatorname{grad} \varphi - \partial \mathbf{A}/\partial t.} \tag{46.12}$$

The first term on the right-hand side of this formula takes into account the electric field generated by electric charges, while the second term describes the field generated in accordance with Faraday's law of electromagnetic induction.

Ambiguity of potentials and gauge transformation. As in the stationary case, the scalar and vector potentials for varying electromagnetic fields are not unique. In other words, the same electromagnetic field can be described by a large number of scalar and vector potentials.

Suppose that a field $\mathbf{E}$, $\mathbf{B}$ is described by potentials $\mathbf{A}$ and φ given by formulas (46.8) and (46.12), and there is a certain arbitrary function $\chi(x, y, z, t)$. We state that the potentials

$$\mathbf{A}' = \mathbf{A} + \operatorname{grad} \chi, \quad \varphi' = \varphi - \partial \varphi/\partial t \tag{46.13}$$

characterize the same field $\mathbf{E}$, $\mathbf{B}$ as that characterized by potentials A and φ. In order to prove this, let us find the fields $\mathbf{E}'$ and $\mathbf{B}'$ described by potentials $\mathbf{A}'$ and φ' by using formulas (46.8) and (46.12):

$$\mathbf{B}' = \operatorname{curl} \mathbf{A}' = \operatorname{curl} \mathbf{A} + \operatorname{curl} \operatorname{grad} \chi = \mathbf{B}, \tag{46.14}$$

where we took into account the fact that curl grad $\chi = 0$ and formula (46.8).

www.EngineeringBooksPDF.com

Then we obtain the following expression for E':

$$E' = -\operatorname{grad}\varphi' - \partial\mathbf{A}'/\partial t = -\operatorname{grad}\varphi - \operatorname{grad}(\partial\chi/\partial t)$$
$$-\partial\mathbf{A}/\partial t - \partial(\operatorname{grad}\chi)/\partial t = -\operatorname{grad}\varphi - \partial\mathbf{A}/\partial t = \mathbf{E}. \quad (46.15)$$

Thus, potentials (46.13) indeed describe the same field as the potentials **A** and φ. Transformations (46.13) are called **gauge transformations**. They make it possible to "calibrate" potentials, i.e. to impose on them a certain condition by using their ambiguity (see Secs. 14, 37, 63).

Sec. 47. Magnetic Field Energy

The formulas for the magnetic field energy of current loops and the expression for energy density are derived. Expressions are given for the energy of a magnetic in an external magnetic field and for body forces acting on compressed magnetics.

Magnetic field energy for an isolated current loop. In order to create an electric current in a fixed loop, it is necessary to connect an extraneous e.m.f. source to the circuit. In a d.c. circuit, the energy supplied to the circuit from the source of extraneous e.m.f. is spent as Joule's heat and for accomplishing work in the load. In this case, the magnetic induction as well as the field energy remain unchanged. The magnetic induction varies with current. Consequently, the source of extraneous e.m.f.s supplies energy to the circuit to create a magnetic field with increasing current. Calculating the work done by the source of extraneous e.m.f. required to increase the current from zero to a finite value, we obtain the energy of the magnetic field associated with this current.

When the magnetic flux piercing a loop changes, an e.m.f. is induced in the loop in accordance with law (46.1). For an isolated loop, the magnetic flux Φ appears due to a magnetic field generated by current in the loop (Fig. 181). As the current increases, the magnetic flux Φ embraced by the current also increases, and the e.m.f. induced in the loop in accordance with Faraday's law is called the e.m.f. of self-induction (self-induced e.m.f.). In accordance with Lenz's law, it is directed so that it opposes an increase in the current. For increasing current it is necessary that the extraneous e.m.f. of the source be opposite to the self-induced e.m.f. and equal in magnitude to it. Thus, *as the current increases, the source of extraneous e.m.f. performs work against the self-in-*

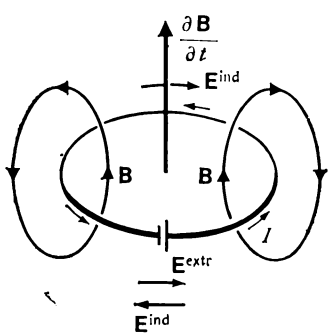


Fig. 181. When the current increases, a source of extraneous e.m.f.s performs the work against the e.m.f. of self-induction

www.EngineeringBooksPDF.com

duced e.m.f. (sometimes called the back-electromotive force). If the quantity of electricity passing through the circuit during time dt is $Q = I\,dt$, the work done by the source of extraneous forces during this time against the self-induced e.m.f. is

$$dA = -\mathcal{E}^{ind} I\,dt = (d\Phi/dt)\,I\,dt = I\,d\Phi \tag{47.1}$$

where $\mathcal{E}^{ind}$ is given by formula (46.1). This work is done when the energy of the source of extraneous e.m.f.s is transformed into the energy of the magnetic

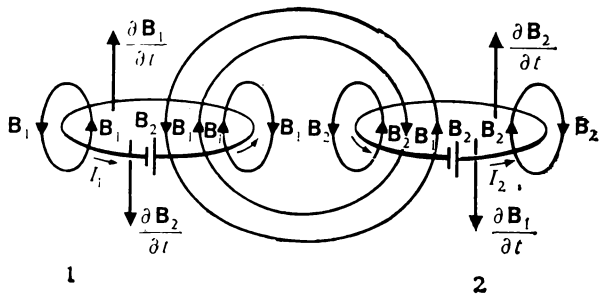


Fig. 182. To the calculation of the magnetic energy of the field of two current loops

field in the current loop. Hence the change in the magnetic field energy is connected with the change in the flux through the relation

$$dW = I\,d\Phi. \tag{47.2}$$

According to the Biot-Savart law (10.10), the magnetic induction of the field is proportional to the current creating the field. Consequently, *for an alternating current flowing in a rigid fixed loop, the pattern of field lines remains unchanged, and at each point the magnetic induction varies in proportion to the current. This means that the magnetic flux through a fixed stationary surface is also proportional to the current*, and hence

$$\Phi = LI, \tag{47.3}$$

where L is a constant proportionality factor which does not depend on current and magnetic induction. This coefficient is called the **inductance of the loop.**

Substituting (47.3) into (47.2), we obtain

$$dW = LI\,dI = d\,(^1/_2\ LI^2). \tag{47.4}$$

Integrating both sides of this equation between $I = 0$ and a certain value I, we obtain the formula

$$W = {}^1/_2\,LI^2, \tag{47.5}$$

which defines **the energy of the magnetic field** generated by the current I flowing in the circuit with inductance L.

Magnetic field energy for several current loops. Similarly we can find **the energy** of the magnetic field of two current loops (Fig. 182). *It should be borne in mind*

www.EngineeringBooksPDF.com

that an e.m.f. induced in each loop increases not only at the expense of the variation of the magnetic flux of the field created by this loop but also due to the change in the magnetic flux of the field created by the other loop. We denote by I_1 and I_2 the currents in the first and second loops and by Φ_{11} and Φ_{12} the magnetic fluxes of the fields generated by currents I_1 and I_2 and piercing the first loop. Similar quantities for the second loop are denoted by Φ_{22} and Φ_{21}. The total fluxes piercing each loop are given by

$$\Phi_1 = \Phi_{11} + \Phi_{12}, \quad \Phi_2 = \Phi_{21} + \Phi_{22}. \tag{47.6}$$

Let L_{11} and L_{22} be the inductances of the loops. Then [see (47.3)]

$$\Phi_{11} = L_{11}I_1, \quad \Phi_{22} = L_{22}I_2. \tag{47.7}$$

From the same considerations as those used in deriving formula (47.3) we conclude that the magnetic flux Φ_{12} through the first loop, created by the current I_2 in the second loop, is proportional to current I_2 in the second loop:

$$\Phi_{12} = L_{12}I_2, \tag{47.8}$$

where L_{12} is a constant called **mutual inductance** of the first and second loops. Similarly, for the second loop we obtain

$$\Phi_{21} = L_{21}I_1. \tag{47.9}$$

Hence [see (47.6)]

$$\Phi_1 = L_{11}I_1 + L_{12}I_2, \quad \Phi_2 = L_{21}I_1 + L_{22}I_2. \tag{47.10}$$

The e.m.f.s induced in the first and second loops are given by

$$\begin{aligned} \mathscr{E}_1^{\text{ind}} &= -\frac{d\Phi_1}{dt} = -\left(L_{11}\frac{dI_1}{dt} + L_{12}\frac{dI_2}{dt}\right), \\ \mathscr{E}_2^{\text{ind}} &= -\frac{d\Phi_2}{dt} = -\left(L_{21}\frac{dI_1}{dt} + L_{22}\frac{dI_2}{dt}\right). \end{aligned} \tag{47.11}$$

In analogy with (47.1), the entire work done by the sources of extraneous e.m.f.s in the loops during the time dt is given by

$$\begin{aligned} dA &= dA_1 + dA_2 = -\mathscr{E}_1^{\text{ind}} I_1\, dt - \mathscr{E}_2^{\text{ind}} I_2\, dt \\ &= (L_{11}I_1\, dI_1 + L_{12}I_1\, dI_2 + L_{21}I_2\, dI_1 + L_{22}I_2\, dI_2), \end{aligned} \tag{47.12}$$

where relations (47.10) are used.

For further calculations, let us prove that $L_{12} = L_{21}$. For this purpose, we calculate Φ_{21} and Φ_{12}:

$$\Phi_{21} = \int_{S_2} \mathbf{B}_1 \cdot d\mathbf{S}_2, \quad \Phi_{12} = \int_{S_1} \mathbf{B}_2 \cdot d\mathbf{S}_1, \tag{47.13}$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are the magnetic inductions of the fields generated by currents I_1 and I_2 respectively, S_1 and S_2 are the integration surfaces stretched over the loops. The magnetic induction at each point is equal to $\mathbf{B}_1 + \mathbf{B}_2$. Denoting by $\mathbf{A}_1$ and $\mathbf{A}_2$ the vector potentials describing the fields $\mathbf{B}_1$ and $\mathbf{B}_2$,

www.EngineeringBooksPDF.com

we have

$$\mathbf{B}_1 = \operatorname{curl} \mathbf{A}_1, \quad \mathbf{B}_2 = \operatorname{curl} \mathbf{A}_2$$

and thus equations (47.13) assume the form

$$\begin{aligned} \Phi_{21} &= \int_{S_2} \operatorname{curl} \mathbf{A}_1 \cdot d\mathbf{S}_2 = \int_{L_2} \mathbf{A}_1 \cdot d\mathbf{l}_2, \\ \Phi_{12} &= \int_{S_1} \operatorname{curl} \mathbf{A}_2 \cdot d\mathbf{S}_1 = \int_{L_1} \mathbf{A}_2 \cdot d\mathbf{l}_1, \end{aligned} \tag{47.14}$$

where L_1 and L_2 are the current loops. The transition to the integration over closed circuits is made in accordance with the Stokes formula. Formula (37.11b) expressing the vector potential in terms of current assumes in this case the following form:

$$\mathbf{A}_1 = \frac{\mu_0}{4\pi} I_1 \int_{L_1} \frac{d\mathbf{l}_1}{r}, \quad \mathbf{A}_2 = \frac{\mu_0}{4\pi} I_2 \int_{L_2} \frac{d\mathbf{l}_2}{r}. \tag{47.15a}$$

Substituting this formula into (47.14), we obtain

$$\Phi_{21} = \frac{\mu_0}{4\pi} I_1 \int_{L_2} \int_{L_1} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r_{21}}; \quad \Phi_{12} = \frac{\mu_0}{4\pi} I_2 \int_{L_1} \int_{L_2} \frac{d\mathbf{l}_2 \cdot d\mathbf{l}_1}{r_{12}}, \tag{47.15b}$$

where $r_{12} = r_{21}$ is the distance between the elements $d\mathbf{l}_1$ and $d\mathbf{l}_2$ of the first and second loops. Comparing (47.15b) with (47.8) and (47.9), we get

$$L_{12} = \frac{\mu_0}{4\pi} \int_{L_1} \int_{L} \frac{d\mathbf{l}_2 \cdot d\mathbf{l}_1}{r_{12}}, \quad L_{21} = \frac{\mu_0}{4\pi} \int_{L_2} \int_{L_1} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r_{21}}. \tag{47.16a}$$

These formulas show that *mutual inductance depends only on geometrical characteristics of the loops and on their mutual arrangement.* Since $d\mathbf{l}_1$ and $d\mathbf{l}_2$ are independent integration variables, the order of integration can be changed. Considering also that $r_{12} = r_{21}$ and $d\mathbf{l}_1 \cdot d\mathbf{l}_2 = d\mathbf{l}_2 \cdot d\mathbf{l}_1$, we conclude that

$$L_{12} = L_{21} \tag{47.16b}$$

In other words, the mutual inductance of the first and the second loop is equal to the mutual inductance of the second and the first loop. Taking this into consideration, we can write

$$L_{12} I_1 \, dI_2 + L_{21} I_2 \, dI_1 = d\left(\tfrac{1}{2} L_{12} I_1 I_2 + \tfrac{1}{2} L_{21} I_2 I_1\right)$$

and, consequently, (47.12) can be represented in the form

$$dA = d\left(\tfrac{1}{2} L_{11} I_1^2 + \tfrac{1}{2} L_{12} I_1 I_2 + \tfrac{1}{2} L_{21} I_2 I_1 + \tfrac{1}{2} L_{22} I_2^2\right). \tag{47.17a}$$

Considering that the work done in increasing the current is equal to the energy of the generated magnetic field, and integrating both sides of Eq. (47.17a) between zero values of current in the loops ($I_1 = 0$, $I_2 = 0$) and the values I_1

www.EngineeringBooksPDF.com

and I_2, we obtain

$$W=\frac{1}{2}(L_{11}I_1^2+L_{12}I_1I_2+L_{21}I_2I_1+L_{22}I_2^2)=\frac{1}{2}\sum_{i=1}^{2}L_{ik}I_iI_k. \quad (47.17b)$$

This formula defines the **energy of the magnetic field** created by currents I_1 and I_2. It can be easily generalized for N loops:

$$W=\frac{1}{2}\sum_{\substack{i=1\\k=1}}^{N}L_{ik}I_iI_k. \quad (47.18)$$

where L_{ik} is called the inductance of the ith loop when $i=k$ and mutual inductance of the ith and kth loops when $i\neq k$. The expression for these coefficients are given by formulas (47.16a) which now assume the form

$$L_{ik}=\frac{\mu_0}{4\pi}\int_{L_i}\int_{L_k}\frac{d\mathbf{l}_i\cdot d\mathbf{l}_k}{r_{ik}} \quad (i\neq k), \quad (47.19)$$

where $d\mathbf{l}_i$, $d\mathbf{l}_k$ are the elements of length of the ith and kth loops L_i and L_k, and r_{ik} is the distance between these elements. It follows from this equation that

$$L_{ik}=L_{ki}, \quad (47.20)$$

which is the generalization of (47.16b) for the case of many current loops.

Magnetic field energy in the presence of magnetics. If the entire space is filled by a homogeneous magnetic, the magnetic induction of the field created by given currents is μ/μ_0 of that in vacuum [see (38.29)]. Consequently, the fluxes Φ and $d\Phi$ appearing in formula (47.1) change by a factor of μ/μ_0, all subsequent calculations being the same. Formulas (47.7) and (47.8) lead to the conclusion that the inductance of the circuit and mutual inductances increase μ/μ_0 times. This means that formulas (47.16a) for mutual inductance in the presence of a magnetic have the same form as in the absence of a magnetic if we replace μ_0 by μ. The same substitution should be made in formulas (47.15a) and (47.15b). Expressions (47.5) and (47.17) for the magnetic field energy remain unchanged, but inductances and mutual inductances in them become μ/μ_0 times larger. Consequently, the *energy of the magnetic field of currents flowing in an unbounded homogeneous magnetic changes by* μ/μ_0 *in comparison with the magnetic field energy of the same currents in vacuum.*

Magnetic energy density. The magnetic field of given currents is distributed over the entire space. Let us express the field energy (47.5) of an isolated current loop in terms of the field vectors. Using formula (47.3), we can write (47.5) in the form

$$W=\tfrac{1}{2}I\Phi. \quad (47.21)$$

Here

$$\Phi=\int_S\mathbf{B}\cdot d\mathbf{S}=\int_S\operatorname{curl}\mathbf{A}\cdot d\mathbf{S}=\int_L\mathbf{A}\cdot d\mathbf{l}, \quad (47.22)$$

www.EngineeringBooksPDF.com

where L and S are the contour loop and the surface stretched over this contour. The potential $\mathbf{A}$ in (47.22) is due to the current I. Thus, the current loop interacts with its own magnetic field. The physical meaning of this interaction consists in that each current element $I\,\mathrm{d}\mathbf{l}$ creates a magnetic field in space, with which other current elements interact. Substituting (47.22) into (47.21), we obtain

$$W=\frac{1}{2}\oint_L I\mathbf{A}\cdot\mathrm{d}\mathbf{l}=\frac{1}{2}\int_V \mathbf{A}\cdot\mathbf{j}\,\mathrm{d}V, \tag{47.23}$$

where a transition to volume currents is made with the help of (9.26). Now we transform the integrand so that it includes only the field vectors and the vector potential. For this purpose, we make use of formulas $\mathbf{B}=\operatorname{curl}\mathbf{A}$ and $\mathbf{j}=\operatorname{curl}\mathbf{H}$, as well as the relation obtained for the vector potential: $\operatorname{div}(\mathbf{A}\times\mathbf{H})=\mathbf{H}\cdot\operatorname{curl}\mathbf{A}-\mathbf{A}\cdot\operatorname{curl}\mathbf{H}$. This gives $\mathbf{A}\cdot\mathbf{j}=\mathbf{H}\cdot\mathbf{B}-\operatorname{div}(\mathbf{A}\times\mathbf{H})$ and hence formula (47.23) assumes the form

$$W=\frac{1}{2}\int \mathbf{H}\cdot\mathbf{B}\,\mathrm{d}V-\int \operatorname{div}(\mathbf{A}\times\mathbf{H})\,\mathrm{d}V. \tag{47.24}$$

In accordance with the Gauss theorem, the second integral is transformed into the integral over the surface bounding the integration volume:

$$\int_V \operatorname{div}\mathbf{A}\times\mathbf{H}\,\mathrm{d}V=\int_S \mathbf{A}\times\mathbf{H}\cdot\mathrm{d}\mathbf{S}. \tag{47.25}$$

If all currents are concentrated in a finite region of space, then at large distances r from this region $A\propto 1/r$ and $H\propto 1/r^2$, i.e. the integrand decreases as $1/r^3$. In this case, the integration surface grows in proportion to r^2, and hence the integral decreases as $1/r$. Consequently, for the entire space, as $r\to\infty$, the second integral in (47.24) vanishes, and the **total energy of the field** is represented by

$$\boxed{W=\frac{1}{2}\int \mathbf{H}\cdot\mathbf{B}\,\mathrm{d}V.} \tag{47.26}$$

It can be said that the field energy is uniformly distributed over the entire space with the volume density

$$\boxed{w=\frac{1}{2}\,\mathbf{H}\cdot\mathbf{B},} \tag{47.27}$$

i.e. the **volume density of the magnetic field energy** at each point is determined by the values of the field vectors at this point. Naturally, it is immaterial by which sources these fields are created.

Inductance. We represent the potential $\mathbf{A}$ in Eq. (47.23) with the help of (37.11a) in the form

$$\mathbf{A}=\frac{\mu}{4\pi}\int \frac{\mathbf{j}'}{r}\,\mathrm{d}V', \tag{47.28}$$

www.EngineeringBooksPDF.com

where the current density and volume density are primed to distinguish them from the same quantities in the integrand of (47.23): these are different volume elements of the same current, the distance between which is denoted in (47.28) by r (see Fig. 183). Substituting (47.28) into (47.23), we get

$$W = \frac{1}{2}\frac{\mu}{4\pi}\int_V\int_V \frac{\mathbf{j}\cdot\mathbf{j}'}{r}\,dV\,dV'$$

$$= \frac{1}{2}I^2\frac{\mu}{4\pi}\frac{1}{I^2}\int_V\int_V \frac{\mathbf{j}\cdot\mathbf{j}'}{r}\,dV\,dV', \quad (47.29)$$

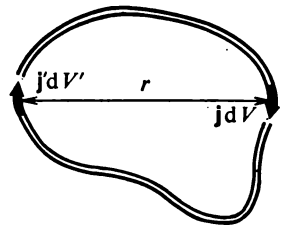


Fig. 183. To the calculation of the self-inductance of a loop

where the numerator and the denominator of the last expression are multiplied by I^2. Comparing (47.29) with (47.5), we obtain

$$L = \frac{\mu}{4\pi}\frac{1}{I^2}\int_V\int_V \frac{\mathbf{j}\cdot\mathbf{j}'}{r}\,dV\,dV'. \quad (47.30)$$

As we go over to volume currents ($I\,d\mathbf{l} \to \mathbf{j}\,dV$), formulas (47.16a) for mutual inductance become

$$L_{ik} = \frac{\mu}{4\pi}\frac{1}{I_iI_k}\int_{V_i}\int_{V_k} \frac{\mathbf{j}_i\cdot\mathbf{j}_k}{r_{ik}}\,dV_i\,dV_k, \quad (47.31)$$

i.e. are similar to (47.30). Formula (47.30), however, cannot be expressed in terms of rectilinear currents. If it is done formally, the integrand in (47.30) assumes the form $I^2d\mathbf{l}\cdot d\mathbf{l}'/r$ and tends to infinity when the elements of integration coincide, viz. when $d\mathbf{l} = d\mathbf{l}'$, since in this case $r = 0$. Hence the integral diverges and the formula for the inductance loses its meaning. This situation is similar to that arising during the calculation of the intrinsic energy of the charge, when the intrinsic energy of the point charge becomes infinite.

The field of a solenoid. By way of an example of application of formulas obtained in this section, let us consider the field of a solenoid. It was shown that the magnetic induction of the field outside a solenoid is equal to zero, while inside the solenoid it is determined by formula (38.40), i.e.

$$B = \mu nI, \quad (47.32)$$

where n is the number of turns per metre of the solenoid length. The magnetic flux through a turn of the solenoid is given by

$$\Phi_1 = BS = \mu nIS, \quad (47.33)$$

where S is the cross-sectional area of the solenoid. The flux through N turns of the solenoid, which fill the length $l = N/n$, is equal to

$$\Phi_N = \Phi_1 N = \mu nISN = \mu ISN^2/l. \quad (47.34)$$

Consequently, the inductance of N turns of the solenoid is given by

$$L_N = \Phi_N/I = \mu SN^2/l. \quad (47.35)$$

www.EngineeringBooksPDF.com

The energy concentrated on the length l is

$$W=\frac{1}{2}L_N I^2=\frac{1}{2}\frac{\mu N^2 I^2}{l}S=\frac{1}{2}\mu n^2 I^2 Sl=\frac{1}{2}HBV. \tag{47.36}$$

where $\mu n^2 I^2 = HB$, $Sl = V$ is the volume of the part of the solenoid, in which the field energy is calculated. Formula (47.36) allows us to determine the field energy in terms of current and inductance as well as in terms of the field energy density.

Let us find the vector potential of the field of a very long solenoid. It is expedient to proceed from formula (47.22). In view of axial symmetry of the problem, we shall make calculations in a cylindrical system of coordinates with the symmetry axis coinciding with the axis of the solenoid. We denote the axial angle by φ and the distance from the axis to the point at which the potential is being calculated by r. For the contour L in (47.22) we choose a circle of radius r, lying in the plane perpendicular to the solenoid axis and with the centre on this axis. Then

$$\Phi=\int\limits_S \mathbf{B}\cdot d\mathbf{S}=\oint\limits_L \mathbf{A}\cdot d\mathbf{l}=\oint\limits_L A_\varphi r\, d\varphi=2\pi r A_\varphi,$$

where we consider that $A_\varphi = \text{const}$ for $r = \text{const}$. Consequently, the vector potential is equal to

$$A_\varphi(r)=\frac{1}{2\pi r}\int\limits_S \mathbf{B}\cdot d\mathbf{S}$$

where S is the area of a circle of radius r. Hence

$$A_\varphi=\begin{cases}\mu nIr/2 & (0<r<a),\\ \mu nIa^2/(2r) & (a<r<\infty).\end{cases}$$

Energy of a magnetic in an external field. Suppose that we have a fixed distribution of currents, which creates in a free space a magnetic field of magnetic induction $B_0(x, y, z) = \mu_0 \mathbf{H}(x, y, z)$ and the energy

$$W_0=\frac{1}{2}\int \mathbf{H}_0\cdot\mathbf{B}_0' dV. \tag{47.37}$$

Suppose that the entire space is filled with a homogeneous magnetic having permeability $\mu = \text{const}$, and the field is created by the same distribution of currents. As was shown earlier [see (38.22)], the magnetic field strength in the magnetic remains unchanged ($\mathbf{H} = \mathbf{H}_0$), while its induction will now be $\mathbf{B} = \mu\mathbf{H}$. Consequently, in the presence of the magnetic the field energy is given by

$$W=\frac{1}{2}\int \mathbf{H}_0\cdot\mathbf{B}\, dV \tag{47.38}$$

This means that the energy of the field increases if the space is filled with a magnetic. The sources of this energy are, for example, extraneous electromotive

www.EngineeringBooksPDF.com

forces required for maintaining current at a constant value after the space has been filled with the magnetic. Since all the sources creating an additional field after filling the space with a magnetic are identical to those which create the field before introducing the magnetic, we can assume that the energy of the magnetic in the external field $\mathbf{H}_0$ is the quantity

$$W_m = W - W_0 = \frac{1}{2} \int (\mathbf{H}_0 \cdot \mathbf{B} - \mathbf{H}_0 \cdot \mathbf{B}_0)\, dV. \tag{47 39}$$

We can transform the integrand in this expression as follows:

$$\mathbf{H}_0 \cdot \mathbf{B} - \mathbf{H}_0 \cdot \mathbf{B}_0 = (\mu - \mu_0) H_0^2 = \frac{\mu - \mu_0}{\mu \mu_0} \mathbf{B} \cdot \mathbf{B}_0 = \mathbf{J}\ \mathbf{B}_0 \tag{47 40}$$

where

$$\mathbf{J} = \chi \mathbf{H} = \frac{\mu - \mu_0}{\mu_0} \frac{\mathbf{B}}{\mu} = \frac{\mu - \mu_0}{\mu \mu_0} \mathbf{B}. \tag{47.41}$$

Consequently, the energy of a magnetic in a magnetic field is

$$\boxed{W_m = \frac{1}{2} \int \mathbf{J} \cdot \mathbf{B}_0\, dV.} \tag{47.42}$$

This expression is similar to formula (18.30) for the energy of a dielectric in an external field, but differs from it in the sign on the right-hand side.

Formula (47.42) was derived for the magnetic which filled the entire space with μ = const. However, this formula has the form of the integral of energy density of the magnetic, and hence it can be expected that it is valid in an arbitrary case. This conclusion can be confirmed by appropriate calculations which are not presented here in view of their cumbersomeness.

Now, we can calculate the energy of a magnetic having a permeability μ_1 placed in a medium of a permeability μ_2. We shall again consider an infinite magnetic and proceed from formula (47.42) as in deriving formula (18.30). The only difference is that in electrostatics the given charge distribution creates in different media the same field D, while in the theory of stationary magnetic field a given current distribution creates in different media the same field $\mathbf{H}$. Thus,

$$W_{m12} = W_{m1} - W_{m2} = \int (\mu_1 - \mu_2)\, \mathbf{H}_1 \cdot \mathbf{H}_2\, dV. \tag{47.43}$$

where

$$W_m = \frac{1}{2} \int (\mathbf{B}_1 \cdot \mathbf{H}_1 \quad \mathbf{B}_0 \cdot \mathbf{H}_0)\, dV. \tag{47.44}$$

Expression (47.43) is similar to formula (18.31) but has the opposite sign in the integral term. Although this formula was derived for an infinite magnetic, it is valid for a bounded magnetic as well. In this case, the integral is extended to the entire volume of the magnetic. The magnetic field strength $\mathbf{H}_2$ is the strength of the field that would be created at the points of the magnetic volume if its permeability were equal to the permeability μ_2 of the medium and $\mathbf{H}_1$

www.EngineeringBooksPDF.com

is the actual magnetic field strength in the magnetic with permeability μ_1, placed in the medium having permeability μ_2.

Suppose that the permeability of the medium changes by an infinitely small quantity $\delta\mu$. Then the energy of the magnetic placed in the magnetic field $\mathbf{H}_1$ changes by δW_m. Putting in (47.43) $\delta\mu = \mu_1 - \mu_2$, $\mathbf{H}_2 = \mathbf{H}$, $\mathbf{H}_1 = \mathbf{H} + \delta\mathbf{H}$ and discarding the quantity $\delta\mu\delta\mathbf{H}\cdot\mathbf{H}$ as a higher order infinitesimal, we obtain

$$\boxed{\delta W_m = \frac{1}{2}\int \delta\mu H^2\, dV,} \tag{47.45}$$

where μ can be a position function and a function of other parameters. This formula differs from the corresponding formula (18.36) for dielectrics only in sign.

Calculation of forces from the expression for energy. Let us consider a system of current loops. When the loops are moved and deformed, mechanical work is performed at the expense of extraneous electromotive forces. The energy of the source of extraneous electromotive forces is spent for generating a magnetic field and performing a mechanical work. The work of extraneous electromotive forces is determined by formula (47.2), while the mechanical work done upon the variation of parameter ξ_i characterizing the configuration of the system is by definition equal to $F_i d\xi_i$, where F_i is the generalized force referred to parameter ξ_i. The law of energy conservation is written in the form

$$\sum_j I_j\, d\Phi_j = dW + \sum_j F_i\, d\xi_i. \tag{47.46}$$

Let us first consider virtual processes in which magnetic fluxes are conserved, i.e. $d\Phi_j = 0$. Equation (47.46) assumes the form

$$0 = (dW)_\Phi + \sum_i F_i\, d\xi_i, \tag{47.47}$$

from which, in view of the independence of $d\xi_i$, we obtain

$$\boxed{F_i = -\left(\frac{\partial W}{\partial \xi_i}\right)_\Phi,} \tag{47.48}$$

where the subscript Φ on the partial derivative shows explicitly that it is taken at constant values of fluxes Φ_j. In order to be able to use formula (47.48), we must express the magnetic field energy as a function of Φ_i and ξ_i as independent parameters.

For practical applications, it is more convenient in many cases to express the generalized force in the form of the derivatives of energy with respect to generalized parameters at constant currents. Considering that [see (47.6)]

$$\Phi_i = \sum_k L_{ik} I_k, \tag{47.49}$$

www.EngineeringBooksPDF.com

the magnetic field energy can be expressed in the form

$$W = \frac{1}{2} \sum \Phi_i I_i. \tag{47.50}$$

For constant currents (I_i = const), this equation gives

$$(dW)_I = \frac{1}{2} \sum_i I_i \, d\Phi_i, \tag{47.51}$$

and hence formula (47.46) is reduced to the form

$$(dW)_I = \sum_i F_i \, d\xi_i. \tag{47.52}$$

It should be noted that this formula is valid only for constant currents. Taking into account the independence ξ_i, we find the expression for generalized forces:

$$\boxed{F_i = \left(\frac{\partial W}{\partial \xi_i}\right)_I,} \tag{47.53}$$

where the subscript I on the partial derivative indicates that it is taken at constant currents. In order to be able to apply (47.53), we should express W as a function of currents and parameters ξ_i.

Let us consider, by way of an example, two coupled current loops for which the magnetic energy is determined by formula (47.17). Using (47.53), let us calculate, for example, the x-component of the force exerted by the first loop on the second one. For the generalized coordinate, we shall take the value of the coordinate x at a certain point of the second loop, assuming the first loop to be fixed. For a virtual displacement associated with this coordinate, we should take the displacement of the second loop along the X-axis without deformations or rotations and express the magnetic energy in terms of this coordinate and other independent parameters in which we are not interested now. The entire dependence of the magnetic energy on x is contained in the mutual inductance $L_{12} = L_{21}$, since the inductances L_{11} and L_{22} do not depend on variation of mutual arrangement of the loops. The generalized force associated with the Cartesian coordinate x is the component F_x of the conventional force. Hence Eq. (47.53) assumes the form

$$F_x = I_1 I_2 \frac{\partial L_{12}}{\partial x}. \tag{47.54}$$

The other components of force are determined in a similar way. The inductance L_{12} is a geometrical quantity, and its dependence on x can be found with the help of formula (47.19).

Clearly, the value of force does not depend on the formula through which it is calculated. Hence the value of force can also be obtained if we proceed from formula (47.48). Let us do this here. We cannot use in (47.48) expression (47.17) for W since it contains currents explicitly. Let us exclude these

www.EngineeringBooksPDF.com

forces with the help of formula (47.10), which give

$$I_1 = \frac{L_{22}\Phi_1 - L_{12}\Phi_2}{L_{11}L_{22} - L_{12}^2}, \quad I_2 = \frac{L_{11}\Phi_2 - L_{21}\Phi_1}{L_{11}L_{22} - L_{12}^2}. \tag{47.55}$$

Substituting these formulas into (47.17), we obtain

$$W = \frac{1}{L_{11}L_{22} - L_{12}^2}\left[\frac{L_{11}\Phi_2^2}{2} - L_{12}\Phi_1\Phi_2 + \frac{L_{22}\Phi_1^2}{2}\right]. \tag{47.56}$$

Now, the magnetic energy is explicitly expressed in terms of fluxes, and we can apply formula (47.48) for $\Phi_i = \text{const}$. The only quantity depending on x in (47.56) is L_{12}, and hence

$$F_x = -\left(\frac{\partial W}{\partial x}\right)_\Phi$$

$$= \frac{1}{(L_{11}L_{22} - L_{12}^2)^2}\left[L_{12}L_{22}\Phi_1^2 - (L_{11}L_{22} + L_{12}^2)\,\Phi_1\Phi_2 + L_{12}L_{11}\Phi_2^2\right]\frac{\partial L_{12}}{\partial x} = I_1 I_2 \frac{\partial L_{12}}{\partial x}, \tag{47.57}$$

where we have taken into account formulas (47.55). As should be expected, (47.57) coincides with (47.54).

Formulas (47.48) and (47.53) should be used depending on the circumstances and the formula involving simpler calculations should be preferred.

Body forces acting on compressible magnetics. Having obtained expression (47.45) for the energy of a magnetic in a magnetic field, we can obtain the expression for forces from the relation between forces and energy in the same way as it was done for dielectrics in Sec. 19. Proceeding from (47.45) and using the same line of reasoning as during the transition from (18.36) to (19.41), we perform similar calculations. It should only be borne in mind that the force for dielectrics is determined at constant charges, i.e. by formula (19.46), and for magnetics, at constant currents, i.e. by formula (47.53). This means that while calculating the derivatives, the energy should be taken with opposite signs. As a result, we obtain the following formula instead of (19.41):

$$\boxed{\mathbf{f} = -\frac{1}{2}H^2 \operatorname{grad}\mu + \frac{1}{2}\operatorname{grad}\left(H^2\rho_m\frac{\partial\mu}{\partial\rho_m}\right).} \tag{47.58}$$

It should be recalled that we are dealing with isothermal processes, and hence the derivative $\partial\mu/\partial\rho_m$ in (47.58) should be calculated at $T = \text{const}$.

It is expedient to write formula (47.58) in a different form:

$$\mathbf{f} = \frac{1}{2}B^2 \operatorname{grad}\left(\frac{1}{\mu}\right) - \frac{1}{2}\operatorname{grad}\left[B^2\rho_m\frac{\partial}{\partial\rho_m}\left(\frac{1}{\mu}\right)\right], \tag{47.59}$$

where we have taken into consideration that $H^2 = B^2/\mu^2$ and $\frac{\partial}{\partial x}\left(\frac{1}{\mu}\right) = -\frac{1}{\mu^2}\frac{\partial\mu}{\partial x}$, etc. In this form, (47.59) is a closer analogue of formula (19.41) since **B** plays the role of a field vector in magnetism and $1/\mu$ is a quantity analogous to ε.

www.EngineeringBooksPDF.com

Let us write formula (47.41) in the form

$$\frac{1}{\mu_0}-\frac{1}{\mu}=\frac{J}{B}. \tag{47.60}$$

Suppose that magnetization J linearly depends on density ρ_m, i.e. $J \propto \rho_m$. Then it follows from (47.60) that

$$\rho_m \frac{\partial}{\partial \rho_m}\left(\frac{1}{\mu}\right)=\frac{1}{\mu}-\frac{1}{\mu_0}. \tag{47.61}$$

Under these conditions, formula (47.59) becomes

$$\boxed{\mathbf{f}=\frac{1}{2}\frac{\mu-\mu_0}{\mu\mu_0}\operatorname{grad} B^2,} \tag{47.62}$$

which coincides with (39.13). Thus, formula (39.13) is valid not only for rigid magnetics but for compressible magnetics as well, for which the magnetization linearly depends on mass density. This is observed for gases and for some liquids.

Energy of a magnetic dipole in an external magnetic field. Since the work required for increasing the magnetic flux through the surfaces stretched over a loop carrying current I is equal to $I\,d\Phi$ ($d\Phi$ is the magnetic flux generated not by the current I in the loop but by other sources of a magnetic field), the energy spent for creating the flux Φ through the surface bounded by the loop with current I is equal to $I\Phi$. For an infinitely small loop magnetic dipole, $\Phi = \mathbf{B}\cdot\mathbf{S}$, $I\Phi = \mathbf{p}_m\cdot\mathbf{B}$. where $\mathbf{p}_m = I\mathbf{S}$ is the magnetic moment of current. Consequently, the **energy of a magnetic dipole in an external magnetic field** is given by

$$\boxed{W=-\mathbf{p}_m\cdot\mathbf{B}.} \tag{47.63}$$

This quantity attains its minimum value when the directions of $\mathbf{p}_m$ and $\mathbf{B}$ coincide. This means that an *external magnetic field tends to rotate the magnetic moment until it coincides with the magnetic induction vector* [see (39.8)].

Why is it possible to calculate the mutual inductance with the help of a formula containing linear currents, while self-inductance cannot be expressed in terms of linear currents?

Due to what property of a magnetic field is the self-inductance of a rigid current loop constant?

The self-inductance and mutual inductance depend only on geometrical characteristics of current loops and their mutual arrangement.

Example 47.1. *Calculate the force with which one solenoid is pulled into or pushed out of another solenoid* (Fig. 184). *The density of the solenoid winding and currents in it are n_1, n_2 and I_1, I_2, respectively, and the cross-sectional areas are the same. The solenoids are sufficiently long and the winding is sufficiently tight so that the field far from the solenoid ends can be described by the formulas for a very long solenoid. The value of x is large, and hence the edge effects can be ignored.*

Let us find the mutual inductance with the help of formulas (47.48) and (47.49). The flux created by the first solenoid through each turn of the second solenoid is equal to $\mu_0 n_1 I_1 S$, and

www.EngineeringBooksPDF.com

the total flux through n_2x turns of the second solenoid in the region of their intersection is

$$\Phi_{21} = \mu_0 n_1 I_1 S n_2 x,$$

whence we obtain the expression for the mutual inductance:

$$L_{21} = \Phi_{21}/I_1 = \mu_0 n_1 n_2 S x \quad (L_{12} = L_{21}) \tag{47.64}$$

Then the force is equal to

$$F_x = I_1 I_2 \frac{\partial L_{12}}{\partial x} = \mu_0 n_1 I_1 n_2 I_2. \tag{47.65}$$

If the currents flow in the same direction, then $I_1 I_2 > 0$, $F_x > 0$, and hence the solenoids repel each other. For oppositely directed currents, $I_1 I_2 < 0$, and $F_x < 0$, which corresponds to the attraction between solenoids.

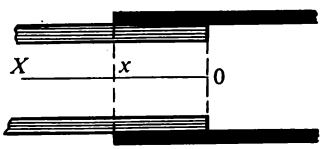


Fig. 184. To the calculation of the force of interaction between two solenoids

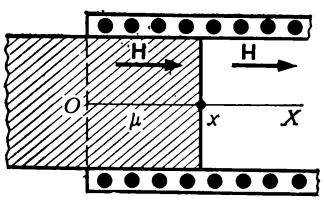


Fig. 185. To the calculation of the force of interaction between a solenoid and a magnet

Example 47.2. *A magnetic having a permeability* μ *is pulled into a solenoid with a circular cross section of area* S, *length* l, *having* n *turns per metre* (Fig. 185). *Find the force acting on the magnetic (ignoring edge effects) if the current in the solenoid is* I.

Since the magnetic susceptibility of the magnetic $\chi \ll 1$, to a first approximation we can assume that the magnetic field strength is everywhere equal to $H_x^{(0)} = H_x = nI$. Consequently, the magnetic energy of the system is given by

$$W = [H_x B_x/2 + H_x^{(0)} B_x^{(0)} (l - x)/2] S,$$

where B_x and $B_x^{(0)}$ are the magnetic inductions in the magnetic and in vacuum respectively. Considering that $B_x = \mu H_x$, $B_x^{(0)} = \mu_0 H_x^{(0)}$, we obtain

$$W = (n^2 I^2/2) [\mu x + \mu_0 (l - x)] S$$

and hence the force is equal to

$$F_x = \left(\frac{\partial W}{\partial x}\right)_I = \frac{1}{2} (\mu - \mu_0) n^2 I^2 S = (w - w_0) S, \tag{47.66}$$

where

$$w = \mu n^2 I^2/2 = H_x B_x/2, \quad w_0 = \mu_0 n^2 I^2/2 = H_x^{(0)} B_x^{(0)}/2 \tag{47.66}$$

is the magnetic energy density on different sides of the boundary on which the force is acting. Thus, the surface density of force $f_x = F_x/S$ is the sum of two forces exerted onto the interface from different sides. The surface density of each force is equal to the magnetic energy density.

Example 47.3. *Calculate the inductance of a coaxial cable of length* l, *whose central core is of radius* r_1 *and the sheath has the internal radius* r_2 *and external radius* r_3 (see Fig. 140). *The permeability of conductors is* μ. *The space between the core and the sheath is filled with a dielectric.*

www.EngineeringBooksPDF.com

Let us first find the magnetic induction. Obviously, the field is axisymmetric and the field lines are the circles with their centres on the cable axis. According to Ampère's circuital law (see Example 35.1), we have

$$B_\varphi(r)=\begin{cases}\dfrac{\mu}{2\pi}\,I\,\dfrac{r}{r_1^2} & (0<r<r_1),\\[2mm] \dfrac{\mu_0}{2\pi}\,\dfrac{I}{r} & (r_1<r<r_2),\\[2mm] \dfrac{\mu}{2\pi}\,\dfrac{I}{r}\,\dfrac{r_3^2-r^2}{r_3^2-r_2^2} & (r_2<r<r_3),\\[2mm] 0 & (r_3<r<\infty).\end{cases} \tag{47.67}$$

In order to calculate the self-inductance of a segment of the cable, we shall use the relation $W=LI^2/2$. Since $W=\frac{1}{2}\int \mathbf{H}\cdot\mathbf{B}\,dV$, we have [see (47.67)]

$$W=\frac{l}{2}\,\frac{\mu I^2}{(2\pi)^2}\int_0^{r_1}\frac{r^2}{r_1^4}\,2\pi r\,dr+\frac{l}{2}\,\frac{\mu_0 I^2}{(2\pi)^2}\int_{r_1}^{r_2}\frac{1}{r^2}\cdot 2\pi r\,dr$$

$$+\frac{l}{2}\,\frac{\mu I^2}{(2\pi)^2}\int_{r_2}^{r_3}\frac{1}{r^2}\left(\frac{r_3^2-r^2}{r_3^2-r_2^2}\right)^2 2\pi r\,dr$$

$$=\frac{l}{2}\,\frac{\mu I^2}{8\pi}+\frac{l}{2}\,\frac{\mu_0 I^2}{2\pi}\ln\frac{r_2}{r_1}+\frac{l}{2}\,\frac{\mu I^2}{2\pi}\left[\frac{r_3^4}{(r_3^2-r_2^2)^2}\ln\frac{r_3}{r_2}-\frac{1}{4}\,\frac{3r_3^2-r_2^2}{r_3^2-r_2^2}\right],$$

whence

$$L=\frac{2W}{I^2}=\frac{1}{2\pi}\left[\mu_0\ln\frac{r_2}{r_1}+\frac{\mu r_3^4}{(r_3^2-r_2^2)^2}\ln\frac{r_3}{r_2}-\frac{\mu r_3^2}{2(r_3^2-r_2^2)}\right]. \tag{47.68}$$

Sec. 48. Quasistationary A.C. Circuits

Basic calculation methods for quasistationary a.c. circuits are described.

Definition. When studying alternating fields and currents, the following two factors should be taken into account:

(1) the finite velocity of propagation of electromagnetic fields (see Sec. 61);

(2) the generation of a magnetic field by a varying electric field.

The quantity $j_d=\partial D/\partial t$ is called the **volume density of displacement current** (see Sec. 57).

At moderate frequencies of a current, these factors can be ignored. In other words, we can assume that electromagnetic waves propagate in space instantaneously and there are no displacement currents, i.e. the magnetic field is generated only by conduction currents. Currents and fields satisfying these condi-

www.EngineeringBooksPDF.com

tions are called **quasistationary**. Let us formulate the quasistationary criteria in an analytical form.

1. If there is a periodic process propagating from the source at a velocity c, the wavelength of this process, i.e. the distance corresponding to the time evolution of the process over a period T, is given by

$$\lambda = cT.$$

The spatial variation of a certain quantity characterizing a process can be neglected only if this quantity is considered in the regions whose linear dimensions l are much smaller than the wavelength ($l \ll \lambda$). This is the *criterion determining the condition under which the finite velocity of propagation of electromagnetic fields can be ignored.*

2. If $D = D_0 \exp(i\omega t)$, then $j_d = \partial D/\partial t = i\omega D = i\omega\varepsilon E$. Hence the effect of the displacement currents can be neglected in comparison with the effect of the conduction currents when

$$|j_d|_{max} \ll |j|_{max}.$$

Since $\mathbf{j} = \gamma\mathbf{E}$ and $\mathbf{j}_d = i\omega\varepsilon\mathbf{E}$, this condition can be written in the form

$$\frac{|j_d|_{max}}{|j|_{max}} = \frac{\omega\varepsilon}{\gamma} \ll 1.$$

Considering that $\gamma \simeq 10^7$ S/m for metallic conductors where $\varepsilon \simeq \varepsilon_0$, we conclude that displacement currents are insignificant in the frequency range

$$\omega \ll \frac{\gamma}{\varepsilon_0} \simeq 10^{18}\ \text{s}^{-1},$$

i.e. right up to the frequencies exceeding those corresponding to the ultraviolet region of the spectrum. This is an approximate estimate since it does not take into account the inertial properties of the medium, which play an essential role at high frequencies. When inertial properties are taken into account, the above estimates become several orders of magnitude lower. However, even after this the frequency range for which displacement currents can be ignored in comparison with conduction currents remains very wide.

However, *for alternating electromagnetic fields in vacuum and in dielectrics, it is necessary to take into account displacement currents as a source of a magnetic field at all frequencies since conduction currents are missing here.* The existence of displacement currents stipulates the existence of electromagnetic waves (see Chap. 9).

As regards the first criterion, *its role is determined by the relative value of wavelength and the spatial dimensions of the region in which the process is investigated.* For example, the wavelength λ corresponding to commercial current at a frequency of 50 Hz is about 6×10^6 m. Consequently, if we are interested in the current distribution over the cables in a power plant or even within a city, we may assume that the current is quasistationary. If, however, the current is transmitted over thousands of kilometers, we must take into account its variation along the transmission line, and then the current cannot be considered

www.EngineeringBooksPDF.com

quasistationary. At very high frequencies corresponding to the wavelengths of several metres, the current cannot be considered quasistationary even within a flat.

Self-inductance. The e.m.f. of induction (46.1) appears upon any variations of a magnetic flux Φ through a current loop. In particular, a closed linear current creates a magnetic flux through the surface bounded by its contour. Consequently, *when the current in a loop changes, an electromotive force is induced in it.* This phenomenon is called **self-induction**. Since the current generates a mag-

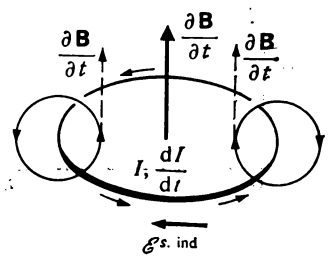


Fig. 186. Emergence of self-induction. Lenz's law

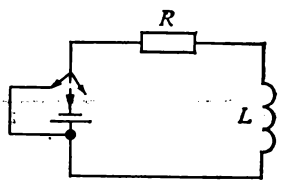


Fig. 187. An *RL*-circuit

netic field in the surrounding space in accordance with the right-hand screw rule, and the electromotive force is related to the flux variation through the left-hand screw rule, we conclude on the basis of Fig. 186 that the e.m.f. of self-induction is directed so as to oppose the variation of current, causing it (**Lenz's law**).

The current in the loop is connected with its own magnetic flux through its contour by formula (47.3):

$$\Phi = LI, \tag{48.1}$$

where L is the inductance of the loop. Hence, formula (46.1) for the e.m.f. of self-induction assumes the form

$$\mathscr{E}^{\text{s. ind}} = -L\frac{dI}{dt}. \tag{48.2}$$

Connection and disconnection of an *RL*-circuit containing a constant e.m.f. If a source of constant extraneous e.m.f., e.g. a battery, is connected to a circuit at the moment $t = 0$ (Fig. 187), the current in the circuit starts growing. However, the self-induction e.m.f. appearing in the circuit due to increasing magnetic induction opposes the extraneous e.m.f. As a result, the increase of current in the circuit slows down. At each instant of time, Ohm's law is observed, which in view of (48.2) can be written in the form of the equation

$$IR = U_0 - L\,dI/dt, \tag{48.3}$$

www.EngineeringBooksPDF.com

where R is the total resistance of the circuit (including the internal resistance of the source). This equation should be solved for the initial condition $I\ (0) = 0$. When we say that Ohm's law is observed at each instant, we assume that current is the same in all parts of the circuit, i.e. the current is quasistationary. The solution of Eq. (48.3) can be found quite easily:

$$I\ (t) = \frac{U_0}{R}\left[1 - \exp\left(-\frac{Rt}{L}\right)\right]. \tag{48.4}$$

The plot of $I\ (t)$ is shown in Fig. 188. The steady-state value of current $I\ (\infty) = U_0/R$, corresponding to Ohm's law for direct current, is attained as a limit for the time going to infinity. Considering the exponential dependence

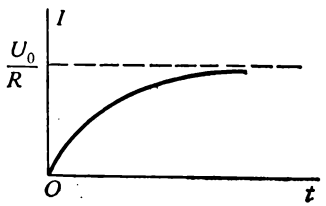


Fig. 188. Increase in the current in a circuit after a constant extraneous e.m.f. is connected

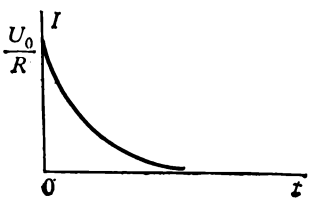


Fig. 189. Decrease in the current in a circuit after a constant extraneous e.m.f. is disconnected

of current on time, we can assume, as usual, that the time during which current in the circuit increases has a value τ for which the exponent becomes equal to -1, i.e.

$$\tau = L/R. \tag{48.5}$$

In circuits with a large inductance the current increases very slowly. For example, if a large induction coil and an incandescent lamp are connected to the circuit, the time during which the lamp attains its full steady glow after closing the circuit is quite long.

When the source of extraneous e.m.f.s is disconnected (see Fig. 187), say, by short-circuiting it, we can see that the current does not immediately drop to zero but decreases gradually. In this case, the equation for current has the form

$$IR = -L\ \mathrm{d}I/\mathrm{d}t \tag{48.6}$$

and must be solved for the initial condition $I\ (0) = U_0/R$:

$$I\ (t) = \frac{U_0}{R}\exp\left(-Rt/L\right). \tag{48.7}$$

The plot of this function is shown in Fig.189. The time of current decrease is given by the same formula (48.5). For a sufficiently large inductance, the incandescent lamp goes out gradually for an appreciable period of time after the extraneous e.m.f. is disconnected. The electromotive force maintaining the

www.EngineeringBooksPDF.com

current in the circuit during this period of time is the e.m.f. of self-induction, and the source of energy is the magnetic energy of the induction coil. The problem of connecting and disconnecting an *RL*-circuit containing an e.m.f. was first considered by Helmholtz in 1855.

Generation of rectangular current pulses. If there is a source of rectangular pulses of voltage, the presence of self-inductance in the circuit hampers the generation of rectangular pulses of current. Current pulses have the form shown

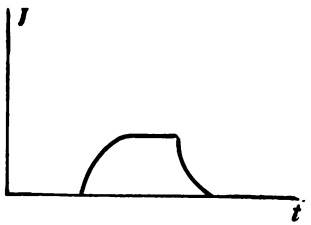


Fig. 190. Shape of a current pulse for a rectangular voltage pulses

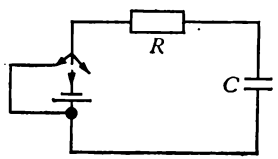


Fig. 191. An *RC*-circuit

in Fig. 190. In order to make their shape as close to rectangular as possible, the inductance of the circuit should be made as small as possible.

***RC*-circuit.** The presence of a capacitor in a circuit makes it impossible for a direct current to pass through it. In this case, the potential difference between the capacitor plates on which appropriate charges are located fully compensates the action of an extraneous e.m.f. However, an alternating current can pass through an *RC*-circuit since in this case the alternating charge on the capacitor plates maintains the current in the circuit. Besides, the potential difference on the capacitor plates does not compensate for the extraneous e.m.f., and a certain current is maintained in the circuit.

Ohm's law for an *RC*-circuit (Fig. 191) is written in the form of the equation

$$IR = U_0 - Q/C, \tag{48.8}$$

where Q is the charge on the capacitor plate and Q/C is the potential difference between the plates. This equation can be conveniently differentiated with respect to t and written as follows:

$$R\frac{dI}{dt} = \frac{dU_0}{dt} - \frac{1}{C}I, \tag{48.9}$$

where $I = dQ/dt$.

Connection and disconnection of an *RC*-circuit containing a constant e.m.f. Suppose that a constant voltage U_0 is applied at a moment $t = 0$. It follows from Eq. (48.8) that $I(0) = U_0/R$, and, for $t > 0$, Eq. (48.9) assumes the form

$$R\frac{dI}{dt} = -\frac{1}{C}I. \tag{48.10}$$

www.EngineeringBooksPDF.com

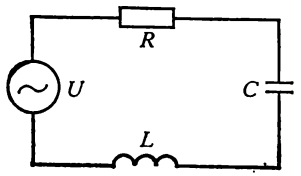


Fig. 192. An *LCR*-circuit containing a source of extraneous e.m.f

The solution of this equation for the initial condition $I(0) = U_0/R$ is given by

$$I(t) = \frac{U_0}{R} \exp[-t/(RC)], \qquad (48.11)$$

i.e. the current in the circuit decreases with time from its maximum value U_0/R to zero. The plot of $I(t)$ is similar to that shown in Fig. 189, and the time of decrease in current is $\tau = RC$. Thus, if the capacitance C is sufficiently large, the current in the circuit may exist for a sufficiently long time after the constant voltage is cut off. An incandescent lamp connected into such a circuit first flares up and then gradually goes out.

After the current has dropped to zero, the capacitor acquires a potential difference equal and opposite to the extraneous e.m.f., so that they compensate each other. When the extraneous e.m.f. is removed, say, by short-circuiting the battery terminals, the potential difference between the capacitor plates is found to be uncompensated. The current appearing in the circuit has the initial value U_0/R, and the law of its decrease is completely identical to (48.11), the time of current decrease being the same.

***LCR*-circuit containing a source of extraneous e.m.f.s.** The diagram of this circuit is shown in Fig. 192. In view of (48.8) and (48.6), the equation for current in this circuit has the form

$$IR = U - L\frac{dI}{dt} - \frac{Q}{C}. \qquad (48.12)$$

Differentiating both sides of this equation with respect to t, we obtain the following equation:

$$L\frac{d^2I}{dt^2} + R\frac{dI}{dt} + \frac{1}{C}I = \frac{d}{dt}U. \qquad (48.13)$$

Various particular cases of the solution to this equation were considered above.

Alternating current. It is most important to analyze a harmonic alternating current, since any other current can be reduced to a harmonic current by using the representation of an arbitrary function in the form of a Fourier's series or Fourier integral.

For further analysis, it is expedient to use the complex form of representation of harmonically varying quantities. We shall be considering steady-state conditions.

If an extraneous e.m.f. varies in accordance with the law

$$U = U_0 e^{i\omega t}, \qquad (48.14)$$

then the current in (48.13) should also vary with time according to the law

$$I = I_0 e^{i\omega t}. \qquad (48.15)$$

www.EngineeringBooksPDF.com

The quantities I, U, I_0 and U_0 in formulas (48.14) and (48.15) are generally complex-valued quantities. It follows from (48.14) and (48.15) that

$$\frac{dU}{dt} = i\omega U, \quad \frac{dI}{dt} = i\omega I, \tag{48.16}$$

and hence Eq. (48.13) assumes the form

$$(-\omega^2 L + i\omega R + 1/C)\, I = i\omega U. \tag{48.17}$$

Dividing both sides of this equation by $i\omega$, we represent it in the form

$$IZ = U, \tag{48.18}$$

where the quantity

$$Z = R + i\,[\omega L - 1/(\omega C)] \tag{48.19a}$$

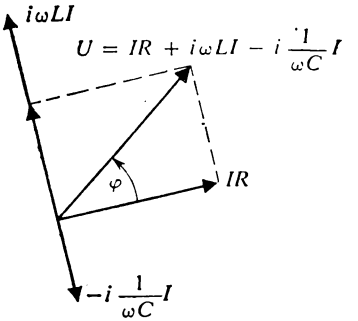


Fig. 193. Vector diagram for the voltage in an a.c. circuit

is called the **impedance.** Equation (48.18) has the form of Ohm's law in which impedance is used instead of resistance. The impedance plays the role of resistance for an alternating current. However, being a complex quantity, it allows us to take into account through (48.18) not only the relation between the amplitudes of current and voltage but the relation between their phases as well.

All the quantities in Eq. (48.18) are complex in the general case. Taking the moduli of both sides of this equation, we can find the relation between the amplitudes of current and voltage:

$$|I|\,|Z| = |U|, \tag{48.19b}$$

where

$$|Z| = \sqrt{R^2 + [\omega L - 1/(\omega C)]^2}. \tag{48.19c}$$

Thus, if we are interested only in the amplitudes of current and voltage, Eq. (48.19b) is completely equivalent to Ohm's law for a direct current. However, the quantity $|Z|$ playing the role of resistance depends, in accordance with (48.19c), on the current frequency.

Vector diagrams. Let us present complex quantities by vectors on a complex plane. A harmonically varying quantity is depicted by a vector rotating at a frequency ω about its origin in the anticlockwise direction. The length of this vector is equal to the amplitude of oscillations of the corresponding physical quantity.

A graphical method of the solution of Eq. (48.18) becomes clear from Fig. 193 if we consider that multiplication of a complex quantity by i corresponds to its rotation through $\pi/2$ in the anticlockwise direction without changing its length, while multiplication by $(-i)$ corresponds to its rotation through $\pi/2$ in the clockwise direction.

Figure 193 shows that the angle φ is determined from the following equation:

$$\tan\varphi = \frac{\omega L - 1/(\omega C)}{R}. \tag{48.20}$$

www.EngineeringBooksPDF.com

Consequently, φ varies between $\pi/2$ and $-\pi l2$ depending on the relation between the impedances of various elements of the circuit and the frequency. The external voltage can change its phase between two extreme positions when it coincides with the voltage on the induction coil and with the voltage across the capacitor. It is more convenient to express this in the form of a relation between the phases of voltages across the elements of the circuit and the phase of the external voltage:

(1) The phase of the voltage across the induction coil ($U_L = i\omega LI$) always leads the phase of the external voltage by an angle lying between 0 and π;

(2) the phase of the voltage across the capacitor [$U_C = -iI/(\omega C)$] always lags behind the phase of the external voltage by an angle lying between 0 and $-\pi$;

(3) The phase of the voltage across the resistor may either lead or lag behind the phase of the external voltage by an angle lying between $+\pi/2$ and $-\pi/2$. It lags behind for a predominantly inductive load, when $\omega L > 1/(\omega C)$, and leads at a predominantly capacitive load when $\omega L < 1/(\omega C)$.

The diagram shown in Fig. 193 also allows us to formulate the following statements concerning the relation between the voltages and currents in different elements of the circuit. It is convenient to use the current as a reference quantity since it has the same phase in all elements of the circuit:

(1) the phase of the voltage across the induction coil leads the phase of the current by $\pi/2$;

(2) the phase of the voltage across the capacitor lags behind the phase of the current by $\pi/2$;

(3) the voltage and current in the resistance coincide in phase;

(4) the external voltage may either lead or lag behind the current in phase, depending on the load.

Kirchhoff's laws. Equation (48.18) makes it possible to solve all problems concerning an alternating current in an LCR-circuit in the same way as the corresponding problems are solved with the help of Ohm's law for a d.c. circuit containing a resistance. The analysis of branched a.c. circuits is similar to that of d.c. circuits (see Sec. 28). Since law (48.19) is valid for an alternating current in a closed circuit, and the law of charge conservation is valid for each junction, Kirchhoff's laws (28.4) and (28.5) for a direct current can be generalized for alternating currents as follows:

(1) *at each junction*,

$$\sum (\pm) I_i = 0; \tag{48.21}$$

(2) *for each closed circuit*,

$$\sum_i (\pm) I_i Z_i = \sum_k (\pm) U_k. \tag{48.22}$$

This generalization of Kirchhoff's laws for branched a.c. circuits was made in 1886 by Lord Rayleigh (J.W. Strutt) (1842-1919). The following remark should be made about the signs of the quantities appearing in (48.21) and (48.22).

www.EngineeringBooksPDF.com

Although each quantity I_i or U_k appearing in these formulas is complex-valued and contains the phase (and hence the sign), appropriate signs should be used while compiling the equations since the same region may belong to different loops and hence is circumvented in opposite directions. A similar remark can be made about the sign of U_k. Solving the equations, we can find the amplitudes and phases of all currents. In view of a complex nature of all quantities, the number of independent equations in this case is twice as large as in the similar case for direct currents.

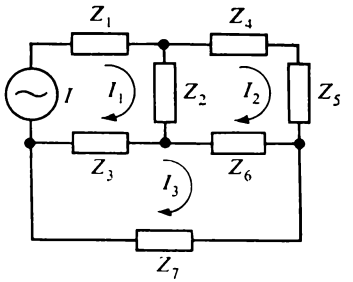


Fig. 194. The mesh-current method

Parallel and series connections of a.c. circuit elements. As for direct currents, it follows from formula (48.18) that for series connection

$$\boxed{Z = Z_1 + Z_2,} \tag{48.23}$$

while for parallel connection we have

$$\boxed{\frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2}.} \tag{48.24}$$

This circumstance makes an analysis of an a.c. circuit similar to that of a d.c. circuit, so there is no need to consider this question in detail.

The reciprocal of impedance is called a.c. **conductance:**

$$Y = 1/Z. \tag{48.25}$$

Thus we can say that for parallel connection a.c. conductances are summed up:

$$Y = Y_1 + Y_2 \tag{48.26a}$$

In terms of conductance, Ohm's law is written in the form

$$I = YU. \tag{48.26b}$$

Mesh-current method. The calculation of branched circuits can be considerably simplified by using the *mesh-current (loop) method which is a direct consequence of Kirchhoff's laws.* A branched circuit consists of a system of simple closed loops. Figure 194 represents a branched circuit consisting of three simple loops. While circumventing a closed loop, in Kirchhoff's law we take the actual current flowing between two junctions. The currents in different branches are generally different. *In the mesh-current method it is assumed that the same current flows in all branches of each closed loop. These currents are called loop (mesh) currents.* The total current passing through a branch of a circuit is equal to the algebraic sum of loop currents for which this branch is common. Kirchhoff's law for each loop is written by taking into account this circumstance, i.e. is

www.EngineeringBooksPDF.com

expressed in terms of loop currents. The total impedance of each branch of the circuit between junctions (Fig. 194) is denoted by the corresponding subscript. The clockwise direction is taken as the positive direction of circumvention.

The equations for the loop currents, whose number coincides with the number of simple loops have the form:

$$\begin{aligned} Z_{11}I_1 + Z_{12}I_2 + Z_{13}I_3 &= U, \\ Z_{21}I_1 + Z_{22}I_2 + Z_{23}I_3 &= 0, \\ Z_{31}I_1 + Z_{32}I_2 + Z_{33}I_3 &= 0, \end{aligned} \tag{48.27}$$

where Z_{11}, Z_{22}, and Z_{33} are intrinsic impedances of the loops, which are equal to the sum of the impedances of the branches of corresponding loops:

$$\begin{aligned} Z_{11} = Z_1 + Z_2 + Z_3, \quad Z_{22} = Z_4 + Z_5 + Z_6 + Z_2, \\ Z_{33} = Z_3 + Z_6 + Z_2, \end{aligned} \tag{48.28}$$

while Z_{12}, Z_{13}, etc. are the mutual impedances of the loops, equal to the impedances of the branches belonging to two loops. Their signs depend on the direction (clockwise or counterclockwise) in which the corresponding branch is passed by the current appearing as a factor of the mutual impedance, relative to the loop current for which the equation is written. For example,

$$Z_{12} = -Z_2, \; Z_{21} = -Z_2, \text{ etc.} \tag{48.29}$$

It can be easily seen that

$$Z_{ij} = Z_{ji}. \tag{48.30}$$

It is almost obvious from what has been said above that Eqs. (48.27) combine the two Kirchhoff's laws. This can be proved more rigorously by obtaining (48.27) from Kirchhoff's laws (48.21) and (48.22), going over to loop currents. We recommend these algebraic calculations as an independent exercise.

The number of equations (48.27) for loop currents is equal to the number of unknown currents. The system of equations is solved by a general rule with the help of the theory of determinants

$$I_1 = U\,(\Delta_{11}/\Delta), \quad I_2 = U\,(\Delta_{12}/\Delta), \quad I_3 = U\,(\Delta_{13}/\Delta), \tag{48.31}$$

where

$$\Delta = \begin{vmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{vmatrix} \tag{48.32}$$

is the determinant of the system; Δ_{11}, Δ_{12}, Δ_{13} are the cofactors of the elements Z_{11}, Z_{12} and Z_{13} in the determinant Δ:

$$\Delta_{11} = \begin{vmatrix} Z_{22} & Z_{23} \\ Z_{32} & Z_{33} \end{vmatrix}, \quad \Delta_{12} = -\begin{vmatrix} Z_{21} & Z_{23} \\ Z_{31} & Z_{33} \end{vmatrix}, \quad \Delta_{13} = \begin{vmatrix} Z_{21} & Z_{22} \\ Z_{31} & Z_{32} \end{vmatrix}. \tag{48.33}$$

Thus, the problem is solved. The generalization of the loop method to an arbitrary number of elementary loops is obvious. We should take care that all

www.EngineeringBooksPDF.com

elementary loops are passed in the same direction and are taken into account in the equations.

The inductance and capacitance characterize the property of a circuit to accumulate energy in the form of electric and magnetic energy. They "smoothen" the curves of variation of current with time in comparison with the corresponding voltage variation curves.

The impedance takes account of not only the ohmic resistance of a circuit but also its inductive and capacitive reactances. Being a complex-valued quantity, the impedance allows us to describe not only the relation between the amplitudes of current and voltage, but also the relation between their phases.

Although electromotive forces and currents in a.c. circuits are represented by complex quantities and hence include phase (and sign), the signs of these quantities should be taken into account while compiling equations since the same branch my belong to different loops and be circumvented in opposite directions.
The mesh-current method assumes that the same current, called the loop current, flows in all branches of each closed loop.
The total current flowing in a branch of the loop in this case is equal to the algebraic sum of loop currents for which this branch is common.

What is the physical meaning of quasistationary criteria?
What determines the signs in the equations expressing Kirchhoff's laws for alternating currents?
What are the advantages of the mesh-current method? When is it expedient to apply this method?

Example 48.1. *Find the self-inductance of n turns of a winding on a toroid of a rectangular cross section, whose inner and outer radii are equal to* r_1 *and* r_2 *respectively, and the height is equal to* a (Fig. 195).

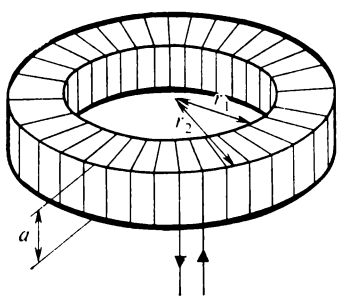


Fig. 195. A toroid with a rectangular cross section

Taking as the path of integration L_0 a circle of radius r, which is concentric with the symmetry axis of the toroid, and applying Ampère's circuital law, we obtain

$$\oint_{L_0} \mathbf{H}\cdot d\mathbf{l} = H_\varphi \cdot 2\pi r = \begin{cases} 0 & \text{for } r < r_1, \\ nI & \text{for } r_1 < r < r_2, \\ 0 & \text{for } r > r_2, \end{cases}$$

where I is the current flowing in the winding of the toroid.

www.EngineeringBooksPDF.com

The magnetic flux through a turn is given by

$$\Phi_1 = \mu a \int_{r_1}^{r_2} H_\varphi \, dr = \frac{\mu a n I}{2\pi} \int_{r_1}^{r_2} \frac{dr}{r} = \frac{\mu a n I}{2\pi} \ln \frac{r_2}{r_1}, \tag{48.34}$$

whence the self-inductance is

$$L = (n\Phi_1/I) = [\mu a n^2/(2\pi)] \, ln \, (r_2/r_1). \tag{48.35}$$

Sec. 49. Work and Power of Alternating Current

The formulas for the work and power of alternating current are derived. Basic physical phenomena associated with the operation of electric motors are discussed.

Instantaneous power. The energy of a source of external e.m.f.s in a circuit with current undergoes the following transformations:

(a) It is transformed into heat as a result of Joule's heating of the conductor [see (27.4)]. If the circuit has a load which performs mechanical work at the expense of the energy of the source of extraneous e.m.f.s, its power is given by a formula similar to (27.4). Hence we assume that the circuit contains only an ohmic resistance R, and the power developed in this resistor is denoted by $P_{tR} = I^2R$.

(b) It is converted into magnetic energy. Since the energy of a magnetic field is defined by formula (47.5), the power developed by the source of extraneous e.m.f.s and leading to a change in the magnetic energy is given by

$$P_{tL} = \frac{dW}{dt} = LI \frac{dI}{dt}. \tag{49.1}$$

The inductive properties of the circuit are characterized by the inductance L. Unlike P_{tR}, the power P_{tL} can be either positive ($dI/dt > 0$) or negative ($dI/dt < 0$). This means that the source of extraneous e.m.f.s gives off energy for increasing the magnetic energy and receives energy when the magnetic energy decreases.

(c) It is transformed into electric energy during electric field variations. The electrical properties of the circuit are characterized by its capacitance C. Since the energy of a capacitor with a charge Q on its plates is given by formula (18.20d), the power of the source of extraneous e.m.f.s required for a change in the electric energy of the field is

$$P_{tC} = \frac{dW}{dt} = \frac{Q}{C} \frac{dQ}{dt} = \frac{Q}{C} I, \tag{49.2}$$

www.EngineeringBooksPDF.com

where $I = \mathrm{d}Q/\mathrm{d}t$ is the current in the circuit. This power can be either positive or negative: as the electric field strength increases, the energy of the source of extraneous e.m.f.s is converted into electric energy, and when the field strength decreases, the electric field energy is transformed into the energy of the source of extraneous e.m.f.s.

The total power developed by the source of extraneous e.m.f.s in a circuit is given by

$$P_t = P_{tR} + P_{tL} + P_{tC}. \tag{49.3}$$

The quantity P_t is often called the power developed by the current, or simply the power of current. We shall, however, use this expression bearing in mind its conditional nature. Similarly, the quantities P_{tR}, P_{tL}, P_{tC} are respectively called the powers of current in a resistor, an induction coil and a capacitor. For the sake of clarity, we shall assume that the ohmic resistance, the inductance and the capacitance are concentrated in different parts of the circuit (see Fig. 192).

The extraneous e.m.f. U is called the **voltage.**

The potential changes across an ohmic resistance by $U_{tR} = IR$. Therefore, U_{tR} is usually called the potential drop across a resistance. The potential difference between the capacitor plates is equal to $U_{tC} = Q/C$. Hence the potential of a capacitor in a circuit changes by U_{tC}. The e.m.f. of self-inductance appearing in an induction coil is $\mathscr{E}^{\mathrm{ind}} = -L\,\mathrm{d}I/\mathrm{d}t$. To compensate this e.m.f., the source of extraneous e.m.f.s spends a certain part of the external e.m.f. ($U_{tL} = L\,\mathrm{d}I/\mathrm{d}t$ is the change in voltage across the induction coil).

Consequently, formulas (49.1) and (49.2) assume the form

$$P_{tL} = U_{tL}\,I, \quad P_{tC} = U_{tC}I. \tag{49.4}$$

In this case [see (49.3)], we obtain

$$P_t = U_{tR}\,I + U_{tL}\,I + U_{tC}I = UI. \tag{49.5}$$

Suppose that the current in a circuit varies in accordance with the law

$$I = I_0 \sin \omega t. \tag{49.6}$$

According to Fig. 193, we can write the following expressions for the true values of U_{tL}, U_{tC} and U_{tR}:

$$U_{tL} = I_0 \omega L \sin (\omega t + \pi/2), \tag{49.7}$$

$$U_{tC} = [I_0/(\omega C)] \sin (\omega t - \pi/2), \tag{49.8}$$

$$U_{tR} = I_0 R \sin \omega t. \tag{49.9}$$

Consequently, instantaneous values of the power developed by current in various circuit elements are defined by the formulas:

$$P_{tL} = I_0^2 \omega L \sin \omega t \sin (\omega t + \pi/2) = I_0^2 \omega L \sin \omega t \cos \omega t, \tag{49.10}$$

$$P_{tC} = [I_0^2/(\omega C)] \sin \omega t \sin (\omega t - \pi/2) = -[I_0^2/(\omega C)] \sin \omega t \cos \omega t, \tag{49.11}$$

$$P_{tR} = I_0^2 R \sin^2 \omega t, \tag{49.12}$$

www.EngineeringBooksPDF.com

which show that the power is always positive *only for resistance R, i.e. the current performs a positive work in this branch. The instantaneous power developed by current in inductance and capacitance is alternating in sign: for a part of time the current performs positive work, i.e. transfers its energy to these elements. During the remaining time the work is negative, i.e. the energy returns from these elements to the source of extraneous e.m.f.s. Thus, the energy is exchanged between inductances, capacitances and the sources of extraneous e.m.f.s. During this process, capacitances and inductances play the role of sources of electromotive forces.*

Mean value of power. In order to calculate the mean value of power developed by the current during a period of oscillations, it is necessary to average the expressions (49.10)-(49.12) over a period of current oscillations. Here we must assume that

$$\langle \sin \omega t \cos \omega t \rangle = 0, \quad \langle \sin^2 \omega t \rangle = 1/2. \tag{49.13}$$

Using these expressions, we obtain from (49.10)-(49.12)

$$P_L = \langle P_{tL} \rangle = 0, \tag{49.14}$$

$$P_C = \langle P_{tC} \rangle = 0, \tag{49.15}$$

$$P_R = \langle P_{tR} \rangle = I_0^2 R/2. \tag{49.16}$$

The mean value of power differs from zero only for resistance R. The mean value of power for inductance and capacitance is equal to zero, i.e. no work is done on these elements by the current. On the average, these elements are neutral from the point of view of energy. Consequently, resistance R is called an active element of the circuit (**resistance**), while inductance and capacitance are called reactive impedances (**reactances**).

Effective (r.m.s.) values of current and voltage. It follows from Fig. 193 that

$$I_0 R = U_0 \cos \varphi, \tag{49.17}$$

and hence formula (49.16) can be written in the form

$$P_R = {}^1/_2 I_0 I_0 R = {}^1/_2 I_0 U_0 \cos \varphi, \tag{49.18}$$

where I_0 and U_0 are the amplitudes of current and external voltage, φ is the phase angle between the current and voltage [see (48.20)], and $\cos \varphi$ is the power factor which determines the effectiveness of power transmission from the source to the load.

For a direct current, the instantaneous power coincides with the mean value of power [see (49.2)]. Since for a direct current $\cos \varphi = 1$, formula (49.18) can be made identical to (27.3) if we write the r.m.s. values instead of the amplitude values I_0 and U_0:

$$I_{\text{eff}} = I_0/\sqrt{2}, \quad U_{\text{eff}} = U_0/\sqrt{2}. \tag{49.19}$$

This gives

$$P_R = I_{\text{eff}} U_{\text{eff}} \cos \varphi. \tag{49.20}$$

The introduction of I_{eff} and U_{eff} allows us to consider formally the a.c. power as if there were no power oscillations. Only the presence of $\cos \varphi$ indicates that we are dealing with alternating current.

www.EngineeringBooksPDF.com

When the values of alternating current and voltage are mentioned in electrical engineering, we mean their effective values. In particular, ammeters and voltmeters are usually graduated in r.m.s. values. For this reason, the maximum value of voltage in an a.c. circuit is almost 1.5 times larger than the value indicated by the voltmeter. This should be borne in mind during calculation of insulators and in the analysis of safety problems.

Power factor. A.c. circuits are mainly designed for transmitting power. Consequently, in designing transmission lines, cos φ should be taken into account.

Suppose that a transmission line has only an ohmic resistance. Then cos $\varphi = 1$, and the power supplied to the load at given values of I_{eff} and U_{eff} is maximum. If a reactive load is connected to the circuit, say, in the form of inductance,

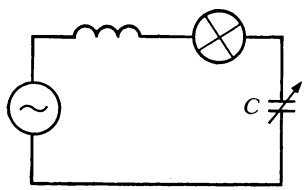


Fig. 196. An increase in the power factor

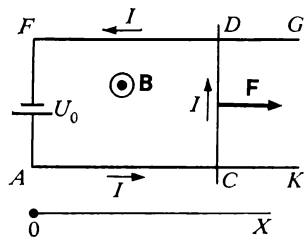


Fig. 197. Operation of a simple motor

cos φ becomes less than unity, and in order to ensure the transmission of the same power, we have to correspondingly increase the value of $I_{\text{eff}}U_{\text{eff}}$, i.e. supply a stronger current to the load. This leads to an increase in energy losses in the form of Joule's heat liberated in the transmission line. For this reason, *it is always desirable to distribute loads in such a way that* $\varphi \simeq 0$, i.e. $\cos \varphi \simeq 1$.

Let us consider, for example, a transmission line for a power supply to an incandescent lamp (Fig. 196), when a large inductance and a variable capacitance are connected in series with the lamp. Suppose that at the initial instant of time the capacitive reactance is equal to zero ($C = \infty$). In this case, for sufficiently large values of $L\omega$ in comparison with resistance R of the lamp, the angle φ attains values close to $\pi/2$, and cos φ is very small. Hence, even if the magnitude of U_{eff} in the circuit is sufficiently high, a very small power is supplied to the lamp, and it glows either dully or does not glow at all. As the capacitance C decreases, the power factor grows (the angle φ decreases, tending to zero) and the glow of the lamp gradually increases. The effective voltage at the terminals of the generator remains unchanged, and the power transmitted from the generator to the line increases. Thus, an increase in the power factor by introducing in the circuit reactive loads which do not consume power makes it possible to improve the efficiency of a transmission line.

Electric motors. A very important application of electric current is the transformation of electric energy into mechanical work, which is realized in electric motors. Their operation is based on the use of Ampère's force acting on a current-carrying conductor in a magnetic field. The first electric motor which

www.EngineeringBooksPDF.com

formed the basis for using electricity to perform mechanical work was constructed in 1839 by B.S. Jacobi (1801-1874).

In order to demonstrate the principle of operation, let us consider a simple d.c. motor (Fig. 197). A source of constant electromotive force U_0 is connected to the circuit $ACDFA$. A rectilinear conductor DC can slide along the conductors FG and AK. It is placed in a uniform magnetic field whose magnetic induction is directed upwards from the plane of the figure. When a current flows in this conductor, it is acted upon by the Lorentz force $F = IlB$. Under the action of this force, the conductor moves and performs a mechanical work, i.e. operates as a motor.

Let us consider the energy balance. The work done during the displacement of the conductor by dx is

$$dA = F\,dx = IlB\,dx, \tag{49.21}$$

and hence the power is given by

$$P_m = dA/dt = IBlv, \tag{49.22}$$

where $v = dx/dt$ is the velocity of the conductor.

On the other hand, the electromotive force induced in the circuit due to motion of the conductor is

$$\mathscr{E}^{ind} = \frac{-d\Phi}{dt} = -\frac{lB\,dx}{dt} = -lBv. \tag{49.23}$$

It is directed against the extraneous electromotive force which generates a current and performs work for overcoming the force (49.23). The power spent by the source of extraneous e.m.f.s is given by

$$P_{ext} = \mathscr{E}^{ind} I = -lBvI. \tag{49.24a}$$

A comparison of formulas (49.24) and (49.22) shows that the *entire power developed by a motor is due to the source of extraneous e.m.f.s.* Besides the useful power (49.22), the source of extraneous e.m.f.s generates power spent as Joule's heat in the ohmic resistance of the wires carrying the current and in the internal resistance of the source. Denoting by R the total ohmic resistance of the wire plus the internal resistance of the source, we obtain the following balance equations for voltages in a closed loop (Kirchhoff's second law):

$$IR = U_0 + \mathscr{E}^{ind} = U_0 - lBv. \tag{49.24b}$$

Multiplying both sides of this equation by I, we obtain

$$I^2R = U_0I - IlBv = U_0I - P_m, \tag{49.25}$$

where expression (49.22) is used. It is expedient to write formula (49.25) in the following final form:

$$P_s = IU_0 = I^2R + P_m, \tag{49.26}$$

i.e. the power developed by a source of extraneous e.m.f. is spent partly as Joule's heat (I^2R) and partly as the work of electric motor (P_m).

www.EngineeringBooksPDF.com

The calculation of energy balance for alternating current is a bit more difficult, but the physical meaning of the phenomenon is the same.

Synchronous motors. To ensure continuous operation of a motor, a certain periodic regime must be chosen. A simple diagram of a synchronous motor is shown in Fig. 197. The magnetic induction in this circuit varies periodically with time.

After conductor CD has moved by a certain distance to the right and performed a certain amount of work, the direction of magnetic induction is reversed. For the same direction of current, the force F also reverses its direction. After that the conductor slows down and starts moving to the left, again performing work, and so on. As a result, we obtain an electric motor whose working part (conductor CD) moves synchronously with a varying magnetic field. Such a

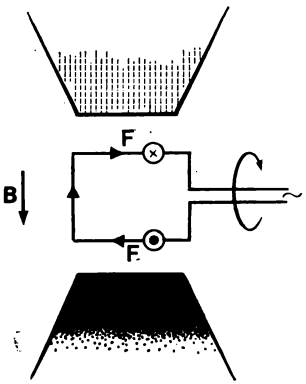


Fig. 198. Operation of a synchronous motor

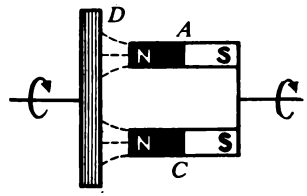


Fig. 199. Emergence of a torque in an asynchronous motor

motor is called **synchronous**. In the above circuit, we can obviously leave the magnetic induction unchanged and vary periodically the direction of current in the moving loop. In this case, the conductor will move synchronously with variations of current in it. Such a motor is also synchronous. We can also vary the induction and the current in the conductor simultaneously. Then the motion of the conductor CD will be synchronized with these variations accordingly.

Synchronous motors used in engineering operate on the basis of the same principle as the simple motor. All three possible types of synchronous motors are used in practice. However, the actual realization of these simple circuits in principle involves quite complicated constructions. As a rule, rotary motion is used in this case.

A simple diagram of the operation of a synchronous motor with a rotary motion is shown in Fig. 198. A loop with an alternating current is in a constant magnetic field. Lorentz forces acting on the conductors of the loop are perpendicular to the magnetic induction and develop a torque which rotates the loop. In order to ensure the action of this torque permanently in the same direction, the speed of rotation of the loop should be equal to the frequency of the alternating cur-

www.EngineeringBooksPDF.com

rent flowing in it, i.e. the synchronism condition should be observed. There exist motors in which the speed of rotation of the loop is an integral number of times less than the frequency of the alternating current supplied to the motor.

The main drawbacks of synchronous motors are the difficulties associated with starting, when the speed of rotation of the loop is synchronized with the frequency of the alternating current, and the possibility of loss of synchronism upon an abrupt change in the load. Methods of overcoming these difficulties have been successfully developed in engineering.

Asynchronous motors. In accordance with Faraday's law, a varying magnetic field generates an electric field [see (46.5)]. If such a vorticity field exists in a conductor, the corresponding electric currents are induced in it. The density of these currents at each point of the conductor is defined by Ohm's law ($\mathbf{j} = \gamma\mathbf{E}$). These currents interact with the magnetic field. Consequently, a varying magnetic field not only induces currents in the conductor but also acts on it with corresponding forces.

Suppose that a varying magnetic field is created by permanent magnets A and C which are fixed to an axle and can rotate around it under the action of an external torque (Fig. 199). A disc D made of a solid conductor is also fixed on an axle and can rotate about it. Moving magnets generate at each point of the disc D an alternating magnetic field, which induces a current of a certain density. The interaction between the current and the magnetic field leads to the appearance of Ampère's force. Thus, rotating magnets exert certain forces on the disc D. Let us calculate the resultant of these forces. In accordance with Lenz's law, the currents induced in a conductor in accordance with Faraday's law tend to weaken the action of the factors that cause these currents. In the case under consideration, the factor responsible for the currents induced in the disc D is the relative motion of the magnets and the disc. This means that the torque applied to the disc tends to rotate it in the direction in which the magnets are rotating. Consequently, the disc starts rotating in the direction in which the magnets rotate, as if entrained by the rotating field of the magnets. The torque exists only until the speed of rotation of the magnets differs from the speed of rotation of the disc, i.e. there is a "slip" between the rotating magnetic field and the disc. The smaller the slip, the smaller the torque acting on the disc. Therefore, the slip increases with the load on the axle of the disc. At a constant speed of rotation of the magnetic field and its magnetic induction, this leads to a decrease in the speed of rotation of the disc.

This principle of setting a disc into motion is the basic feature of the operation of asynchronous motors. However, in order to be able to call this motor an electric motor, it is necessary to ensure the rotation of the magnetic field without a mechanical drive. For this purpose electromagnets fed by alternating current are used.

Generation of rotating magnetic field. Two electromagnets creating crossed magnetic fields (Fig. 200) are fed by alternating currents with the phase difference of $\pi/2$. In the circuit shown in Fig. 200, this is achieved by including electromagnets of inductance L and resistance R into the circuits of electromagnets.

www.EngineeringBooksPDF.com

As a result, two varying magnetic fields are generated in the space between the poles of the electromagnets. The magnetic inductions of these fields with a phase difference close to $\pi/2$ obey a harmonic law. The sum of the magnetic inductions B_1 and B_2 of these fields is a vector **B** rotating about the point O (Fig. 201).

If we place a bulky conductor in the space between the magnets (see Fig. 200), e.g. a cylinder with the axis of rotation perpendicular to the plane of the figure, it will be set in rotation in the rotating field so that its direction of rotation coincides with that of the field. The physical processes occurring in this case

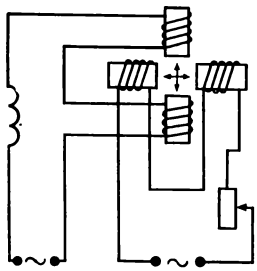
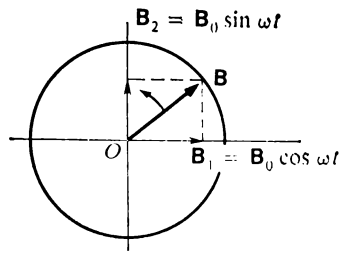


Fig. 200. Schematic diagram for generating a rotating magnetic field



g. 201. Addition of two mutually perpendicular harmonic oscillations with the phase difference of $\pi/2$

are similar to those taking place when the field is created by rotating permanent magnets. Squirrel-cage rotors are used instead of a solid cylinder (Fig. 202).

It is much more convenient to generate a rotating magnetic field with the help of a three-phase current, since in this case there is no need to artificially create the phase difference between currents that feed different electromagnets (see Sec. 52).

Clearly, the speed of rotation of an asynchronous motor may continuously vary and there is no multiple proportion between this speed and the frequency of the supplied current. For this reason, such a motor is called **asynchronous**, and the possibility of a continuous variation of its speed of rotation is a significant advantage of this motor.

The current in the windings of an electromagnet depends on the slip: the larger the slip, the stronger the current. Hence, when a motor is started and the slip is maximum, a very strong current passes through motor windings, which may damage them. In order to avoid this, a starter rheostat which at the starting moment is set at a sufficiently high resistance is included into the feed circuit. As the speed of rotation of the motor increases, the resistance of the rheostat is decreased.

As in the case of synchronous motors, practical realization of asynchronous motors is a complicated technical problem. However, even in the most complicated constructions the basic principles of operation remain the same.

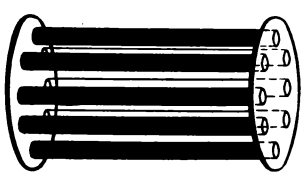


Fig. 202. A squirrel-cage rotor of an asynchronous motor

www.EngineeringBooksPDF.com

Asynchronous motors may operate not only with a rotating magnetic field but with a pulsed field as well. This becomes obvious if we take into account the fact that a pulsed field is equivalent to two fields rotating in opposite directions. One field ensures the rotation of the rotor of an asynchronous motor, while the field rotating in the opposite direction produces, on the average, no effect on the rotor motion.

Load matching with a generator. An a.c. generator producing an e.m.f. has a certain internal resistance, capacitance and inductance, i.e. has a certain impedance

$$Z_g = R_g + iX_g, \tag{49.27}$$

where R_g is the resistance and X_g is the reactance which is equal to the difference between the inductive and capacitive reactances. The load fed by the generator is also characterized by the impedance

$$Z_l = R_l + iX_l. \tag{49.28}$$

The power is liberated only in the resistance R_l. The generator and the load are series-connected into the circuit, U_g being the electromotive force of the generator.

The power developed in the load R_l is expressed in accordance with formula (49.16):

$$P_l = {}^1/_2 I_0^2 R_l, \tag{49.29}$$

where I_0^2 is the square of the current amplitude in the load. On the basis of (48.19b), we have

$$I_0^2 = |I|^2 = \frac{|U_g|^2}{|Z_g + Z_l|^2} = \frac{|U_g|^2}{(R_g + R_l)^2 + (X_g + X_l)^2}. \tag{49.30}$$

Using this expression, we write formula (49.29) as follows:

$$P_l = \frac{|U_g|^2}{2} \frac{R_l}{(R_g + R_l)^2 + (X_g + X_l)^2}. \tag{49.31}$$

Let us find the conditions under which this power attains its maximum.

Reactances X_g and X_l can assume positive or negative values. Clearly, the expression (49.31) assumes the maximum value only if

$$\boxed{X_g + X_l = 0.} \tag{49.32}$$

This means that the power factor should have the maximum value ($\cos \varphi = 1$). When condition (49.32) is satisfied, expression (49.31) becomes

$$P_l = \frac{|U_g|^2}{2} \frac{R_l}{(R_g + R_l)^2}. \tag{49.33}$$

The power varies with the resistance and attains the maximum value when $\partial P_l / \partial R_l = 0$, i.e. if

$$\boxed{R_l = R_g.} \tag{49.34}$$

www.EngineeringBooksPDF.com

If conditions (49.32) and (49.34) are observed, the *generator delivers the maximum power to the load. In this case it is said that the load is completely matched with the generator.*

Considering (49.34), we can write the expression for the maximum power developed in the generator load:

$$P_{l\max} = \frac{|U_g|^2}{2} \frac{1}{4R} = \frac{\langle U_0^2 \rangle}{4R}, \tag{49.35}$$

where $\langle U_0^2 \rangle$ is the mean square amplitude of the generator voltage.

The matching between the generator and the load is always very important when it is necessary to supply the maximum power to the load. For example, an input resistance of a radio receiver should be matched with the resistance of the aerial (generator) and of the transmission line (see Sec. 54).

Foucault (eddy) currents. *The currents induced in bulky conductors in an alternating magnetic field are called Foucault or eddy currents.* Sometimes they are useful and sometimes harmful.

Foucault currents play a useful role in the rotor of an a.c. motor driven by a rotating magnetic field, since the very operation of an a.c. motor requires the emergence of eddy currents. Being conduction currents, eddy currents partially dissipate energy as Joule's heat. This energy loss in the rotor of an a.c. asynchronous motor is useless but we have to put up with it, just avoiding overheating of the motor. However, along with these currents, Foucault currents also appear in the cores of electromagnets of the asynchronous motor, which are usually made of ferromagnetics (conductors). These currents are of no importance for the operation of electromagnets, but they heat the cores, thus deteriorating their characteristics. These currents should be eliminated as harmful factors. This can be done by manufacturing the cores in the form of thin plates separated by layers of insulators, arranged in such a way that eddy currents cross the plates. Owing to this, eddy currents cannot develop in sufficiently thin plates and have an insignificant volume density.

Joule's heat liberated by eddy currents is used in heating and even in melting metals, when it is advantageous and expedient as compared to other methods of heating. If a metal is heated by high-frequency currents only the surface layer of the conductor is heated on account of skin effect (see Sec. 53).

The instantaneous power developed by the current in inductances and capacitances may be alternating in sign, while the power developed in resistances is always positive.

Sec. 50. Resonances in A.C. Circuits

Resonances in an a.c. circuit and properties of an oscillatory circuit are discussed.

Voltage resonance. Let us consider a circuit in which R, L and C are connected in series with a generator (see Fig. 192), and determine the frequency dependence of the amplitude of current, I_0, and the phase difference φ between the cur-

www.EngineeringBooksPDF.com

rent and the external voltage. On the basis of (48.18) and (48.20), we have

$$I_0 = \frac{U_0}{\sqrt{R^2 + [L\omega - 1/(\omega C)]^2}}, \tag{50.1}$$

$$\tan\varphi = \frac{\omega L - 1/(\omega C)}{R}. \tag{50.2}$$

The dependences $I_0(\omega)$ and $\varphi(\omega)$ are plotted in Figs. 203 and 204. The maximum value of current I_0 is attained at the frequency

$$\omega_0 = 1/\sqrt{LC}, \tag{50.3}$$

called the **resonance frequency of the circuit.** In this case, the current amplitude is equal to U_0/R and the phase difference $\varphi = 0$. In other words, the circuit

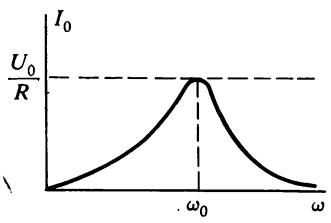


Fig. 203. Frequency dependence of current at voltage resonance

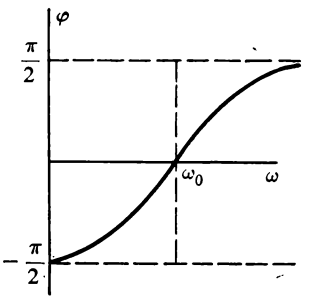


Fig. 204. Frequency dependence of the phase shift φ at voltage resonance

as if does not contain a capacitor or an induction coil. Thus, at this frequency the voltages across the capacitor and the induction coil completely compensate each other, being equal in value (but always opposite in phase). Hence this resonance is also called the **voltage resonance**. The vector diagram of voltage resonance is shown in Fig. 205. At resonance ($\omega = \omega_0$) the circuit behaves as a pure resistor.

If a current of constant frequency ω is passed through the circuit, a change, say, in the inductance, also causes a resonant variation of I_0. The maximum value of I_0 is attained at $L = 1/(\omega^2 C)$ [see (50.1) and (50.3)]. If an incandescent lamp is included in the circuit, its brightness increases as the resonance is approached, attains its maximum value at resonance and then decreases.

Current resonance. Let us consider the circuit shown in Fig. 206. Obviously, the current flowing through the circuit is

$$I = I_L + I_C = U\left(\frac{1}{R + i\omega L} + i\omega C\right) = U\left(\frac{R - i\omega L}{R^2 + \omega^2 L^2} + i\omega C\right)$$
$$= U\frac{R}{R^2 + \omega^2 L^2} - i\frac{U}{R^2 + \omega^2 L^2}[\omega L - \omega C(R^2 + \omega^2 L^2)]. \tag{50.4}$$

www.EngineeringBooksPDF.com

Consequently, under the condition

$$\omega L - \omega C\,(R^2 + \omega^2 L^2) = 0, \tag{50.5}$$

the circuit behaves as a purely ohmic resistance. The phase difference between the external voltage and the current is equal to zero. Dividing all terms of Eq. (50.5) by $\omega^2 LC$, we can write it in the form

$$\frac{1}{\omega C} - \omega L = \frac{R^2}{\omega L}. \tag{50.6}$$

In most cases of practical importance, the condition $\omega L \gg R$ is satisfied. Hence the solution of Eqs. (50.6) and (50.5) can be written as follows:

$$\omega_0 = 1/\sqrt{LC}. \tag{50.7}$$

At this resonance frequency, the impendance attains its maximum value, while the current I_0 in the circuit has its minimum value. Currents I_L and I_C,

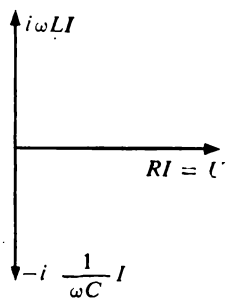


Fig. 205. Vector diagram of voltage at resonance

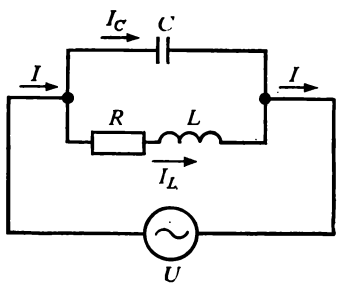


Fig. 206. Circuit with current resonance

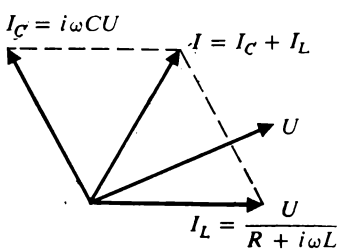


Fig. 207. Vector diagram of currents in a circuit with a capacitor and an induction coil connected in parallel

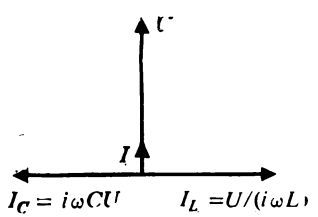


Fig. 208. Vector diagram of currents at resonance

however, are not at their lowest values in this case. A vector diagram of currents in the circuit is shown in Fig. 207. As resonance conditions are approached, the current diagram assumes the form shown in Fig. 208. Thus, much stronger currents flow in the *RLC*-circuit as compared to the currents supplied to this

www.EngineeringBooksPDF.com

circuit. The charge in the RLC-circuit passes from the capacitor to the induction coil and vice versa, i.e. the current oscillates in this circuit. It can be seen from Fig. 208 that the currents in the capacitor and in the induction coil are at resonance and are mutually compensated. Hence this resonance is called the **current resonance**.

Oscillatory circuit. In both the cases considered above, the circuit shown in Fig. 192 behaves as a resonance system performing forced oscillations under the action of an external force. The oscillations of current in an LC-circuit were first investigated by Thomson, who in 1853 obtained formula (50.7). This formula was later termed **Thomson's formula** ($T = 2\pi \sqrt{LC}$). To analyze the current oscillations in a circuit, we can use the results of the theory of forced mechanical vibrations of a material point. For this purpose, it is necessary to find the quantities in electric oscillations corresponding to force, displacement and velocity in mechanical vibrations. Let us write the equation for forced mechanical vibrations:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = F/m, \tag{50.8}$$

where x is the displacement of the point from its equilibrium position, m its mass, F the external force, $\gamma = b/(2m)$ the damping factor, and b the coefficient of friction. The dots over x indicate derivatives with respect to time.

Let us now transform Eqs. (48.12) and (48.13) for an electric circuit. Considering that $I = \mathrm{d}Q/\mathrm{d}t$, we can write Eq. (48.12) in the form

$$L\frac{\mathrm{d}^2Q}{\mathrm{d}t^2} + R\frac{\mathrm{d}Q}{\mathrm{d}t} + \frac{1}{C}Q = U. \tag{50.9}$$

Dividing both sides by L, we obtain the equation

$$\ddot{Q} + (R/L)\dot{Q} + [1/(LC)]Q = U/L, \tag{50.10}$$

which is analogous to Eq. (50.8). The role of displacement in an electric circuit is played by charge Q on capacitor plates, the inductance L serves as the mass, the e.m.f. U replaces the force, while the coefficient of friction is replaced by the ohmic resistance R. The natural frequency of the oscillatory circuit is equal to $\omega_C = 1/\sqrt{LC}$ [see (50.3)]. Current $I = \mathrm{d}Q/\mathrm{d}t$ plays the role of velocity. Since in mechanical vibrations of a point we usually consider its deviation from the equilibrium position, amplitude of vibrations, etc., it is more convenient to use Eq. (50.10) instead of (48.13) while considering an electric circuit. Besides, instead of charge Q on the capacitor plates it is more expedient to use the voltage $U_C = Q/C$ across the capacitor. In terms of this quantity, Eq. (50.10) assumes the form

$$\ddot{U}_C + 2\gamma\dot{U}_C + \omega_0^2 U_C = \omega_0^2 U,$$

where $\gamma = R/(2L)$ and $\omega_0 = 1/\sqrt{LC}$. All properties of these oscillations are obtained from a direct comparison of the quantities γ, ω_0, U and U_C of an electric oscillatory circuit with the corresponding quantities characterizing mechan-

www.EngineeringBooksPDF.com

ical vibrations of a point. In the absence of a resistance ($R = 0$), the natural frequency of the oscillatory circuit is equal to $\omega_0 = (LC)^{-1/2}$. These oscillations are undamped. In the presence of an ohmic resistance in the circuit, the oscillations become damped, and the damping time is

$$\tau_d = 1/\gamma = 2L/R. \tag{50.12}$$

For the frequency of damped oscillations, it is customary to take the frequency

$$\Omega = \sqrt{\omega_0^2 - \gamma^2}. \tag{50.13}$$

The **damping factor (logarithmic decrement)** is

$$\Theta = \gamma T, \tag{50.14}$$

where $T = 2\pi/\omega_0$ is the period of natural oscillations.

The amplitude resonance curves and the phase resonance curves are similar to the corresponding curves for mechanical vibrations.

The ***Q*-factor** is defined by the relation

$$Q = \frac{U_{C\,\mathrm{res}}}{U_{C\,\mathrm{stat}}} = \frac{U_{C0\,\mathrm{res}}}{U_0} = \frac{\omega_0}{2\gamma} = \frac{\omega_0 L}{R} = \frac{1}{R}\sqrt{\frac{L}{C}}, \tag{50.15}$$

where $U_{C0\,\mathrm{res}}$ is the amplitude of voltage across the capacitor at resonance and U_0 is the amplitude of the external e.m.f. applied to the circuit. Thus, in a high-quality circuit, the amplitude of oscillations of the capacitor voltage may be much larger than the amplitude of voltage applied to the circuit.

The width of the resonance curve is

$$2\Delta\omega = \omega_0/Q = R/L. \tag{50.16}$$

It should be recalled that the *width* $2\Delta\omega$ *of the resonance curve is defined not with respect to the amplitude of oscillations, but to the square of the amplitude.*

Sec. 51. Mutual Inductance Circuits

The basic methods for calculating circuits are discussed. The operation of a transformer is described.

Mutual inductance. Any a.c. circuit is a source of a varying magnetic field. In accordance with Faraday's law of electromagnetic induction, electromagnetic forces are induced by this field in other loops lying within its range, thus changing the current in these loops. Thus, the *loops are mutually connected through electromagnetic induction.*

Equation for a system of conductors taking into account the self-inductance and mutual inductance. The total magnetic flux through the kth loop is defined

www.EngineeringBooksPDF.com

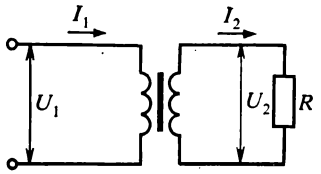


Fig. 209. Transformer

by the expression

$$\Phi_k = \sum_{i=1}^{N} L_{ki} I_i, \tag{51.1}$$

which is a generalization of formulas (47.6) and (47.10) to the case of several current loops in accordance with the principle of superposition. Here L_{kk} is the inductance of the kth loop, while L_{ki} (for $k \neq i$) is the mutual inductance of the kth and ith loops. The total number of conductors is equal to N.

For the sake of simplicity, we assume that there are no capacitors in the circuits. Then, taking into account the electromagnetic induction for current in the ith loop, we obtain the following equation:

$$I_k R_k = U_k - d\Phi_k/dt, \tag{51.2}$$

where U_k is the extraneous e.m.f. in the kth loop. Substituting (51.1) into (51.2) we obtain the following system of equations for determining the current in all loops:

$$I_k R_k = U_k - \sum_{i=1}^{N} L_{ki} \frac{dI_i}{dt} \qquad (k = 1, 2, \ldots, N). \tag{51.3}$$

This linear system of N equations in N unknown currents I_k is complete and can, in principle, be easily solved. The only nontrivial problem is to determine the mutual inductances and inductances of the loops. These quantities are assumed to be given in Eq. (51.3).

The case of two loops. By way of an example, let us consider the system of equations for two loops:

$$I_1 R_1 = U_1 - \left(L_{11} \frac{dI_1}{dt} + L_{12} \frac{dI_2}{dt} \right), \tag{51.4}$$

$$I_2 R_2 = U_2 - \left(L_{21} \frac{dI_1}{dt} + L_{22} \frac{dI_2}{dt} \right), \tag{51.5}$$

where L_{11} and L_{22} are inductances of the first and second loops respectively, and L_{12}, L_{21} are the mutual inductances of the loops.

Further calculations can be considerably simplified if we consider the operation of an a.c. transformer (Fig. 209).

Transformer. A transformer has two conductors in the form of coils wound on a closed core made of a high-permeability material. Owing to this circumstance, the magnetic fluxes created by the currents flowing through the conductors are concentrated almost completely inside the core. The conductors are called the **transformer windings**. The winding to which the extraneous e.m.f. source is connected is called the **primary** winding (or primary), while the winding to which the load is connected is called the **secondary** winding (or secondary).

www.EngineeringBooksPDF.com

Let us mark the quantities corresponding to the primary and secondary by subscripts 1 and 2 respectively. Equations (51.2) can be written in the form

$$I_1R_1 = U_1 - d\Phi_1/dt, \tag{51.6}$$

$$I_2R_2 = -d\Phi_2/dt, \tag{51.7}$$

where R_1 is the ohmic resistance of the primary, R_2 is the sum of the ohmic resistances of the secondary and the load which is assumed to be purely resistive for the sake of simplicity, Φ_1 and Φ_2 are the total magnetic fluxes through the primary and secondary respectively, and U_1 is the extraneous e.m.f. applied to the primary.

The resistance R_1 of the primary winding is quite small and the voltage drop across it due to the ohmic resistance can be assumed to be much smaller than U_1, i.e. $I_{10}R_1 \ll U_{10}$, where I_{10} and U_{10} are the amplitudes of current and voltage in the primary. Hence we can neglect the product I_1R_1 as compared to V_1 in (51.6) and write it in the form

$$U_1 = d\Phi_1/dt. \tag{51.8}$$

Under normal conditions, the ohmic resistance of the load is much higher than the ohmic resistance of the secondary. Hence, to a great degree of accuracy, R_2 in (51.7) is equal to the load resistance. Consequently, I_2R_2 on the left-hand side of (51.7) is equal to the voltage U_2 across the terminals of the secondary of the transformer. Hence we can write (51.7) in the following form:

$$U_2 = -d\Phi_2/dt. \tag{51.9}$$

Since the extraneous e.m.f. varies exponentially [$U_1 \propto \exp(i\omega t)$], all quantities vary according to the same law. Consequently, $d\Phi_1/dt = i\omega\Phi_1$ and $d\Phi_2/dt = i\omega\Phi_2$. Since the entire magnetic flux is concentrated in the core, each turn of the primary and secondary is pierced by the same magnetic flux Φ_0. Hence the magnetic fluxes through the primary and secondary are given by

$$\Phi_1 = \Phi_0N_1, \tag{51.10}$$

$$\Phi_2 = \Phi_0N_2, \tag{51.11}$$

where N_1 and N_2 denote the number of turns in the primary and secondary respectively. Taking into account (51.10) and (51.11), we can write (51.8) and (51.9) in the form

$$U_1 = i\omega N_1\Phi_0, \tag{51.12}$$

$$U_2 = -i\omega N_2\Phi_0. \tag{51.13}$$

Dividing the left- and right-hand sides of (51.12) and (51.13) termwise and taking the moduli, we obtain

$$|U_1|/|U_2| = N_1/N_2. \tag{51.14}$$

Considering that $|U_1| = U_{10}$ and $|U_2| = U_{20}$ are the voltages in the primary and secondary, we can write (51.14) in the form

$$U_{10}/N_1 = U_{20}/N_2, \tag{51.15}$$

www.EngineeringBooksPDF.com

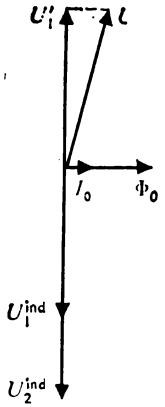


Fig. 210. Vector diagram of a transformer at no-load

i.e. the ratio of voltages in the secondary and in the primary is equal to the ratio of the number of turns in the secondary to the number of turns in the primary.

Neglecting the energy losses in the transformer, we can write the law of conservation of energy in the form

$$I_1U_1 = I_2U_2. \tag{51.16}$$

Taking the moduli, we obtain on the basis of (51.15) the following relation:

$$I_{10}N_1 = I_{20}N_2, \tag{51.17}$$

where I_{10} and I_{20} are the currents in the primary and secondary respectively.

Formulas (51.15) and (51.17) describe the law of transformation of voltages and currents in a transformer. They are strictly valid for an ideal transformer in which there is no dissipation of the magnetic flux and no energy losses. For a real transformer, these relations are satisfied to a considerably high degree of accuracy.

Vector diagram of a transformer at no-load. A transformer operates at no-load when the secondary is open. We shall neglect the phase-lag of the magnetic flux in comparison with the phase of the current in the primary on account of a certain hysteresis in the reversal of magnetization of the core material. This lag is negligibly small, and hence we can assume that the magnetic flux is in phase with the current in the primary, called the no-load current. The current in the secondary is equal to zero. From the formula

$$U^{ind} = -\,d\Phi/dt, \tag{51.18}$$

it follows that U^{ind} lags behind the flux Φ by $\pi/2$. Hence the vector diagram of a transformer at no-load has the form shown in Fig. 210: U_1 is the external voltage applied to the primary, U_1^{ind} is the voltage in the primary due to self-induction, U_2^{ind} is the voltage in the secondary due to mutual induction, I_0 is the no-load current, and Φ_0 is the magnetic flux per turn of the transformer windings. As before, the losses and dissipation of the magnetic flux in the transformer are neglected.

According to the law of electromagnetic induction, we have

$$U_1^{ind} = -\frac{d\Phi_0}{dt} N_1, \tag{51.19}$$

$$U_2^{ind} = -\frac{d\Phi_0}{dt} N_2, \tag{51.20}$$

since the total magnetic fluxes through the primary and secondary windings are given by

$$\Phi_1 = \Phi_0 N_1, \quad \Phi_2 = \Phi_0 N_2. \tag{51.21}$$

www.EngineeringBooksPDF.com

It must be borne in mind that the no-load current is very small, like the ohmic resistance of the primary winding in comparison with its inductive reactance. Hence (see Fig. 210)

$$U_1 \approx U_1' \approx -U_1^{\text{ind}}, \qquad (51.22)$$

i.e.

$$U_1^{\text{ind}} \approx -U_1. \qquad (51.23)$$

Dividing termwise the left- and right-hand sides of Eq. (51.20) by the corresponding sides of Eq. (51.19) and taking into account (51.23), we get

$$| U_2^{\text{ind}} | / | U_1 | \approx N_2/N_1. \qquad (51.24)$$

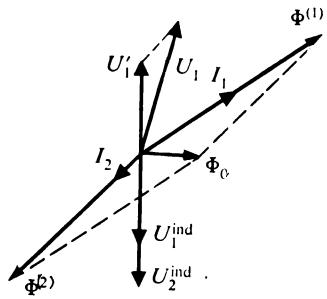


Fig. 211. Vector diagram of a transformer with a load containing an inductive reactance element

Vector diagram of a transformer under load. In a transformer under load, the magnetic flux Φ_0 through each turn of the windings is created by the primary as well as by the secondary. The e.m.f. of self-induction in the primary must always compensate the external voltage, i.e. the sum of magnetic fluxes $\Phi^{(1)}$ and $\Phi^{(2)}$ created by the currents in the primary and secondary must be approximately equal to the no-load flux Φ_0 or, in other words, $\Phi_0 = \Phi^{(1)} + \Phi^{(2)}$. This means that the voltage in the secondary will also satisfy condition (51.24) for the transformer under load.

It should be noted that the magnetic fluxes $\Phi^{(1)}$ and $\Phi^{(2)}$ are not the total fluxes Φ_1 and Φ_2 through the primary and secondary. Actually, $\Phi^{(1)}$ and $\Phi^{(2)}$ are the magnetic fluxes through a turn of each winding, created in the core by the currents I_1 and I_2 respectively. The total magnetic fluxes through the primary and secondary are given by $\Phi_1 = N_1 (\Phi^{(1)} + \Phi^{(2)})$, $\Phi_2 = N_2 (\Phi^{(1)} + \Phi^{(2)})$.

The vector diagram of a transformer under load is shown in Fig. 211. The currents I_1 and I_2 are much stronger than the no-load current I_0. Hence the magnetic fluxes $\Phi^{(1)}$ and $\Phi^{(2)}$ created by these currents are much denser than the magnetic flux Φ_0. Since $\Phi^{(1)} + \Phi^{(2)} = \Phi_0$ (complex numbers), we get

$$\Phi^{(1)} \approx -\Phi^{(2)}, \quad | \Phi^{(1)} | \approx | \Phi^{(2)} |. \qquad (51.25)$$

Let us consider the equations

$$| \Phi^{(1)} | = \text{const} | I_1 | N_1, \quad | \Phi^{(2)} | = \text{const} | I_2 | N_2, \qquad (51.26)$$

which become obvious if we consider that $\Phi^{(1)}$ and $\Phi^{(2)}$ are magnetic fluxes created by each winding. In this case (51.25) becomes the equality

$$| I_1 | N_1 = | I_2 | N_2. \qquad (51.27)$$

This equation can be written in the following more convenient form:

$$\frac{| I_2 |}{| I_1 |} = \frac{N_1}{N_2}, \qquad (51.28)$$

which, as expected. is identical to (51.17).

www.EngineeringBooksPDF.com

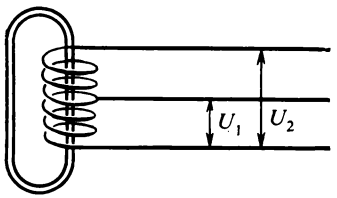


Fig. 212. Autotransformer

The first transformers were constructed by P.N. Yablochkov (1847-1894) in 1877 and by F.I. Usagin (1855-1919) in 1882.

Autotransformer. An **autotransformer** has an optimum construction, saving the winding material (Fig. 212). The physical principles of its operation and formulas are analogous to those considered above. The difference in operation of an ordinary transformer and an autotransformer is that, in the latter case, the primary and secondary are connected electrically, while the windings in the former are insulated. Hence, for example, the static electric charges in an autotransformer can go over from the primary to the secondary winding, which is impossible in an ordinary transformer. These peculiarities of transformers and autotransformers must be taken into consideration in many cases.

Transformer as a circuit element. The current in a secondary circuit is (see Fig. 209)

$$I_2 = U_2/R. \tag{51.29}$$

Considering that $I_1N_1 = I_2N_2$ and $U_1/N_1 = U_2/N_2$, we obtain from (51.29)

$$\frac{N_1}{N_2} I_1 = \frac{1}{R} \frac{N_2}{N_1} U_1. \tag{51.30}$$

Consequently, the resistance R in the secondary circuit of a transformer at the input is represented by the effective resistance

$$R_{\text{eff}} = \frac{U_1}{I_1} = \left(\frac{N_1}{N_2}\right)^2 R. \tag{51.31}$$

This means that a *transformer can be used for matching a power source with the load in order to obtain maximum power output* [see (49.34)]. For example, it can be used to match a high internal resistance of an amplifier with a low resistance of a loud speaker. Complex impedances are transformed in the same way as (51.31).

Real transformer. It can be seen from (51.31) that as regards primary, an ideal transformer can be represented as a pure resistance. The inductance of the primary is not manifested in any way on account of the mutual cancellation of the magnetic fluxes created by the currents in the primary and in the secondary. In other words, a transformer in a circuit behaves as a transducer of the effective resistance, having no inductance of its own.

The relations obtained above are valid for an ideal transformer. A **real transformer** has an inductance as well as a capacitance. The equivalent diagram of a real transformer is shown in Fig. 213. The inductances L_1 and L_2 of the primary and secondary are due to the dissipation of the magnetic flux, which means that there is no complete compensation of magnetic fluxes created by the currents in the primary and secondary. The resistances R_1 and R_2 are the ohmic resistances of the primary and secondary windings. The inductance L_0 in the primary is due to the magnetic flux corresponding to the no-load current

www.EngineeringBooksPDF.com

in the primary. The capacitances C_1 and C_2 in the windings appear due to the capacitive coupling between the turns of these windings.

It can be concluded from the equivalent diagram of a transformer that at very low frequencies the transformer ceases to operate on account of the fact

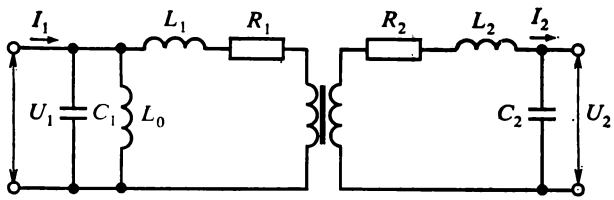


Fig. 213. Schematic diagram of a real transformer

that the inductive reactance ωL_0 becomes very low and most of the current passes through the inductance L_0. At quite high frequencies also, the transformer does not work because the current mainly passes through the capacitor C_1, by-passing the transformer windings. The limits of normal operation of the transformer are always indicated in its specifications.

What are the physical conditions under which current resonance and voltage resonance can be realized?
What is the correspondence between the parameters characterizing an *RCL* oscillatory circuit and the parameters of a mechanical vibrational system with friction?
What is the physical meaning of matching the load to the generator?
Give examples of the situations when eddy currents are useful and when they are undesirable.

Why must the core of an autotransformer be closed?
What are the main advantages and drawbacks of synchronous and asynchronous motors?
What is the role of "slip" in an asynchronous motor?
What does it depend on?
How should the transformer be connected to match the generator and the load if the load resistance is extremely small?
What i sthe difference between an ideal and a real transformer?

Sec. 52. Three-Phase Current

The basic physical phenomena occurring in a three-phase circuit are described.

Definition. The current considered so far was characterized by its amplitude and phase, and was called **single-phase current.** The combination of three identical single-phase currents, mutually displaced in phase by one third of the period, constitutes a **three-phase current.**

www.EngineeringBooksPDF.com

Generation of three-phase current. Let us consider an a.c. generator with three isolated windings arranged at an angle of 120° from one another in which cur-

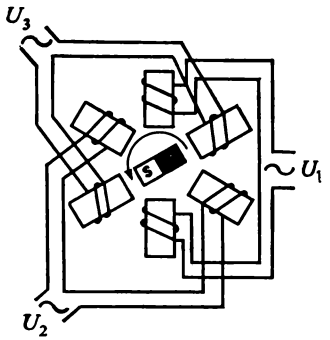


Fig. 214. Three-phase generator

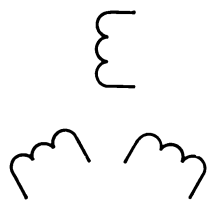


Fig. 215. Schematic diagram of the windings of a three-phase generator

rent is generated (Fig. 214). A rotating magnetic field created by a rotating permanent magnet generates identical voltages in the windings but with a phase shift:

$$U_1 = U_0 \sin \omega t,\ U_2 = U_0 \sin (\omega t + 2\pi/3),\ U_3 = U_0 \sin (\omega t - 2\pi/3). \tag{52.1}$$

The windings of the generator can be conveniently represented in the form of the diagram shown in Fig. 215.

Star connection of generator windings. If three generator windings are used independently of one another, the three-phase current generator will just be a combination of three single-phase current generators and will contain no new elements. In particular, the supply of electric power to the load will require three pairs of conducting wire.

If the windings are interconnected in a certain manner, the three-phase current acquires certain properties which are very useful for technical purposes. There are two types of winding connection: the **star connection** and the **delta connection.**

The star connection and the vector diagram for voltages in the windings are shown in Fig. 216*a* and *b*. In this case, we have a common point *O* at the same potential for all the windings. The voltage across each winding is called **phase voltage.** The conductor connected to the common-potential point is called the **neutral wire,** while the conducting wires connected to the free terminals of the windings are called **phase wires.** Thus, the *phase voltage is the voltage between the phase wire and the neutral wire.* The voltage between phase wires is called **line voltage.** It can be seen from the vector diagram that the amplitudes $U_{0\,l}$ and $U_{0\,ph}$ of the line and phase voltages are related as follows:

$$U_{0\,l} = 2U_{0\,ph} \sin 60° = U_{0\,ph} \sqrt{3}. \tag{52.2}$$

www.EngineeringBooksPDF.com

In particular, if $U_{0\,ph} = 127$ V, $U_{0\,l} = 220$ V. The current I_{ph} in the winding is called the **phase current**, while the current I_l in the line is called the line

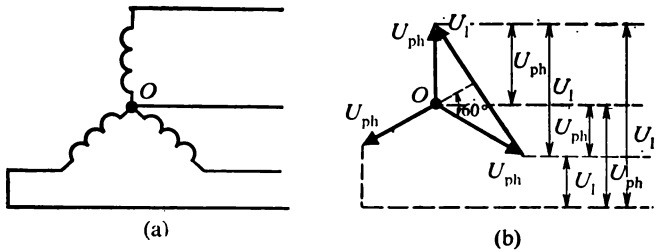


Fig. 216. Star connection of windings of a three-phase generator (*a*) and the corresponding vector diagram of voltage (*b*)

current. In a star connection, the phase and line currents are equal ($I_{ph} = I_l$). If identical loads R are connected to each winding, the total current through the neutral wire will be equal to zero, since

$$I_1 + I_2 + I_3 = \frac{1}{R}(U_1 + U_2 + U_3) = 0. \tag{52.3}$$

This is so because, as can be seen from the vector diagram,

$$\sum_i U_1 = 0.$$

The star connection of generator windings allows us to use four cables instead of six in electric power transmission, which is an important achievement.

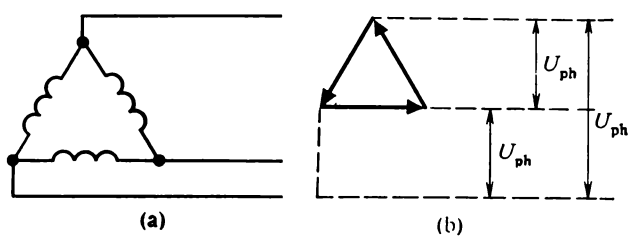


Fig. 217. Delta connection of windings of a three-phase transformer (*a*) and the corresponding vector diagram of voltage (*b*)

Delta connection of generator windings. Such a connection and the corresponding vector diagram are shown in Fig. 217*a* and *b*. In this case, $U_{0\,ph} = U_{0l}$. From the vector diagram of currents (Fig. 218), it can be seen that

$$I_{0l} = 2I_{0\,ph}\cos 30° = I_{0\,ph}\sqrt{3}. \tag{52.4}$$

$$I_{1\,ph} + I_{2\,ph} + I_{3\,ph} = 0. \tag{52.5}$$

www.EngineeringBooksPDF.com

When the generator windings are delta-connected without any load, there is no short-circuit current in the windings. This, however, is true only for the

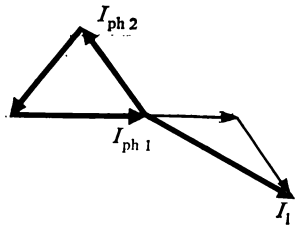


Fig. 218. Vector diagram of currents for star-connected windings

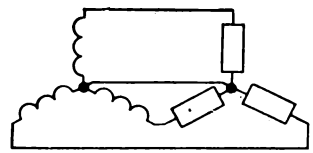


Fig. 219. Star-star connection

fundamental harmonic. Higher-harmonic currents due to non-linearity of vibrations always exist in the windings. Hence, as a rule, the windings of powerful generators are not delta-connected.

Load connection. Loads can also be star- or delta-connected and applied to a three-phase generator whose windings are also star-connected or delta-connect-

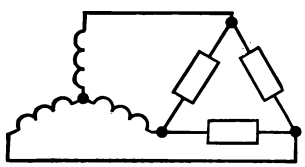


Fig. 220. Star-delta connection

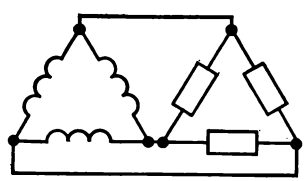


Fig. 221. Delta-delta connection

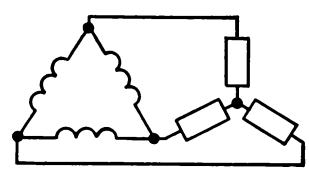


Fig. 222. Delta-star connection

ed. Thus, we have four possible combinations in which a generator and a load can be connected (Fig. 219-222).

Each of these combinations has its own peculiarities.

In a star-star connection, all loads have a phase voltage. In accordance with (52.3), the current in the neutral wire is very small for nearly identical loads. However, the neutral wire cannot be removed, since this would lead to a line voltage $U_{0l} = U_{0\,ph}\sqrt{3}$ between each pair of loads. This voltage would be distributed between loads in accordance with their resistances. But such a dependence of voltage on load is inadmissible, and hence the neutral wire must be always retained and no fuses should be connected to it.

In the case of a star-delta connection (Fig. 220), a line voltage $U_{0l} = U_{0\,ph}\sqrt{3}$ is applied to each load irrespective of the load resistance.

A phase voltage independent of the load resistance acts on all loads in a delta-delta connection (Fig. 221).

For a delta-star connection (Fig. 222), the voltage at each load is equal to $U_{0\,ph}/\sqrt{3}$.

www.EngineeringBooksPDF.com

Generation of a rotating magnetic field. If a three-phase current is applied to the windings of a generator (see Fig. 214), a rotating magnetic field corresponding to the field of a rotating magnet generating current is produced in the space between the windings. If a squirrel-cage rotor is used in place of the magnet, it will start rotating. In other words, the generator will operate as an asynchronous motor. Thus the construction of electric motors is considerably simplified if a three-phase current is used. This is also a significant advantage of three-phase currents.

Dolivo-Dobrovolskii (1862-1919) was the first to obtain a rotating magnetic field. In 1889, he also constructed the first asynchronous motor and then transmitted electric power over large distances with the help of a three-phase current. Three-phase current has ensured a wide and effective application of current in engineering.

What are the main advantages of three-phase current over single-phase current? Draw the circuit diagrams for star- and delta-connections of loads and generators. Enumerate the relations between phase and line voltages and currents.

Sec. 53. Skin Effect

The physical pattern of the emergence of skin effect is presented and the basic theory of skin effect and its consequences are described. The concept of anomalous skin effect is introduced.

Essence of the phenomenon. Direct current is distributed uniformly over the cross section of a rectilinear conductor. *Due to inductive interaction between different current elements during the passage of an alternating current, the current density is redistributed over the cross section of the conductor so that the current is concentrated predominantly in the surface layer of the conductor.* The tendency of alternating current to flow near the surface of a conductor is called the **skin effect.**

Physical pattern of the emergence of skin effect. Let us consider a cylindrical current-carrying conductor (Fig. 223). A magnetic field is created around the conductor, its lines of force being concentric circles whose centre lies on the axis of the conductor. With increasing current, the magnetic induction also increases while the shape of the lines of force remains unchanged. Hence, at each point of the conductor, the derivative $\partial \mathbf{B}/\partial t$ is directed along the tangent to the magnetic field line and thus the lines $\partial \mathbf{B}/\partial t$ are also circles coinciding with the magnetic field lines. In accordance with the law of electromagnetic induction, the varying magnetic field

$$\operatorname{curl} \mathbf{E} = -\partial \mathbf{B}/\partial t \tag{53.1}$$

www.EngineeringBooksPDF.com

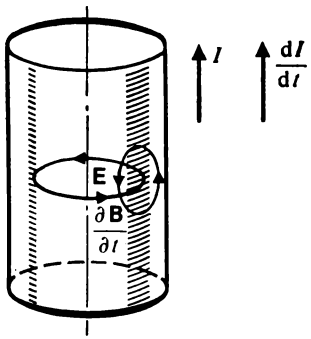


Fig. 223. Physical pattern of the emergence of skin effect

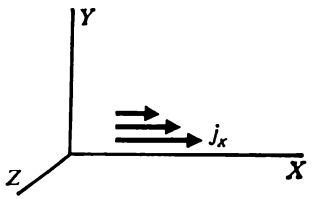


Fig. 224. Skin effect in an infinite conductor with a plane boundary

induces an electric field, whose lines of force are closed curves around the magnetic field line (Fig. 223). *Near the conductor axis, the vector of induced electric field is directed against the vector of the electric field creating the current, while in more remote regions the two vectors coincide in direction. Consequently, the current density decreases in the axial regions and increases near the surface of the conductor, thus leading to the skin effect.*

Basic theory. To begin with, we must obtain an equation for the skin effect. We start with the Maxwell equation

$$\text{curl } \mathbf{B} = \mu \mathbf{j} \tag{53.2}$$

and Eq. (53.1). Substituting into (53.2) the expression for $\mathbf{j}$ in accordance with Ohm's law,

$$\mathbf{j} = \gamma \mathbf{E}, \tag{53.3}$$

and differentiating both sides of the equation thus obtained with respect to time, we obtain

$$\text{curl } \frac{\partial \mathbf{B}}{\partial t} = \mu\gamma \frac{\partial \mathbf{E}}{\partial t} \tag{53.4}$$

or, by taking into account Eq. (53.1),

$$-\text{curl curl } \mathbf{E} = \mu\gamma \frac{\partial E}{\partial t}. \tag{53.5}$$

Since

$$\text{curl curl } \mathbf{E} = \text{grad div } \mathbf{E} - \nabla^2 \mathbf{E} \tag{53.6}$$

and

$$\text{div } \mathbf{E} = 0,$$

we finally get

$$\nabla^2 \mathbf{E} = \gamma\mu \frac{\partial \mathbf{E}}{\partial t}. \tag{53.7}$$

In order to simplify the solution of this equation, we assume that the current flows through an infinite homogeneous conductor occupying the half-space

www.EngineeringBooksPDF.com

$y > 0$ along the X-axis (Fig. 224). The plane $Y = 0$ is the surface of the conductor. Thus,

$$j_x = j_x(y, t),\ j_y = j_z = 0, \tag{53.8}$$

$$E_x = E_x(y, t),\ E_y = E_z = 0. \tag{53.9}$$

This gives [see (53.7)]

$$\frac{\partial^2 E_x}{\partial y^2} = \gamma\mu \frac{\partial E_x}{\partial t}. \tag{53.10}$$

Since all quantities in (53.10) depend harmonically on t, we can put

$$E_x(y, t) = E_0(y)\, e^{i\omega t}. \tag{53.11}$$

Substituting (53.11) into (53.10) and dividing both sides of the equation thus obtained by $\exp(i\omega t)$, we obtain the following equation for $E_0(y)$:

$$\frac{d^2 E_0}{dy^2} = i\gamma\mu\omega E_0. \tag{53.12}$$

The general solution of this equation has the form

$$E_0 = A_1 e^{-ky} + A_2 e^{ky}. \tag{53.13}$$

Considering that

$$k = \sqrt{i\gamma\mu\omega} = \alpha(1+i), \quad \alpha = \sqrt{\gamma\mu\omega/2}, \tag{53.14}$$

we obtain

$$E_0(y) = A_1 e^{-\alpha y} e^{-i\alpha y} + A_2 e^{\alpha y} e^{i\alpha y}. \tag{53.15}$$

As we move away from the surface of the conductor ($y \to \infty$), the second term in (53.15) increases indefinitely, which is a physically inadmissible situation. Consequently, $A_2 = 0$ in (53.15) and only the first term remains as a physically acceptable solution. In this case, the solution of the problem has the following form (if we take into account Eq. (53.11)):

$$E_x(x, t) = A_1 e^{-\alpha y} e^{i(\omega t - \alpha y)}. \tag{53.16}$$

Taking the real part of this expression and going over to current density with the help of the relation $\mathbf{j} = \gamma\mathbf{E}$, we obtain

$$j_x(y, t) = \gamma A_1 e^{-\alpha y} \cos(\omega t - \alpha y). \tag{53.17}$$

Considering that $j_x(0, 0) = j_0$ is the amplitude of current density at the surface of the conductor, we arrive at the following distribution of the volume current density in a conductor:

$$j_x(y, t) = j_0 e^{-\alpha y} \cos(\omega t - \alpha y). \tag{53.18}$$

Skin depth. At the surface of a conductor, the volume current density has its maximum value. As we move from the surface, this quantity decreases and at a distance $\Delta = 1/\alpha$ it is equal to 1/e of its value at the surface. Hence, the entire current is practically concentrated in the layer Δ called the **skin depth.** On the basis of (53.14), it is equal to

$$\Delta = [2/(\gamma\mu\omega)]^{1/2}. \tag{53.19}$$

www.EngineeringBooksPDF.com

Obviously, the skin depth may be very small for very high frequencies ω. For example, the skin depth $\Delta = 4$ mm for a good conductor like copper ($\gamma = 10^7\ \Omega^{-1}\cdot m^{-1}$) and at $\omega = 10^4\ s^{-1}$. If the frequency is increased by a factor of 100 and reaches $\omega = 10^6\ s^{-1}$, the skin depth decreases by a factor of 10 ($\Delta \simeq 0.4$ mm). This means that at quite high frequencies, the entire current in not very thin conductors flows only in a very small part of the conductor cross section near its surface. Hence the situation will be the same if we remove the material from the cylindrical region inside the conductor and leave only a cylindrical shell of the thickness equal to the skin depth. If the conductor is quite thick and the frequency is not very high, the current flows through the entire cross section becoming weaker only slightly near the axis. For example, the skin effect is manifested very weakly in normal conductors in the domestic power supply at a frequency of 50 Hz.

Frequency dependence of ohmic resistance of a conductor. Since the effective cross-sectional area through which a current flows decreases with increasing frequency, the *resistance of a conductor increases with frequency.*

Frequency dependence of inductance of a conductor. The energy of a magnetic field through which a current flows is equal to

$$W_m = {}^1/_2 LI^2. \tag{53.20}$$

If the current flows through a hollow cylinder, the field outside the cylinder will be the same as for the same current passing through a solid cylinder. Inside the hollow cylinder, there is no field. Hence, the energy of the field created by the current flowing through a hollow cylinder is less than that of the field generated by the same current passing through a solid cylinder. This means that the magnetic field energy W_m decreases on account of the skin effect. Hence, on the basis of (53.20) it can be stated that the *inductance of a conductor decreases with increasing frequency.*

Hardening of metals by high-frequency currents. Owing to the skin effect, at high frequencies Joule's heat is released predominantly in the surface layer. This allows the hardening of a thin surface layer of a conductor without any significant changes in the temperature of its interior. This phenomenon is used in industry for hardening metals by the method which is quite important from the technological point of view.

Anomalous skin effect. The mechanism described above for the emergence of skin effect assumes that a moving electron continuously loses its energy in order to overcome the ohmic resistance of the conductor, thus leading to liberation of Joule's heat. Obviously, such an idealization is possible only if the motion of electrons takes place in regions whose linear dimensions are much larger than the mean free path of an electron between collisions with the atoms of the substance. Consequently, the *above theory is valid only under the condition that the skin depth is much larger than the mean free path of freely moving electrons.* Such a relation between the mean free path and the skin depth is observed over quite a wide range. For example, even at a frequency of 10 GHz and a temperature of 300 K, the skin depth in copper is about 1 μm, while the mean free path is of the order of 0.01 μm. However, *the situation radically changes at very low*

www.EngineeringBooksPDF.com

temperatures, since the conductivity rapidly increases and so does the mean free path, while the skin depth decreases. For example, at liquid helium temperature (4.2 K) the conductivity of pure copper increases about 10^4 times. This leads to an increase in the mean free path of electrons l y a factor of 10^4 and a decrease in the skin depth by a factor of $\sqrt{10^4} = 10^2$. Thus, the mean free path and the skin depth become equal to 100 and 0.01 μm respectively. Under these conditions, the mechanism leading to the emergence of the skin effect is no longer operative. *The effective thickness of the layer in which the current is concentrated changes. This phenomenon is called the anomalous skin effect.*

Under the conditions of anomalous skin effect, only electrons, whose velocity is parallel to the surface of the conductor can traverse the entire free path within the normal skin depth. All other electrons manage to leave the "normal" skin depth during their free motion and thus considerably change the direction of motion. This leads to a decrease in the conductivity of the material and a change in the effective "anomalous" skin depth Δ'. In order to roughly estimate the value of this quantity, it can be assumed that a fraction of conduction electrons has the order of Δ'/l of the number of electrons which would ensure the conduction in the framework of "normal" skin effect (l is the mean free path of electrons). A decrease in this fraction would lead to a decrease in the conductivity which can approximately be taken into account by the substitution $\gamma \to \beta\gamma\ (\Delta'/l)$, where β is the numerical coefficient of the order of unity. Making this substitution in formula (53.19), we get

$$\Delta' = [2l/(\beta\gamma\mu\omega)]^{1/3}. \tag{53.21}$$

Owing to inductive interaction between different current elements in the case of an alternating current, a redistribution of the current density takes place over the cross section of a conductor. As a result, the current is predominantly concentrated in the surface layer of the conductor.

What is the physical reason behind the dependence of the resistance and inductance of a conductor on the a.c. frequency?
What are the conditions under which the skin effect appears?

Sec. 54. Four-Terminal Networks

The terminology and basic concepts of the theory of four-terminal networks are described.

Definition. *An electric circuit with two input and two output terminals through which electric energy is transmitted is called a four-terminal network.* Such a network is schematically shown in Fig. 225. Examples of four-terminal networks are mode transducers, frequency filters, transformers, etc. Our aim is to find the relation between the voltages and currents at the input and output termi-

www.EngineeringBooksPDF.com

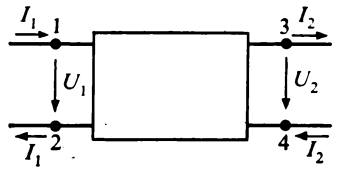


Fig. 225. Four-terminal network

nals of a four-terminal network. If such a network does not contain any power sources, it is called a passive network. In the presence of power sources, the network is called active. It is assumed that the current at the output terminal *2* is the same as the current at the input terminal *1*. Similarly, the current at the output terminal *3* is equal to the current at the input terminal *4*.

Equations. Suppose that a four-terminal network contains n independent loops. We can then compose n equations for the current loops of the type (48.27):

$$\begin{aligned} &\sum_{i=1}^{n} Z_{1i}I_i = U_1, \\ &\sum_{i=1}^{n} Z_{2i}I_i = -U_2, \\ &\sum_{i=1}^{n} Z_{ki}I_i = 0 \quad (k=3, 4, \ldots, n). \end{aligned} \tag{54.1}$$

The minus sign of U_2 in the second equation appears due to the fact that here a certain direction of circumvention is chosen as positive, and relative to this direction U_1 and U_2 are passed in opposite directions (see Fig. 225). The solution of this system of equations is

$$\begin{aligned} I_1 &= \frac{\Delta_{11}}{\Delta}U_1 - \frac{\Delta_{21}}{\Delta}U_2, \\ I_2 &= \frac{\Delta_{12}}{\Delta}U_1 - \frac{\Delta_{22}}{\Delta}U_2, \end{aligned} \tag{54.2}$$

where Δ and Δ_{ij} are the determinant and the corresponding complements of the system of equations (54.1). Consequently, there exist linear dependences of the type (54.2) between currents and voltages of a passive four-terminal network. These dependences can be conveniently written in the form

$$I_1 = B_{11}U_1 + B_{12}U_2, \quad I_2 = B_{21}U_1 + B_{22}U_2. \tag{54.3}$$

The coefficients B_{ij} have the dimensions of conductivity. *Hence Eqs.* (54.3) *are called the equations of a four-terminal network with coefficients in the form of conductivities.*

It is not difficult to solve Eqs. (54.3) for voltages

$$U_1 = A_{11}I_1 + A_{12}I_2, \quad U_2 = A_{21}I_1 + A_{22}I_2, \tag{54.4}$$

where coefficients A_{ij} have the dimensions of resistances (impedances). *These equations are called the equations of a four-terminal network with coefficients in the form of resistances.*

www.EngineeringBooksPDF.com

Reciprocity theorem. Since the coefficients Z_{ij} in Eqs. (54.1) are symmetric for a passive four-terminal network [see [48.30)], i.e.

$$Z_{ij} = Z_{ji}, \tag{54.5}$$

we can show that the coefficients A_{ij} in (54.4) are also symmetric in this case:

$$A_{12} = A_{21}. \tag{54.6}$$

Hence

$$(U_2/I_1)_{I_2=0} = (U_1/I_2)_{I_1=0}, \tag{54.7}$$

i.e. *for a given input current, the output voltage at the open pair of terminals remains unchanged if the input and output terminals of such a network are interchanged (reciprocity theorem for a passive four-terminal network).*

Impedance of a four-terminal network. The impedance A_{21} is called the mutual impedance of a four-terminal network, since it follows from the second of Eqs. (54.4) for an open output circuit ($I_2 = 0$) that

$$A_{21} = U_2/I_1. \tag{54.8a}$$

Under the same condition, we get for the first of Eqs. (54.4)

$$A_{11} = U_1/I_1. \tag{54.8b}$$

This means that A_{11} is the input impedance of a four-terminal network when the output circuit is open. In accordance with the reciprocity theorem, the terms A_{12} and A_{22} have the same meaning.

Simple four-terminal networks. With the help of Eqs. (54.3) and (54.4), we can connect the input voltage and current with the output voltage and current of a four-terminal network through the relations

$$U_1 = D_{11}U_2 + D_{12}I_2, \quad I_1 = D_{21}U_2 + D_{22}I_2, \tag{54.9}$$

where D_{ij} can be easily expressed in terms of B_{ij} and A_{ij}, viz. the quantities which appear in Eqs. (54.3) and (54.4); the coefficient D_{12} has the dimensions of resistance; D_{21} has the dimensions of conductivity, while the coefficients D_{11} and D_{22} are dimensionless.

A four-terminal network is called longitudinally symmetric if interchanging the input and output terminals leaves the currents and voltages in the circuits connected to these terminals unchanged. Since such an interchange is possible, we obtain the following expression for symmetric four-terminal networks:

$$D_{11} = D_{22}. \tag{54.10}$$

Simplest forms of symmetric pi-section and T-section four-terminal networks are shown in Figs. 226 and 227, while nonsymmetric networks are shown in Figs. 228 and 229. The easiest way to find the coefficients D_{ij} for a four-terminal network is by using the mesh-current method. For this purpose, a system of equations is first set up, and then the forces corresponding to the currents in the internal loops are excluded from these equations. The two remaining equations

www.EngineeringBooksPDF.com

containing U_1, U_2, I_1, and I_2 are transformed into equations of the type (54.9) and a comparison of these equations with (54.9) gives the values of D_{ij}.

For a longitudinally symmetric pi-section four-terminal network (Fig. 226), we obtain

$$D_{11} = 1 + ZY/2, \quad D_{12} = Z, \quad D_{11} = Y(1 + ZY/4). \tag{54.11}$$

For a longitudinally symmetric T-section four-terminal network (Fig. 227), we have

$$D_{11} = 1 + ZY/2, \quad D_{12} = Z(1 + ZY/4), \quad D_{21} = Y. \tag{54.12}$$

As a result of direct verification, we find that

$$D_{11}^2 - D_{12}D_{21} = 1, \tag{54.13}$$

i.e. the determinant of the coefficients of transformation (54.9) is equal to unity in the case of longitudinally symmetric pi- and T-section four-terminal networks.

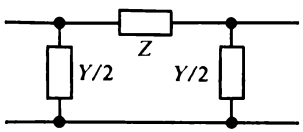


Fig. 226. Longitudinally symmetric pi-section four-terminal network

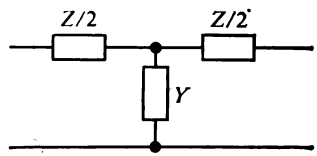


Fig. 227. Longitudinally symmetric T-section four-terminal network

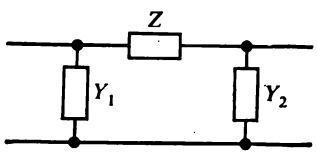


Fig. 228. Nonsymmetric pi-section four-terminal network

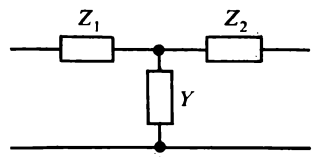


Fig. 229. Nonsymmetric T-section four-terminal network

The expressions for the coefficients of nonsymmetric four-terminal networks are somewhat complicated and will not be given here.

Input and output impedances. For a four-terminal network, these quantities are defined as the ratio of the corresponding voltages and currents:

$$Z_{in} = U_1/I_1, \quad Z_{out} = U_2/I_2. \tag{54.14}$$

Taking into account Eqs. (54.10)-(54.13), we obtain from Eq. (54.9)

$$Z_{in} = \frac{Z_{out} + D_{12}/D_{11}}{1 + Z_{out}D_{21}/D_{11}}. \tag{54.15}$$

Thus, *a four-terminal network transforms the output impedance into input impedance.* When the output is short-circuited ($Z_{out} = 0$), the input impedance

www.EngineeringBooksPDF.com

of a four-terminal network is

$$Z_{0\,in} = D_{12}/D_{11}, \tag{54.16}$$

while for an open output circuit ($Z_{out} = \infty$), the input impedance is defined by the expression

$$Z_{\infty in} = D_{11}/D_{21}. \tag{54.17}$$

Gain factor. The transformation of currents and voltages is characterized by the ratio of their values at the output to the values at the input. In the same way as for (54.15), we obtain

$$U_2/U_1 = Z_{out}/(Z_{out}\,D_{11} + D_{12}), \tag{54.18}$$

$$I_2/I_1 = 1/(D_{11} + Z_{out}D_{21}). \tag{54.19}$$

If the four-terminal network does not transform the impedance, i.e. *if the input and output impedances are identical, the output impedance is said to be matched with the system.* Substituting the value

$$Z_{ch} = Z_{in} = Z_{out} \tag{54.20}$$

of the impedance into (54.15), we obtain

$$Z_{ch} = \sqrt{D_{12}/D_{21}}. \tag{54.21}$$

This quantity is called the **characteristic (wave) impedance** of a four-terminal network. Consequently, a four-terminal network is matched with the transmission line if its input and output impedances are equal to the characteristic impedance. In this case, Eqs. (54.18) and (54.19) assume the form

$$U_2/U_1 = 1/(D_{11} + \sqrt{D_{12}D_{21}}), \tag{54.22}$$

$$I_2/I_1 = 1/(D_{11} + \sqrt{D_{12}D_{21}}). \tag{54.23}$$

Using the relation

$$\cosh g = D_{11}, \tag{54.24}$$

we can determine the gain factor g. In this case, we obtain on the basis of (54.13)

$$\sinh g = \sqrt{\cosh^2 g - 1} = \sqrt{D_{12}D_{21}}. \tag{54.25}$$

Using (54.24) and (54.25), we can transform formulas (54.22) and (54.23) as follows:

$$U_2 = U_1 e^{-g}, \tag{54.26}$$

$$I_2 = U_1 e^{-g}. \tag{54.27}$$

It should be noted that expressions (54.26) and (54.27) are valid only under the conditions of perfect matching. In the case of mismatching, formulas (54.18) and (54.19) should be used.

www.EngineeringBooksPDF.com

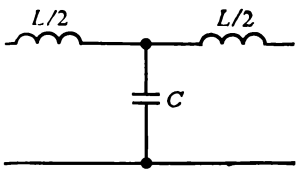


Fig. 230. Low-pass filter

With the help of the gain factor and characteristic impedance, formulas (54.18) and (54.19) can be represented in the form

$$U_2/U_1 = Z_{out}/(Z_{out}\cosh g + Z_{ch}\sinh g) \quad (54.28)$$

$$I_2/I_1 = Z_{ch}/(Z_{ch}\cosh g + Z_{out})\sinh g). \quad (54.29)$$

Like all other quantities appearing in formulas (54.26)-(54.29), the gain factor is a complex quantity:

$$g = \alpha + i\beta. \quad (54.30)$$

It can be seen from (54.26) and (54.27) that under matching conditions, the real part of the gain factor determines the change in the amplitudes of voltage and current at the output of a four-terminal network as compared to their values at the input, while the imaginary part determines the phase shift. The real part of the gain factor is just the logarithm of the ratio of amplitudes:

$$\alpha = \ln(U_1/U_2). \quad (54.31)$$

Since g depends on frequency, the spectral composition, and hence the shape of a multifrequency signal, change as it passes through a four-terminal network. The nature of variation of frequency and phase spectrum of a signal can be determined with the help of the formulas obtained in this section.

Sec. 55. Filters

The principle of operation and properties of filters are described.

Definition. A **filter** is a device that changes the amplitude of oscillations depending on their frequency. If a filter is made in the form of a four-terminal network, the transmission coefficient should noticeably vary with frequency.

Low-pass filter. Let us consider a bridged T-section four-terminal network shown in Fig. 230. A comparison with Fig. 227 shows that in the formulas obtained above we must put

$$Z = i\omega L, \quad Y = i\omega C. \quad (55.1)$$

The characteristic resistance is given, according to (54.24) and (54.11), by

$$Z_{ch} = \sqrt{\frac{Z}{Y}}\sqrt{1+\frac{ZY}{4}} = \sqrt{\frac{L}{C}}\sqrt{1-\frac{\omega^2 LC}{4}}. \quad (55.2)$$

Taking into account (54.11), we obtain the following expression for the transmission coefficient g [see (54.24)]:

$$\cosh g = 1 - \omega^2 LC/2. \quad (55.3)$$

www.EngineeringBooksPDF.com

Taking into account expression (54.30) for g, we can write Eq. (55.3) in the form

$$\cosh(\alpha + i\beta) = \cosh\alpha\cos\beta + i\sinh\alpha\sin\beta = 1 - \omega^2 LC/2, \tag{55.4}$$

whence

$$\cosh\alpha\cos\beta = 1 - \omega^2 LC/2, \tag{55.5}$$

$$\sinh\alpha\sin\beta = 0. \tag{55.6}$$

Equation (55.6) has the following solutions:

$$\beta = \pi n \quad (n = 0, 1, 2, \ldots), \tag{55.7}$$

for which $\cos\beta = \pm 1$. However, the hyperbolic cosine is always greater than or equal to unity, i.e. $\cosh\alpha \geqslant 1$. Hence it follows from (55.5) that $\cos\beta = -1$ and we can put $\beta = \pi$. Under these conditions, Eq. (55.5) assumes the form

$$1 + \cosh\alpha = \omega^2 LC/2. \tag{55.8}$$

Since $\cosh\alpha \gg 1$, Eq. (55.8) has a solution only for sufficiently high frequencies:

$$\omega \geqslant \omega_c, \tag{55.9}$$

where

$$\omega_c = 2/\sqrt{LC} \tag{55.10}$$

is the cut-off frequency. Taking into account (55.9), we conclude from (55.2) that the characteristic resistance is purely imaginary:

$$Z_{out} = i\sqrt{\frac{L}{C}}\sqrt{\frac{\omega^2 LC}{4} - 1} = i\sqrt{\frac{L}{C}}\sqrt{\frac{\omega^2}{\omega_c^2} - 1}\,. \tag{55.11}$$

The real part of the transmission coefficient can be determined from Eq. (55.8). It can be seen that it increases very rapidly with frequency. Hence on the basis of (54.26) and (54.27) we may conclude that the amplitudes of oscillations at the four-terminal network output for $\omega \geqslant \omega_c$ rapidly decrease with increasing frequency.

Another solution of Eq. (55.6) has the form

$$\sinh\alpha = 0, \quad \alpha = 0. \tag{55.12}$$

Then Eq. (55.5) assumes the form

$$\cos\beta = 1 - \omega^2 LC/2. \tag{55.13}$$

This equation has a solution only for $\cos\beta \geqslant -1$, i.e. at frequencies

$$\omega \leqslant \omega_c = 2/\sqrt{LC}, \tag{55.14}$$

for which the first solution did not fit. The characteristic resistance in this case is real:

$$Z_{ch} = \sqrt{\frac{L}{C}}\sqrt{1 - \frac{\omega^2}{\omega_c^2}}. \tag{55.15}$$

www.EngineeringBooksPDF.com

Since $\alpha = 0$ here, the frequencies $\omega \leqslant \omega_c$ pass without attenuation. However, there is a frequency-dependent phase shift determined by Eq. (55.13).

Figure 231 shows the dependence of the amplitude of oscillations at the output on the amplitude at the input. The four-terminal network considered above is a filter capable of passing low frequencies, less than a certain cut-off frequency

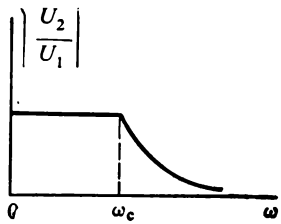


Fig. 231. Frequency characteristic of a low-pass filter

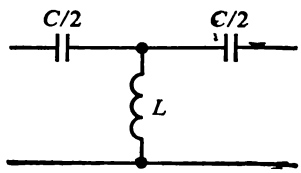


Fig. 232. High-pass filter

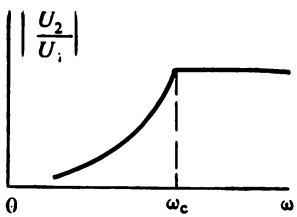


Fig. 233. Frequency characteristic of a high-pass filter

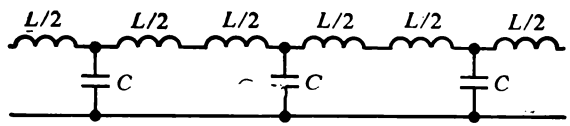


Fig. 234. Iterative filter comprising T-section networks

ω_c. The oscillations with frequencies higher than the cut-off frequency attenuate very rapidly. The filter operates as a shutter for frequencies which considerably exceed the cut-off frequency. The frequency range $\omega \leqslant \omega_c$ is called the pass band.

High-pass filter. A four-terminal network shown in Fig. 232 is calculated as in the previous case and operates as a high-pass filter with the frequency characteristic shown in Fig. 233.

Iterative filter. If we connect the output terminals of a four-terminal network shown in Fig. 230 to the input terminals of a similar four-terminal network and continue the process, we shall obtain a filter shown in Fig. 234. It can be analyzed with the help of the same methods. However, the basic properties of such a filter can be outlined without a detailed calculation, since consecutive units constituting the filter have identical characteristic resistances and operate in matching mode at each frequency. The cut-off frequency for all units is the same. Consequently, this filter will have the same pass band $\omega \leqslant \omega_c$, and the attenuation of frequencies $\omega \geqslant \omega_c$ will be more pronounced. The frequency

www.EngineeringBooksPDF.com

characteristic of this filter has the form similar to that shown in Fig. 231 but with a steeper decrease in amplitudes at $\omega > \omega_c$ (Fig. 235).

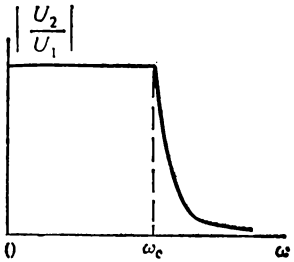


Fig. 235. Frequency characteristic of an iterative filter comprising T-section networks

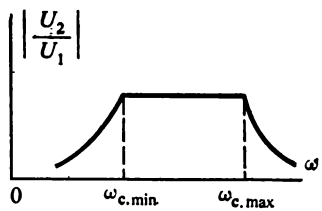


Fig. 236. Frequency characteristic of a band filter

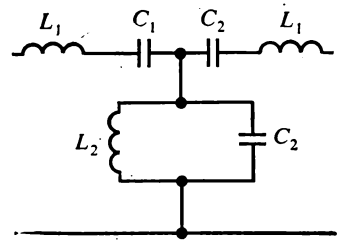


Fig. 237. Band filter

Band filter. *This is a filter that passes only a certain frequency band betweena certain minimum and maximum frequencies*:

$$\omega_{c.\,min} \leqslant \omega \leqslant \omega_{c.\,max}. \tag{55.16}$$

The frequency characteristic of this filter is shown in Fig. 236.

Such a filter can, in principle, be realized in the form of a sequence of low-pass and high-pass filters. The high-pass filter should cut off all frequencies below $\omega_{c.max}$ and pass high frequencies, while the low-pass filter should pass all frequencies below $\omega_{c.max}$, and cut off all frequencies exceeding $\omega_{c.max}$. In practice, however, more complicated networks are used (see, for example, Fig. 237). Such a filter is also a four-terminal network and can be analyzed by similar methods.

What are the physical processes underlying the low-pass and high-pass filters? Describe the construction of a band filter.

Sec. 56. Betatron

The operating principle of betatron is considered and basic concepts of the theory of stability of electron motion in a betatron are analyzed. The energy limit attainable in betatrons is discussed.

Function. Betatron is a device in which an eddy induced electric field acts on free electrons in vacuum. It is intended for accelerating electrons to a high energy (of the order of several hundreds megaelectronvolts). An acceleration to higher energies is hampered by the energy losses due to bremsstrahlung appear-

www.EngineeringBooksPDF.com

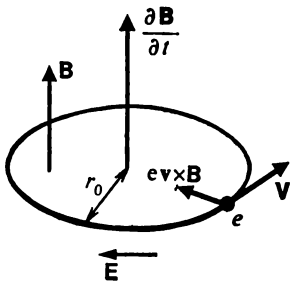


Fig. 238. To the derivation of the betatron condition

ing as a result of the accelerated motion of electrons in circular orbits. The mechanism of acceleration used in a betatron is unable to compensate these losses, and the acceleration cycle is terminated.

Operating principle. The basic idea is to specify the conditions under which an electron moving in an increasing magnetic field would be accelerated by an eddy electric field and simultaneously held by the magnetic field in a circular orbit of constant radius.

It turns out that this condition can be met. It is called the **betatron condition.**

The betatron condition. Let us write the equation of motion of an electron in a circle of constant radius in an increasing magnetic field, assuming that such a motion is possible. The solution of this equation will give us the conditions under which this motion can be realized.

We denote the radius of the orbit by r_0 and the electron momentum always directed along a tangent to the circular orbit by p (Fig. 238). Faraday's law of electromagnetic induction used for determining the electric field strength in the orbit gives the following equation:

$$2\pi r_0 E = -\,\mathrm{d}\Phi/\mathrm{d}t. \tag{56.1}$$

On the other hand, the equation of motion has the form

$$\mathrm{d}p/\mathrm{d}t = eE. \tag{56.2}$$

It follows from (56.1) and (56.2) that

$$\mathrm{d}p/\mathrm{d}t = -\frac{e}{2\pi r_0}\frac{\mathrm{d}\Phi}{\mathrm{d}t}. \tag{56.3}$$

Since $r_0 = \text{const}$, we can integrate both sides of this equation over t between 0 and t:

$$p_t - p_0 = -\,[e/(2\pi r_0)]\,(\Phi_t - \Phi_0), \tag{56.4}$$

where the subscripts t and 0 denote the values of the corresponding quantities at the moment t and at the initial moment $t = 0$. Newton's equation for the centripetal acceleration has the form

$$mv^2/r_0 = -\,evB, \tag{56.5}$$

where m is the relativistic mass. It follows from this equation that $p = mv = -\,eBr_0$. Then [see (56.4)]

$$B_t - B_0 = \frac{1}{2}\left(\frac{\Phi_t}{\pi r_0^2} - \frac{\Phi_0}{\pi r_0^2}\right). \tag{56.6}$$

www.EngineeringBooksPDF.com

Since the magnetic induction vector **B** is perpendicular to the plane of the orbit, and the magnetic flux is given by

$$\Phi = \int_S B \cdot dS \tag{56.7}$$

($S = \pi r_0^2$ is the area bounded by the orbit), then

$$\Phi/(\pi r_0^2) = \langle B \rangle \tag{56.8}$$

is the average magnetic induction of the field over the area S bounded by the orbit. Assuming that at the initial instant the magnetic field is absent ($B_0 = 0$, $\Phi_0 = 0$), and combining (56.6) and (56.8), we obtain

$$B_t = {}^1/_2 \langle B_t \rangle. \tag{56.9}$$

This is the betatron condition: *the magnetic induction on the electron orbit is equal to half the average induction of the field piercing the orbit.* Consequently, the magnetic induction should decrease from the centre to the orbit in accordance

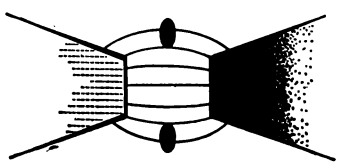


Fig. 239. Schematic diagram of a betatron

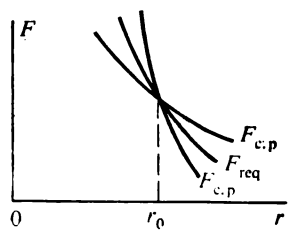


Fig. 240. To the derivation of radial stability condition for electrons in a betatron

with a certain law such that condition (56.9) is observed. For this purpose, it is necessary to properly select the shape of the poles of electromagnets creating the magnetic field (Fig. 239). Since for a given shape of magnetic poles the shape of the lines of force does not depend on the current and magnetic induction, condition (56.9) is found to be fulfilled for any current in the electromagnet. This means that *there is no need to take care of the law of variation of current.* The only question we should be interested in is the stability of motion. If some factors draw an electron out of the motion strictly along the circle of radius r_0, will the forces striving to hold it in the acceleration mode near the circle appear, or will it no longer be accelerated and be lost?

There are two possible directions of deflecting an electron from its orbit: either along the radius or along the vertical from the plane of its motion.

Radial stability. The magnetic induction of the field within the orbit is usually represented in the form

$$B = \text{const}/r^n. \tag{56.10}$$

www.EngineeringBooksPDF.com

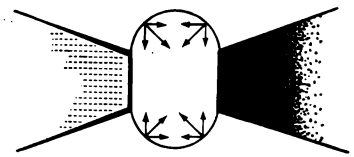


Fig. 241. Schematic diagram for ensuring the vertical stability of electron motion in a betatron

The rate of variation of magnetic induction is characterized by the quantity n. The centripetal force $F_{c.p}^{req}$ required to ensure the motion of the electron in a circle of radius r and the actual centripetal force $F_{c.p}$ at the same distance r from the centre are given by

$$F_{c.p}^{req} = mv^2/r = A_1 r, \quad F_{c.p} = evB = A_2/r^n, \quad (56.11)$$

where A_1 and A_2 are constants (v = const). The plots of these quantities for $n > 1$ and $0 < n < 1$ are shown in Fig. 240. When $r = r_0$, equality (56.5) is fulfilled, and the motion in a circle of radius r_0 takes place. If for some reason or other the electron is displaced by $r > r_0$, then for $n > 1$ the centripetal force $F_{c.p} < F_{c.p}^{req}$. This means that factors which tend to move the electron away from the orbit of radius r_0 dominate. Therefore, for $n > 1$, the motion turns out to be unstable. When $n < 1$, the centripetal force $F_{c.p} > F_{c.p}^{req}$, and the factors tending to return the electron to the orbit of radius r_0 dominate. As a result, radial stability is attained. An analysis of the case when $r < r_0$ leads to the same conclusion. Consequently, the **radial stability condition** has the form

$$0 < n < 1. \quad (56.12)$$

Vertical stability. This stability is always ensured when the magnetic induction of the field decreases towards the periphery ($n > 0$) since in this case the lines of force are convex outwards (Fig. 241). When an electron deviates from the midplane, a component of the Lorentz force tends to return it to this plane (Fig. 241). Thus, when condition (56.12) is satisfied, vertical stability of motion is also ensured, i.e. *inequality* (56.12) *is the general condition of stability of electron motion in a betatron.*

Betatron oscillations. Upon small deviations of electrons from the equilibrium orbit ($r = r_0$), they perform small harmonic oscillations about this orbit both in the radial and in vertical directions. Such oscillations are called betatron oscillations. Their amplitude determines the cross section of the vacuum toroidal chamber in which electrons move. Usually, the linear dimensions of the cross section of the chamber constitute about 5% of the radius of the orbit.

Energy limit attainable in betatron. As was mentioned above, this limit is determined by the electron energy losses for bremsstrahlung (see Chap. 10). The maximum energy practically attainable in betatrons does not exceed 300 MeV.

Problems

8.1. Calculate the inductance of a section of a two-wire line of length l, neglecting the internal inductance of the wires. The radii of the wires are equal to r_0 and the distance between the wires is d.

8.2. A current of density j passes through a straight infinite circular cylindrical conductor having a cylindrical cavity of a circular cross section. The axes of the cylinder and of

www.EngineeringBooksPDF.com

the cavity are collinear (Fig. 98). Find the magnetic induction of the field inside the cavity ($\mu = \mu_0$).

Hint: See Problem 2.9.

8.3. A very long solenoid with the density of winding equal to n turns per metre length has the cross-sectional area S. Current I passes through the winding. Two very long iron rods having permeability μ are pulled into the solenoid at both its ends so that the rods tightly fit the winding and are separated by a very small gap within the solenoid. Find the force with which the rods attract each other.

8.4. A U-shaped electromagnetic whose winding consists of n turns is characterized by the following parameters: cross-sectional area S, length l, permeability of the magnet material μ, and the distance between the poles d. A current I passes through the winding. A strip of the same material is placed in contact with the magnet poles. Find the force with which the strip is attracted by the magnet.

8.5. A horizontal metallic rod rotates at a frequency ν about a vertical axis passing at a distance equal to $1/k$ of its length from one of its ends. The length of the rod is l. Find the potential difference between the ends of the rod if it rotates in a uniform vertical magnetic field with magnetic induction B. Assume that $k = 3$, $l = 1.2$ m, $\nu = 6$ s^{-1} and $B = 10^{-2}$ T.

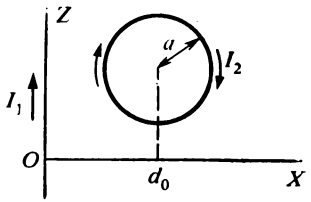


Fig. 242. Mutual arrangement of interacting linear and circular currents

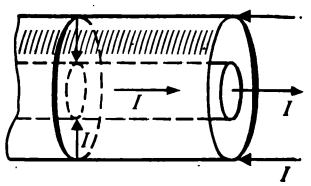


Fig. 243. Coaxial cable with a movable diaphragm

8.6. A sinusoidal magnetic field with a magnetic induction amplitude $B_0 = 0.5$ T is formed between the circular poles of a large electromagnetic fed by an alternating current of frequency $\nu = 1$ kHz. Assuming that the magnetic field is uniform, find the maximum electric field strength in the gap between the poles at a distance $r = 0.1$ m from its centre.

8.7. A short-circuited solenoid of radius b with n turns rotates at an angular velocity ω about the diameter of one of the turns in a uniform magnetic field of magnetic induction **B**. The axis of rotation is perpendicular to the magnetic induction vector. The resistance and inductance of the solenoid are equal to R and L respectively. Find the current in the solenoid as a function of time.

8.8. A superconducting ring which can move only in the vertical direction lies on a table above a conducting loop. A current I is passed through the loop. As a result, the superconducting ring is lifted. The mutual inductance of the loop and the ring lifted to a height x is $L_{12}(x)$. The inductance of the superconducting ring is L_{11}, the mass of the ring is m, the acceleration due to gravity is g. Find the height h to which the ring will be lifted.

8.9. A current $I_0 \sin \omega t$ is passed through a coil A_1 and a corresponding current is induced accordingly in coil A_2. The self-inductances and mutual inductance of the coils are L_1, L_2 and L_{12}. The resistance of coil A_2 is R_2. Suppose that ξ_i is a certain generalized coordinate which characterizes the position of coil A_2. Find the average generalized force F_i associated with the coordinate ξ_i.

8.10. A very long straight conductor and a circular conductor of radius a lie in a plane (Fig. 242). The distance between the centre of the annular conductor and the straight one is equal to d. Find their mutual inductance.

8.11. Currents I_1 and I_2 pass through the straight and circular conductors described in Problem 8.10. Find the force acting on the circular conductor.

www.EngineeringBooksPDF.com

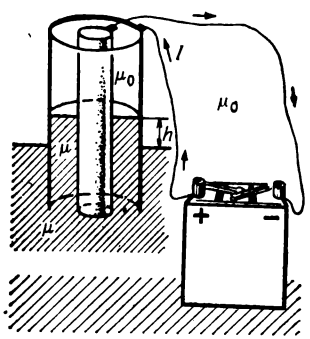


Fig. 244. Pulling a magnetic into the space between coaxial cables carrying current

8.12. Find the mutual inductance of a toroid winding (see Fig. 195) and a straight infinite conductor coinciding with the symmetry axis of the toroid.

8.13. Find the inductance of a toroid winding with a circular cross section of radius r, having n turns. The major radius of the toroid is R.

8.14. A coaxial cable whose core and sheath have infinite conductivity and radii r_1 and r_2 is short-circuited by a movable diaphragm (Fig. 243). Find the force acting on the diaphragm when current I flows in the cable.

8.15. A hollow cylinder of radius r_2 and a cylindrical conductor of radius r_1 having a very high conductivity and coaxial with the cylinder are immersed in a conducting liquid magnetic having permeability μ and mass density ρ (Fig. 244). The current I passes through the circuit. Calculate the height to which the liquid magnetic rises in the cylinder.

8.16. A dielectric cylinder of radius a rotates about its axis at an angular velocity ω in a constant magnetic field whose magnetic induction vector $\mathbf{B}$ is parallel to ω. Find the polarization of the cylinder and the surface charge density of bound charges. The permittivity of the cylinder material is ε.

8.17. A thin conducting disc of conductivity γ is placed in a varying magnetic field whose magnetic induction is given by $\mathbf{B} = \mathbf{B} \cos(\omega t + \varphi)$ and directed at right angles to the plane of the disc. Find the density of eddy currents induced in the disc.

8.18. Find the inductance of a toroid winding consisting of n turns of rectangular cross section with side a if the major radius of the toroid is R.

8.19. A circular loop of radius a rotates about its diameter at a constant angular velocity ω in a uniform magnetic field of magnetic induction $\mathbf{B}$. The ohmic resistance of the loop is R, the axis of rotation is perpendicular to $\mathbf{B}$. Find the current $I(t)$, the torque $M(t)$ decelerating the rotation of the loop and the average power $\langle P \rangle$ spent for maintaining a constant speed of rotation of the loop. For the reference point $t = 0$, take the instant of time when the plane of the loop is perpendicular to $\mathbf{B}$.

8.20. A branch of a circuit consists of two cylindrical coaxial tubes of radii a_1 and a_2 ($a_2 > a_1$) of length l. At one end, the tubes are connected by a conducting plate. Calculate the inductance of the branch of the circuit.

8.21. Two plane closed circular wire loops of radii a_1 and a_2 lie in the same plane separated by a distance d from each other. Assuming that d is sufficiently large so that dipole approximation can be used, find the mutual inductance of the loops.

8.22. The magnetic induction $\mathbf{B}_0$ of the field between the plane parallel poles of an electromagnet can be considered uniform and constant. A plate of area S, made of a paramagnetic having paramagnetic susceptibility χ_p, is inserted into the space beween the poles. The plate surfaces are parallel to the surfaces of the poles of the electromagnet. Find the force acting on the plate.

8.23. Find the radial force acting on a toroid whose parameters are given in Problem 8.13, if current I flows in it.

8.24. Two identical loops with inductance $L = L_{11} = L_{22}$ are arranged so that their mutual inductance $L_{12}^{(0)} = 0$. Superconducting currents I_0 pass through the loops. Then their mutual arrangement is altered so that their mutual inductance becomes L_{12}. Find the currents in the altered state.

8.25. An electric circuit has four junctions. Three of them coincide with the vertices of an equilateral triangle, while the fourth junction coincides with its centre (the point of intersection of medians and bisectors). The capacitances of the branches between the vertices of the triangle are equal to C ($R = 0$, $L = 0$). The inductances connected between the vertices of the triangle and its centre are equal to L ($R = 0$, $C = 0$). Find the resonance frequency of the circuit.

www.EngineeringBooksPDF.com

Answers

8.1. $L=\frac{\mu_0}{\pi} l \ln \frac{d}{r_0}$. 8.2. $\mathbf{B}=(\mu_0/2)\,\mathbf{j}\times\mathbf{r}$. 8.3. $F=\frac{S}{2}\frac{(\mu-\mu_0)\,\mu}{\mu_0} n^2 I^2$.
8.4. $F=\frac{S}{(l+d)^2}\frac{\mu^2}{\mu_0} n^2 I^2$. 8.5. $U=\pi\nu l^2\frac{k-2}{k} B=9.1$ V. 8.6. $E=B_0\omega r/2=156$ V/m.
8.7. $I=\pi b^2 nB\omega\,(R^2+\omega^2L^2)^{-1/2}\sin(\omega t+\varphi_0)$. 8.8. $h=\frac{1}{2}\frac{I^2}{mg}\frac{1}{L_{11}}\{[L_{12}(0)]^2-[L_{12}(h)]^2\}$,
8.9. $F_i=-\frac{1}{2}\frac{I_0^2\omega^2 L_2 L_{12}}{R^2+\omega^2L_2^2}\frac{\partial L_{12}}{\partial\xi_i}$. 8.10. $L_{12}=\mu_0\,(d-\sqrt{d^2-a^2})$. 8.11. $F_x=-\mu_0 I_1 I_2\times$
$\left(\frac{d}{\sqrt{d^2-a^2}}-1\right)$. 8.12. $L_{12}=\frac{\mu_0 nd}{2\pi}\ln(r_2/r_1)$. 8.13. $L=\mu_0 n^2\,(R-\sqrt{R^2-r^2})$.
8.14. $F=\frac{\mu_0 I^2}{4\pi}\ln(r_2/r_1)$. 8.15. $h=\frac{(\mu-\mu_0)\,I^2\ln(r_2/r_1)}{4\pi^2\rho g\,(r_2^2-r_1^2)}$. 8.16. $\mathbf{P}=(\varepsilon-\varepsilon_0)\,B\omega\mathbf{r}$, $\sigma_b=$
$(\varepsilon-\varepsilon_0)\,B\omega a$. 8.17. $\mathbf{j}=(1/2)\,\gamma\omega\mathbf{B}_0\times\mathbf{r}\sin(\omega t+\varphi)$. 8.18. $L=\frac{\mu_0 n^2 a}{2\pi}\ln\left(\frac{2R+a}{2R-a}\right)$.
8.19. $I(t)=\frac{\pi a^2\omega B}{\sqrt{R^2+\omega^2L^2}}\sin(\omega t-\varphi)$, $\tan\varphi=\omega L/R$, $M(t)=-\frac{\pi^2a^4B^2\omega}{\sqrt{R^2+\omega^2L^2}}\sin\omega t\times$
$\sin(\omega t-\varphi)$; $\langle P\rangle=\frac{1}{2}I_0^2R=\frac{1}{2}\frac{(\pi a^2\omega B)^2}{R^2+\omega L^2}R$. 8.20. $L=[\mu_0 l/(2\pi)]\ln(a_2/a_1)$. 8.21. $L_{12}=$
$\mu_0\pi a_1^2a_2^2/(4d^3)$. 8.22. $F=\chi_p SB_0^2/[2\mu_0(1+\chi_p)]$. 8.23. $F=-\mu_0 I^2 n^2(R/\sqrt{R^2-r^2}-1)$.
8.24. $I=I_0L/(L+L_{12})$. 8.25. $\omega_0=(3LC)^{-1/2}$.

www.EngineeringBooksPDF.com

CHAPTER 9

Electromagnetic Waves

A varying magnetic field generates a varying electric field which, in turn, generates a varying magnetic field, and so on. As a result, coupled electric and magnetic fields are created, which form an electromagnetic wave. It is "detached from" charges and currents that have generated this wave. The mode of existence of an electromagnetic wave excludes its immobility in space and the constancy of its field intensities in time.

Sec. 57. Displacement Current

The physical content of displacement current is considered. The role of displacement current in Maxwell's equations is analyzed.

The nature of displacement current. There is no direct current in a circuit containing a capacitor, while an alternating current can flow in it. The magnitude of a quasistationary conduction current is the same for all series-connected elements of the circuit. The conduction current due to the motion of electrons cannot pass through a capacitor since its plates are separated by a dielectric. Consequently, we have to conclude that in a capacitor a certain process closes the conduction current, i.e. ensures in a certain sense the charge exchange between the capacitor plates without transporting a charge between them. The current associated with this process is called **displacement current**.

Let us consider an a.c. circuit including a parallel-plate capacitor (Fig. 245). An electric field of strength $E = \sigma/\varepsilon$ exists between the capacitor plates, where σ is the charge density on the plates and ε is the permittivity of the medium between the plates. The electric displacement between the capacitor plates is given by $D = \sigma = Q/S$, where Q is the charge on each plate and S is the surface area of the plate. The current in the circuit is $I = \partial Q/\partial t$. Hence it follows that

$$I_{\mathrm{d}} = S\frac{\partial D}{\partial t}, \tag{57.1}$$

i.e. a change in the electric displacement between the capacitor plates is the process that closes the conduction current in the circuit. The quantity I in formula (57.1) has a subscript "d" (displacement) in order to emphasize that this

www.EngineeringBooksPDF.com

quantity is not the conduction current between the plates although $I = I_d$. The density of the displacement current in the space between the plates is given by $j_d = I_d/S = \partial D/\partial t$. Considering that the direction of $\mathbf{j}_d$ at each point between the plates of a parallel-plate capacitor coincides with the direction of $\partial D/\partial t$, we can write the following differential equation instead of (57.1):

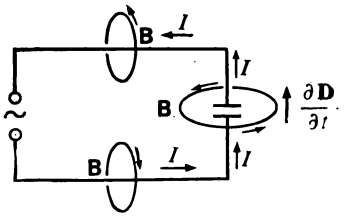


Fig. 245. Displacement current

$$\boxed{\mathbf{j}_d = \partial \mathbf{D}/\partial t.} \qquad (57.2)$$

Taking into account the local nature of this relation, it should be expected that it is independent of the nonlocal model (parallel-plate capacitor) for which it was obtained. This is actually so. Formula (57.2) defines the volume density of the displacement current $\mathbf{j}_d$. The existence of displacement current was theoretically postulated in 1864 by Maxwell and subsequently confirmed in experiments carried out by other scientists.

Why do we call the rate of variation of displacement the displacement current density? The mathematical equality of the quantity $S\partial D/\partial t$, which characterizes the process occurring between the capacitor plates, to the conduction current outside the capacitor plates, i.e. the equality of two quantities referring to different regions of space and having different physical nature, generally does not express any physical law. Consequently, we can call $S\ \partial D/\partial t$ a "current" only formally. In order to attach to this term a physical meaning, it is necessary to prove that $S\ \partial D/\partial t$ possesses properties typical of current although it is not associated with the motion of electric charge as the conduction current. The main property of the conduction current is its ability to generate a magnetic field. Consequently, *the decisive question is whether or not the displacement current generates a magnetic field in the same way as the conduction current does or, to be more precise, whether the quantity* (57.2) *produces the same magnetic field as the volume density of conduction current equal to this quantity does. Maxwell gave an affirmative answer to this question.*

The experimental verification of the correctness of this answer consists in the following. In accordance with Ampère's circuital law, the circulation of vector **B** around a contour embracing the current is equal to $\mu_0 I$. The circulation can be measured with the help of a Rogovskii belt. Moving it along the contour, we note that the circulation remains unchanged even when the Rogovskii belt embraces a capacitor. This just means that the displacement current generates in the capacitor the same magnetic field as that created by the corresponding conduction current. However, the most striking confirmation of generation of a magnetic field by the displacement current is the existence of electromagnetic waves. If the displacement current did not create a magnetic field, electromagnetic waves would not exist.

Maxwell's equations including displacement current. The generation of a magnetic field by a conduction current is described by the equation

$$\operatorname{curl} \mathbf{H} = \mathbf{j}. \qquad (57.3)$$

www.EngineeringBooksPDF.com

Considering the generation of a field by the displacement current, it is necessary to generalize this equation in the following form:

$$\operatorname{curl} \mathbf{H} = \mathbf{j} + \mathbf{j}_d. \tag{57.4}$$

Then, taking into account (57.2), we finally obtain the following equation:

$$\boxed{\operatorname{curl} \mathbf{H} = \mathbf{j} + \partial \mathbf{D}/\partial t,} \tag{57.5}$$

which is one of Maxwell's equations.

Relativistic nature of displacement current. Upon a coordinate transformation of fields, electric and magnetic fields generate each other (see Sec. 11). If there is a nonuniform magnetic field in a certain system of coordinates, this field will vary with time in another system of coordinates, and an electric field will appear simultaneously. This is the evidence of the fact that a varying electric field generates a magnetic field. This does not imply, however, that the generation of the magnetic field by a varying electric field is a new fundamental phenomenon in electricity and magnetism. This situation is similar to that considered in detail in connection with the electromagnetic induction in Secs. 45 and 46. *The generation of a magnetic field by a varying electric field is a fundamental phenomenon in nature.*

The formal equality of the displacement current in a capacitor and the conduction current in the wires connected to its plates does not express any physical law. A new physical law states that the displacement current creates the same magnetic field as does the conduction current corresponding to it.

Example 57.1. *There are two layers of a weakly conducting material, having conductivities* γ_1 *and* γ_2 *and permittivities* ε_1 *and* ε_2, *between the plates of a parallel-plate capacitor. The thickness of the layers are* a_1 *and* a_2 *respectively* (Fig. 246). *The area of capacitor plates is* S. *Investigate the process of establishment of current in the circuit if at the moment* $t = 0$ *a constant potential difference* U_0 *was applied across the capacitor plates. Analyze the processes originating after interruption of the circuit and upon shunting the source of extraneous e.m.f.s.*

At the moment when the voltage is applied, the surface charge cannot appear immediately at the boundary between the layers. Therefore, at the initial instant of time the system under consideration behaves as if the conductivity of the substance between the plates were equal to zero, i.e. as an ideal capacitor. Hence, the displacement appearing in the space between the plates is

$$D = \varepsilon_1 E_1 = \varepsilon_2 E_2, \tag{57.6}$$

where E_1 and E_2 are the electron field strengths in the first and second layers respectively. This formula takes into account the continuity of D. Since the potential difference between the plates is U_0, we have

$$\int_{(1)}^{(2)} \mathbf{E} \cdot d\mathbf{l} = a_1 E_1 + a_2 E_2 = U_0, \tag{57.7}$$

where the path integral from the first plate to the second is taken along the normal to the plates. It follows from (57.6) and (57.7) that

$$D = \varepsilon_1 \varepsilon_2 U_0 / (\varepsilon_2 a_1 + \varepsilon_1 a_2). \tag{57.8}$$

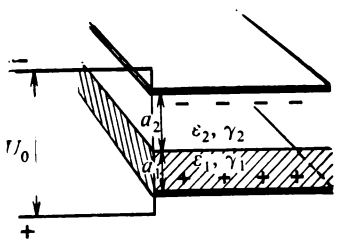


Fig. 246. Two-layered parallel-plate capacitor with a leak

At the initial instant, the entire current is a displacement current. It is equal to infinity since the potential difference is applied instantaneously, and D instanta-

www.EngineeringBooksPDF.com

neously grows from zero to a value determined from formula (57.8). The surface charge density on the plates also increases instantaneously from 0 to $\sigma_1 = -\sigma_2 = D$.

The instantaneous variations of electric displacement from zero to finite values are due to a very high rate of generation of polarization of the substance under the action of an external field. Polarization may appear during a time typical of intramolecular processes.

At the instants of time following the application of the voltage, the conduction current starts to increase, and, after a sufficiently long period of time ($t \to \infty$), the equilibrium value of the current density is established:

$$j = \gamma_1 E_1 = \gamma_2 E_2 = \gamma_1\gamma_2 U_0/(\gamma_2 a_1 + \gamma_1 a_2), \tag{57.9}$$

where relation (57.7) is taken into account. Since the conductivity is nonuniform, the boundary between the layers is charged with the surface density

$$\sigma = D_{2n} - D_{1n} = \varepsilon_2 E_2 - \varepsilon_1 E_1 = (\varepsilon_2\gamma_1 - \varepsilon_1\gamma_2)\, U_0/(\gamma_2 a_1 + \gamma_1 a_2), \tag{57.10}$$

where we used the boundary condition (17.36), since the electric field strength does not depend on time.

Under transient conditions, before stationary values of (57.9) and (57.10) have been attained, the conduction currents in the first and second layers are different, and the charge density at the interface between the layers increases with time. Under such conditions, it is the sum of volume densities of conduction and displacement currents (called the total volume density of current) that has the same value in both layers:

$$j_t = \gamma_1 E_1 + \frac{\partial}{\partial t}(\varepsilon_1 E_1) = \gamma_2 E_2 + \frac{\partial}{\partial t}(\varepsilon_2 E_2)\,. \tag{57.11}$$

Eliminating E_2 from (57.11), we obtain, using (57.7), the following equation for E_1:

$$\frac{dE_1}{dt} + \frac{E_1}{\tau} = \frac{\gamma_2 U_0}{\varepsilon_2 a_1 + \varepsilon_1 a_2}, \tag{57.12}$$

where

$$\tau = (\varepsilon_1 a_2 + \varepsilon_2 a_1)/(\gamma_1 a_2 + \gamma_2 a_1). \tag{57.13}$$

A similar equation can be obtained for E_2.

The solution of these equations under initial condition (57.8) has the form

$$E_1 = \frac{\gamma_2 U_0}{\gamma_2 a_1 + \gamma_1 a_2}\,(1 - e^{-t/\tau}) + \frac{\varepsilon_2 U_0}{\varepsilon_2 a_1 + \varepsilon_1 a_2}\, e^{-t/\tau}, \tag{57.14}$$

$$E_2 = \frac{\gamma_1 U_0}{\gamma_2 a_1 + \gamma_1 a_2}\,(1 - e^{-t/\tau}) + \frac{\varepsilon_1 U_0}{\varepsilon_2 a_1 + \varepsilon_1 a_2}\, e^{-t/\tau}. \tag{57.15}$$

As $t \to \infty$ these solutions assume the form (57.9) as expected.

The surface charge density between the layers varies according to the law

$$\sigma = \varepsilon_2 E_2 - \varepsilon_1 E_1 = \frac{\varepsilon_2\gamma_1 - \varepsilon_1\gamma_2}{\gamma_2 a_1 + \gamma_1 a_2}\,(1 - e^{-t/\tau})\, U_0. \tag{57.16}$$

For $t = 0$, the surface charge density $\sigma = 0$ as well. As $t \to \infty$, it tends to (57.10) as it should be expected.

The total current density can be obtained from (57.11) by taking into account Eqs. (57.14) and (57.15):

$$j_t = \gamma_1 E_1 + \frac{\partial}{\partial t}(\varepsilon_1 E_1) = \gamma_2 E_2 + \frac{\partial}{\partial t}(\varepsilon_2 E_2) = \left[\frac{\gamma_1\gamma_2}{\gamma_2 a_1 + \gamma_1 a_2} + \left(\gamma_1 - \frac{\varepsilon_1}{\tau}\right)\right.$$
$$\left. \times \left(\frac{\varepsilon_2}{\varepsilon_2 a_1 + \varepsilon_1 a_2} - \frac{\gamma_2}{\gamma_2 a_1 + \gamma_1 a_2}\right) e^{-t/\tau} + \frac{\varepsilon_1\varepsilon_2}{\varepsilon_2 a_1 + \varepsilon_1 a_2}\,\delta(t)\right] U_0, \tag{57.17}$$

where $\delta(t)$ is the delta function. It appears due to the fact that at $t = 0$ the displacement instantaneously grows from 0 to (57.8). In other words, while calculating the time derivative

www.EngineeringBooksPDF.com

in (57.17), we have

$$\frac{\partial(\varepsilon_1 E_1)}{\partial t}=\varepsilon_1\frac{\partial E_1}{\partial t}+\frac{\varepsilon_1\varepsilon_2 U_0}{\varepsilon_2 a_1+\varepsilon_1 a_2}\,\delta(t), \tag{57.18}$$

while for calculating $\partial E_1/\partial t$ in (57.18), we use expression (57.14) which is valid for all $t > 0$.

The above analysis shows that the voltage distribution in various branches of the circuit at the moment when the external voltage is applied may considerably differ from the distribution under steady-state conditions. This circumstance should be taken into account while calculating the circuits.

When the circuit is disconnected, $j_t = 0$, and hence Eqs. (57.11) assume the form

$$\gamma_1 E_1+\frac{\partial(\varepsilon_1 E_1)}{\partial t}=0,\quad \gamma_2 E_2+\frac{\partial(\varepsilon_2 E_2)}{\partial t}=0. \tag{57.19}$$

The fields are independent of each other. Under steady-state conditions, we obtain, in accordance with (57.14) and (57.15),

$$E_{10}=\gamma_2 U_0/(\gamma_2 a+\gamma_1 a_2),\quad E_{20}=\gamma_1 U_0/(\gamma_2 a_1+\gamma_1 a_2). \tag{57.20}$$

The solution of Eq. (57.19) under initial conditions (57.20) has the form

$$E_1=\frac{\gamma_2 U_0}{\gamma_2 a_1+\gamma_1 a_2}\,e^{-1/\tau_1},\qquad E_2=\frac{\gamma_1 U_0}{\gamma_2 a_1+\gamma_1 a_2}\,e^{-t/\tau_2}, \tag{57.21}$$

where $\tau_1 = \varepsilon_1/\gamma_1$, $\tau_2 = \varepsilon_2/\gamma_2$.

The potential difference across disconnected terminals varies in accordance with the law

$$U=aE_1+a_2E_2=\frac{U_0}{\gamma_2 a_1+\gamma_1 a_2}\,[\gamma_2 a_1 e^{-1/\tau_1}+\gamma_1 a_2 e^{-1/\tau_2}]. \tag{57.22}$$

The surface charge density at the interface between the layers in the capacitor is given by

$$\sigma=\varepsilon_2E_2-\varepsilon_1E_1=\frac{U_0}{\gamma_2 a_1+\gamma_1 a_2}\,[\varepsilon_2\gamma_1 e^{-t/\tau_1}-\varepsilon_1\gamma_2 e^{-t/\tau_2}]. \tag{57.23}$$

Upon shunting the source of extraneous e.m.f.s, $U_0 = 0$ and Eqs. (57.7) and (57.12) become

$$a_1E_1+a_2E_2=0, \tag{57.24}$$

$$\frac{dE_1}{dt}+\frac{E_1}{\tau}=0, \tag{57.25}$$

where τ is defined by (57.13). The initial condition for $t = 0$ can be found from (57.10) combined with (57.24):

$$\varepsilon_2E_{20}-\varepsilon_1E_{10}=-\left(\frac{\varepsilon_2 a_1}{a_2}+\varepsilon_1\right)E_{10}=\frac{\varepsilon_2\gamma_1-\varepsilon_1\gamma_2}{\gamma_2 a_1+\gamma_1 a_2}\,U_0. \tag{57.26}$$

The solution of Eq. (57.25) for the initial value of E_{10} from (57.26) is

$$E_1=-E_2a_2/a_1=-\frac{(\varepsilon_2\gamma_1-\varepsilon_1\gamma_2)\,a_2U_0}{(\varepsilon_2 a_1+\varepsilon_1 a_2)(\gamma_2 a_1+\gamma_1 a_2)}\,e^{-t/\tau}. \tag{57.27}$$

The current in the circuit and the surface charge density between the layers are given by

$$I=\left[\left(\frac{\gamma_1\varepsilon_2-\gamma_2\varepsilon_1}{\varepsilon_1 a_2+\varepsilon_2 a_1}\right)^2\frac{a_1a_2U_0}{(\gamma_2 a_1+\gamma_1 a_2)}\,e^{-t/\tau}-\frac{\varepsilon_1\varepsilon_2U_0}{\varepsilon_1 a_2+\varepsilon_2 a_1}\,\delta(t)\right]S, \tag{57.28}$$

$$\sigma=\frac{\varepsilon_2\gamma_1-\varepsilon_1\gamma_2}{\gamma_2 a_1+\gamma_1 a_2}\,U_0e^{-t/\tau}. \tag{57.29}$$

www.EngineeringBooksPDF.com

The term in (57.28) containing the delta function appears due to the fact that at the moment of shunting the source of extraneous e.m.f.s, the displacement vector **D** abruptly changes from the value corresponding to formula (57.9) for steady-state conditions to the value corresponding to the initial conditions at $t = 0$, given by formula (57.26).

Sec. 58. Maxwell's Equations

The physical meaning, conditions of applicability, completeness and compatibility of Maxwell's equations are considered.

Maxwell's equations. Equations (57.5), (46.5), (36.4) and (17.30), obtained in previous sections as a result of generalization of experimental facts, form the system of Maxwell's equations:

$$\begin{aligned} &\operatorname{curl} \mathbf{H} = \mathbf{j} + \partial \mathbf{D}/\partial t, \quad \text{(I)} \qquad \operatorname{div} \mathbf{B} = 0, \quad \text{(III)} \\ &\operatorname{curl} \mathbf{E} = -\partial \mathbf{B}/\partial t, \quad \text{(II)} \qquad \operatorname{div} \mathbf{D} = \rho. \quad \text{(IV)} \end{aligned} \tag{58.1a}$$

These equations, which are called the **field equations**, are applicable for describing all macroscopic electromagnetic phenomena. While considering a specific situation, we should take into account the electromagnetic properties of material media. In many cases, this is achieved through relations (17.31), (38.24) and (16.5)

$$\mathbf{D} = \varepsilon \mathbf{E}, \ \mathbf{B} = \mu \mathbf{H}, \ \mathbf{j} = \gamma \mathbf{E} \ \text{(V)}, \tag{58.1b}$$

which are usually called constitutive relations. There are, however, many phenomena (for example, nonlinear phenomena) for which the constitutive relations have a different form and the derivation of these equations is an independent scientific problem.

Physical meaning of Maxwell's equations. Equation (I) expresses the law according to which a magnetic field is generated by conduction currents and displacement currents, which are two possible sources of a magnetic field.

Equation (II) expresses the law of electromagnetic induction and points to a varying magnetic field as one of the possible sources of an electric field. The other source of an electric field is associated with electric charges. The field generated by the charges is described by equation (IV) which expresses Coulomb's law. The physical meaning of equation (III) was considered in detail in connection with (36.4).

The constitutive equations (V) express the relations between field vectors and currents, which take into account the properties of a material medium. Dielectric properties, which are described phenomenologically by polarization, are taken into consideration in permittivity ε. Magnetic properties which are phenomenologically described by magnetization are taken into account in permeability μ. The conducting properties of the medium are included in conductivity γ.

www.EngineeringBooksPDF.com

Field equations are linear equations based on the principle of superposition, which is an independent experimental fact.

Conditions of applicability of Maxwell's equations. The substantiation of Eqs. (58.1) shows that they are applicable under the following conditions:

(1) material bodies are at rest relative to the electromagnetic field;

(2) material constants ε, μ and γ may depend on coordinates, but they should be independent of time and field vectors;

(3) there are no permanent magnets and ferromagnetic bodies in the field.

The motion of the medium can be taken into consideration in the simplest manner as follows. The presence of the medium in electric and magnetic phenomena is ultimately reduced to the presence of charges in the medium and their motion. Consequently, we may proceed from Maxwell's equations for vacuum ($\varepsilon = \varepsilon_0$, $\mu = \mu_0$) and take into account the medium in the same way as it was done in Secs. 17 and 38 but considering the motion of charges as well. As a result, the form of field equations (58.1a) remains unchanged, and the motion of the medium is accounted for by modifying the constitutive relations (58.1b) which become dependent on the velocity of the medium and are considerably complicated. Moreover, they are no longer the relations between two quantities (say **D** and **E**) but become "linked". For example, the conduction current density now depends on the magnetic induction in addition to the electric field strength, and so on.

The field outside permanent magnets and ferromagnetics can be described with the help of Maxwell's equations if we assume that their magnetization is known. However, it is impossible to solve the problem by using Maxwell's equations in the presence of ferromagnetics in space if, for example, currents are given. Maxwell's equations are inapplicable in this case.

Completeness and compatibility of Maxwell's equations. Using the constitutive equations (58.1b), we can exclude the quantities **D**, **H**, and **j** from the field equations (58.1a), as a result of which they are transformed into equations in terms of vectors **E** and **B**, i.e. in six unknown independent components of these quantities. On the other hand, the number of scalar equations in (58.1a) is equal to eight. It turns out that we have eight equations in six unknown quantities, i.e. the number of equations exceeds the number of unknowns, which is inadmissible since the system of equations seems to be overdetermined.

Actually, however, the system is not overdetermined and no difficulties of such kind are encountered. This is due to the fact that equations (I), (IV) and (II) and (III) have the same differential results and are not independent, although we cannot say that some of them are obtained from other equations.

In order to prove the identity of differential results of equations (II) and (III), let us apply to both sides of Eq. (II) the divergence operation and differentiate both sides of Eq. (III) with respect to time. In both cases we obtain the same equation $\partial \operatorname{div} \mathbf{B}/\partial t = 0$.

Let us prove that, taking into account the law of charge conservation

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} = 0, \tag{58.2}$$

www.EngineeringBooksPDF.com

we can treat Eq. (IV) as the differential result of Eq. (I). For this purpose, we apply the divergence operation to both sides of Eq. (I):

$$\operatorname{div} \mathbf{j} + \partial \operatorname{div} \mathbf{D}/\partial t = 0, \tag{58.3}$$

where div curl $\mathbf{H} = 0$. Comparing (58.3) and (58.2), we find that the following equation must hold:

$$\operatorname{div} \mathbf{D} = \rho, \tag{58.4}$$

which coincides with Eq. (IV). Thus, we have proved that Eq. (IV) is the differential result of Eq. (I), taking into account the law of charge conservation.

The existence of two differential relations between Eqs. (I-IV) makes this system compatible. A more detailed analysis shows that the system of equations is complete, and its solution is unique for the given boundary and initial conditions. The proof of the uniqueness of the solution generally boils down to the following. If there are two different solutions, their difference will also be a solution on account of linearity of Maxwell's equations but only for zero charges and currents and zero initial and boundary conditions. Using the expression for the electromagnetic field energy and the law of energy conservation, we conclude that the difference of the solutions is identically equal to zero, which means that the solutions are identical. This proves the uniqueness of the solution of Maxwell's equations.

Sec. 59. The Law of Conservation of Electromagnetic Energy. Energy Flux

The mathematical formulation of the law of conservation of energy is given and the concept of electromagnetic energy flux is analyzed.

Formulation of the law of conservation of energy. The electric and magnetic field energy is defined by formulas (18.16) and (47.26). The forces responsible for the work in electric and magnetic fields were investigated in Secs. 19 and 39. The work of alternating current was defined in Sec. 49, and the heating effect of current was studied in Sec. 27. The law of conservation of energy requires that all these processes be represented in the form of the law of conservation and mutual transformation of different forms of energy. Since the sources of electromagnetic energy are separated in space from the energy consumers, the concept of energy transport, characterized by the energy flux, is introduced.

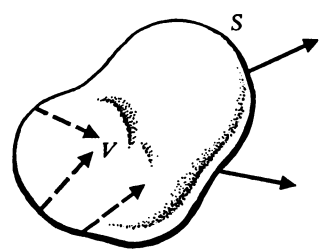


Fig. 247. To the formulation of the law of conservation of energy

Let us consider a certain closed volume V which contains an electromagnetic field and currents (Fig. 247). Joule's heat liberated by the current in

www.EngineeringBooksPDF.com

this volume is given by

$$P = \int_V \mathbf{j} \cdot \mathbf{E}\, dV. \tag{59.1}$$

In order to simplify the calculations, it is assumed that there are no other energy transformations in this volume. Substituting into this equation the expressions for $\mathbf{j}$ from Eq. (58.1a), we obtain

$$P = \int_V \mathbf{E} \cdot \operatorname{curl} \mathbf{H}\, dV - \int_V \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}\, dV. \tag{59.2}$$

Using formula (A.15), we have

$$\operatorname{div} (\mathbf{E} \times \mathbf{H}) = \operatorname{curl} \mathbf{E} \cdot \mathbf{H} - \mathbf{E} \cdot \operatorname{curl} \mathbf{H} \tag{59.3}$$

and hence

$$P = -\int_V \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{H}\, dV - \int_V \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}\, dV - \int_V \operatorname{div} (\mathbf{E} \times \mathbf{H})\, dV, \tag{59.4}$$

where $\operatorname{curl} \mathbf{E} = -\partial \mathbf{B}/\partial t$. Considering that $\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{1}{2} \frac{\partial (\mathbf{H} \cdot \mathbf{B})}{\partial t}$ and $\frac{\mathbf{E} \cdot \partial \mathbf{D}}{\partial t} = \frac{1}{2} \frac{\partial (\mathbf{E} \cdot \mathbf{D})}{\partial t}$ and transforming the last integral in (59.4) into an integral over the surface σ bounding the volume V in accordance with the Gauss theorem we finally obtain

$$P = -\frac{\partial}{\partial t} \left[\frac{1}{2} \int_V (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H})\, dV \right] - \int_\sigma \mathbf{E} \times \mathbf{H} \cdot d\boldsymbol{\sigma} \tag{59.5}$$

Here we denoted the surface by σ in order to use the letter S for designating the flux density of electromagnetic energy.

Energy flux. The quantity

$$W = \frac{1}{2} \int_V (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H})\, dV \tag{59.6}$$

characterizes the electromagnetic energy contained in volume V. The quantity

$$\boxed{\mathbf{S} = \mathbf{E} \times \mathbf{H}} \tag{59.7}$$

is the **density of the energy flux through the surface** bounding the volume V, called **Poynting's vector.** This quantity was obtained by J.H. Poynting (1852-1914) in 1884. However, ten years earlier, i.e. in 1874, N.A. Umov (1846-1915) carried out a general analysis of the energy flow in bodies, which was characterized by the corresponding energy flux. (Therefore, vector (59.7) is also called the Umov-Poynting vector.) It is more convenient to write Eq. (59.4) in the form

$$\boxed{\frac{\partial W}{\partial t} = -P - \int_\sigma \mathbf{S} \cdot d\boldsymbol{\sigma},} \tag{59.8}$$

www.EngineeringBooksPDF.com

i.e. *a change in the electromagnetic field energy in a volume occurs at the expense of the work of conduction currents in this volume and the energy flux through the surface bounding the volume.* If the electric field energy remains unchanged ($\partial W/\partial t = 0$), then (see (59.8))

$$P = -\int_{\sigma} \mathbf{S} \cdot d\boldsymbol{\sigma}. \tag{59.9}$$

Consequently, *all the work performed in a closed volume is at the expense of the electromagnetic energy flux through the surface bounding the volume.*

Equation (59.8) expresses the law of conservation of elecromagnetic energy.

It should be emphasized that Eq. (59.8) is just an expression for the energy conservation law and not its proof.

The law of conservation of energy as a universal law of nature is presumed to be known while developing the theory of electricity and magnetism. Proceeding from this law as a universal law, we can find the mathematical expression for the volume density of energy of electric and magnetic fields and the energy density of these fields as well as the flux density of electromagnetic energy. It is also possible to establish the relation between these quantities, which expresses the idea of transformation of electromagnetic energy. The physical quantity P in formula (59.8) takes into account the possibility of interconversion of different forms of energy.

Sec. 60. Transmission of Electromagnetic Energy along Transmission Lines

The physical aspects of energy transmission along the transmission lines are discussed and the main characteristics of transmission lines are given.

Compensation of energy losses due to liberation of Joule's heat. Let us consider a portion of a circular conductor of radius r, carrying a direct current of volume density $\mathbf{j}$ (Fig. 248). In accordance with the differential form of Ohm's law, the electric field parallel to the conductor axis is given by

$$\mathbf{E} = \mathbf{j}/\gamma. \tag{60.1}$$

On account of the boundary condition of continuity of tangential components of electric field, the same field exists outside the conductor near its surface.

Let us use formula (59.9) for calculating the flux of electromagnetic energy through a closed surface of a cylinder whose lateral surface coincides with the surface of the conductor of length l and the bases are the circular cross sections of the conductor.

The magnetic field strength on the surface of the conductor is directed along a tangent to the surface and lies in the plane perpendicular to the axis of the

www.EngineeringBooksPDF.com

conductor (and to the vector j) (see Fig. 248). Its value is given by

$$H = j\pi r^2/(2\pi r) = j/(2r). \tag{60.2}$$

Thus, the Poynting vector (59.7) is directed along the radius to the conductor axis and its magnitude is

$$S = EH = j^2 r/(2\gamma). \tag{60.3}$$

This means that *the electromagnetic energy flows into the conductor from the surrounding medium through the lateral surface.* There is no energy flow through

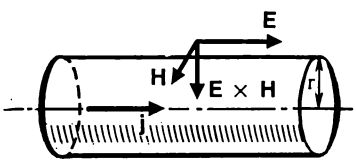


Fig. 248. Compensation of energy losses on liberation of Joule's heat

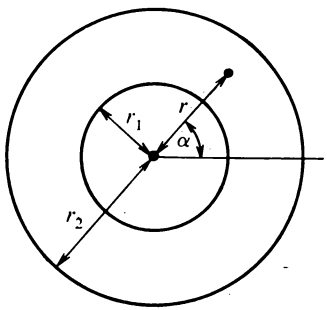


Fig. 249. Transmission of electromagnetic energy along a current-carrying cable

the cylinder bases. The amount of energy flowing per second into a portion of length l of the conductor is given by

$$P = S \cdot 2\pi r l = (j^2/\gamma)\, \pi r^2 l. \tag{60.4}$$

According to Joule's law, the amount of heat liberated over the length l of the conductor per second is

$$P' = (j^2/\gamma)\, \pi r^2 l. \tag{60.5}$$

A comparison of formulas (60.4) and (60.5) shows that the entire energy liberated in a conductor in the form of heat upon the passage of electric current is supplied from the surrounding medium through the lateral surface of the conductor. Consequently, *the energy supplied by electric current flows in the space surrounding the conductor. The wires play the role of guides along which the electromagnetic energy is transmitted.* The flux density of electromagnetic energy at each point of space is determined by the Poynting vector.

Energy transmission along a cable. A current passes in the core of a cable in one direction and in its sheath in the opposite direction (Fig. 249). The space between the core and the sheath is filled with a dielectric. In order to simplify calculations, we assume that the resistance of the cable is negligibly small and can be neglected, i.e. we can assume that the energy is transmitted without losses. Then the potential is constant along the core and the sheath, and the potential drop between them takes place in the energy consumer (load) and

www.EngineeringBooksPDF.com

in the source (extraneous e.m.f.). Suppose that the potential drop across a load is U. This means that the potential difference between the core and the sheath is equal to U. Consequently, there exists an electric field between them. Since the problem is axially symmetric and the current flows along the cable without resistance, this field is directed along the radius, and tangential component E_α is absent. The Z-axis of the cylindrical coordinate system coincides with the cable axis. The magnetic field lines are concentric circles with the centre on the cable axis. The field strength has a nonzero value only in the space between the core and the sheath and is equal to zero outside the cable. The radial component of the Poynting vector is equal to zero. The Maxwell equation div $\mathbf{D} = \rho$ for the space between the core and the sheath has the following form:

$$\operatorname{div} \mathbf{E} = \frac{1}{r} \frac{\partial}{\partial r} (rE_r) = 0, \tag{60.6}$$

where we use the representation of the divergence operation in cylindrical coordinates and take into account the fact that the axial and tangential components of E are missing. It follows from this equation that

$$E_r = a_0/r, \tag{60.7}$$

where a_0 is the integration constant determined from the conditions of the problem. The potential difference between the core and the sheath is given by

$$U = \int_{r_1}^{r_2} E_r \, dr = a_0 \ln (r_2/r_1), \tag{60.8}$$

which allows us to determine the magnitude of the constant $a_0 = U/\ln (r_2/r_1)$. Using this value in formula (60.7), we obtain

$$E_r = \frac{U}{\ln (r_2/r_1)} \frac{1}{r}. \tag{60.9}$$

The magnetic field strength in the cable is

$$H_\alpha = I/(2\pi r), \tag{60.10}$$

in accordance with Ampère's circuital law and in view of the axial symmetry of the problem. Combining (60.9) and (60.10), we obtain

$$S_z = E_r H_\alpha = \frac{1}{2\pi} \frac{UI}{\ln (r_2/r_1)} \frac{1}{r^2}. \tag{60.11}$$

This quantity is the rate of flow of electromagnetic energy parallel to the cable axis in the space between the core and the sheath. There is no energy flow outside the cable or in the central core and sheath since there is no electric field in them at all in the assumption of zero resistance. The electromagnetic power flowing in 1 s through the cross-sectional area of the cable is given by

$$P = \int_{\sigma} S_z \, d\sigma = \frac{1}{2\pi} \int_0^{2\pi} d\alpha \int_{r_0}^{r_2} \frac{dr}{r} \cdot \frac{UI}{\ln (r_2/r_1)} = UI. \tag{60.12}$$

www.EngineeringBooksPDF.com

The power developed when the current I passes through a load at a potential difference U is

$$P_1 = IU. \tag{60.13}$$

A comparison of formulas (60.12) and (60.13) shows that *the entire energy used by a consumer flows along the cable in the space between the core and the sheath in the form of electromagnetic energy.*

The situation is basically the same if an alternating current of a moderate frequency passes through a cable. When the current in the cable reverses its direction, the components E_r and H_α of the field vectors also reverse the direction, the direction of the Poynting vector remaining the same. Therefore, although

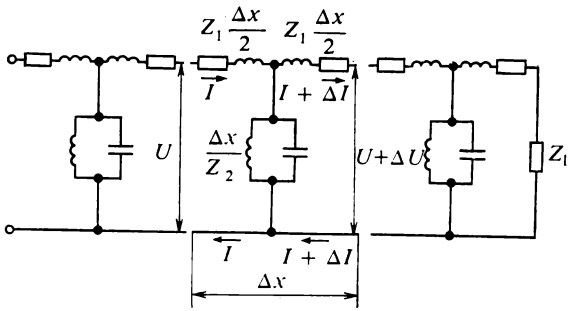


Fig. 250. Equivalent diagram of an a.c. transmission line

the direction of current is reversed, the direction of transmission of electromagnetic energy remains the same: power is always transmitted from the source to the consumer.

In other types of transmission lines, the mode of energy transmission does not change in principle and only the configuration of the fields and the paths along which the energy is transmitted become more complicated.

Transmission line for an alternating current. The methods presented in Sec. 8 can be used to describe the current in transmission lines for not very high frequencies and sufficiently short distances, when quasi-stationary conditions can be assumed to be satisfied. Otherwise, the situation becomes more complicated, which can be seen clearly from the fact that the current in different regions of the line is different at the same instant of time. Any portion of a conductor has a certain inductance and capacitance, which makes the entire transmission line a circuit with continuously distributed resistances, capacitances and inductances.

Equations for current and voltage. First of all, we must find the law according to which the current and voltage between the conductors vary along the line. The schematic diagram of distribution of inductance, capacitance and resistance is shown in Fig. 250. We denote by L, C, and R the inductance, capacitance, and resistance per metre of the transmission line. Impedances Z_1 and Z_2 are also referred to 1 m length. The branch Δx of the line has a series-connected im-

www.EngineeringBooksPDF.com

pedance corresponding to the complex impedance

$$Z_1 \Delta x = (R_1 + i\omega L)\, \Delta x, \tag{60.14}$$

and a parallel-connected impedance Z_2 corresponding to the complex conductivity

$$\frac{1}{Z_2} \Delta x = \left(\frac{1}{R_2} + i\omega C\right) \Delta x. \tag{60.15}$$

Suppose that voltage U is applied to the beginning of the branch Δx of the line, the current being equal to I. At the end of the branch, these quantities are equal to $U + \Delta U$ and $I + \Delta I$ respectively. Here and below, the losses through the insulation are neglected.

Let us apply Kirchhoff's second law to the entire external circuit, taking the counterclockwise direction of circumvention as the positive direction:

$$-Z_1 \frac{\Delta x}{2} (I + \Delta I) - Z_1 \frac{\Delta x}{2} I = U + \Delta U - U. \tag{60.16}$$

Dividing this equation by Δx, we obtain

$$-Z_1\, \Delta I/2 - Z_1 I = \Delta U/\Delta x. \tag{60.17}$$

If $\Delta x \to 0$, the first term on the left-hand side of (60.17) tends to zero ($\Delta I \to$ $\to 0$), and we get

$$\frac{dU}{dx} = -Z_1 I. \tag{60.18}$$

Similarly, Kirchhoff's law applied to the subcircuit containing the impedance $Z_2/\Delta x$ gives

$$\frac{Z_2}{\Delta x} \Delta I - Z_2 \frac{\Delta x}{2} I = -U. \tag{60.19}$$

As $\Delta x \to 0$, we obtain

$$\frac{dI}{dx} = -\frac{1}{Z_2} U. \tag{60.20}$$

Differentiating both sides of (60.18) with respect to x and expressing dI/dx with the help of (60.20), we obtain the following equation for U:

$$\frac{d^2U}{dx^2} = \frac{Z_1}{Z_2} U. \tag{60.21}$$

Similarly, differentiating (60.20) with respect to x and using Eq. (60.18), we arrive at the following equation for current:

$$\frac{d^2I}{dx^2} = \frac{Z_1}{Z_2} I. \tag{60.22}$$

Equations (60.21) and (60.22) are called the **transmission line equations.**

Characteristic impedance and propagation constant. The general solution of the transmission line equations has the form (say, for U)

$$U = Ae^{-\alpha x} + Be^{\alpha x}. \tag{60.23}$$

www.EngineeringBooksPDF.com

Substituting this equation into (60.21), we obtain the following expression for α which is called the **propagation constant**:

$$\alpha = \sqrt{Z_1/Z_2}. \tag{60.24}$$

The solution of Eq. (60.22) also has a similar form:

$$I = A_1 e^{-\alpha x} + B_1 e^{\alpha x}. \tag{60.25}$$

Substituting the solutions of (60.23) and (60.25) into (60.18) and (60.20), we can find the relation between the constants A, B, A_1 and B_1:

$$A_1 = A/Z_1, \quad B_1 = -B/Z_1, \tag{60.26}$$

where

$$Z_1 = \sqrt{Z_1 Z_2} \tag{60.27}$$

is the **characteristic impedance of the line.** In order to clarify the meaning of this quantity, we assume that a line of length l terminates at a load whose impedance is equal to the characteristic impedance (Fig. 250). Using Eqs. (60.23)-(60.27) for the output voltage of the line, i.e. across the load Z_1, we can write

$$U_{\text{load}} = I_{\text{load}} Z_1, \tag{60.28}$$

or

$$Ae^{-\alpha l} + Be^{\alpha l} = Z_1 \left(\frac{A}{Z_1} e^{-\alpha l} - \frac{B}{Z_1} e^{\alpha l} \right). \tag{60.29}$$

Hence it follows that $B = 0$, and $A = U_{1n}$ is the input voltage of the line for $x = 0$. Thus, the voltage and current in the line are defined by the following expressions:

$$U = U_{1n} e^{-\alpha x}, \quad I = U_{1n} e^{-\alpha x}/Z_1. \tag{60.30}$$

Consequently, the input impedance of the line is equal to the characteristic impedance:

$$Z_{1n} = U_{1n}/I_{1n} = Z_1. \tag{60.31}$$

This means that *if a line terminates in a load with the characteristic impedance, its input impedance is equal to the characteristic impedance regardless of the line length, i.e. in this case the current is transmitted along the line at a constant ratio between the voltage and current.*

Characteristic resistance. In most practical cases, ohmic resistances of the elements of a line are much smaller than the corresponding inductive and capacitive reactances ($R_1 \ll \omega L$, $1/R_2 \ll \omega C$) and can therefore be neglected. Under this condition, the characteristic impedance

$$Z_1 = \sqrt{Z_1 Z_2} = \sqrt{\frac{R_1 + i\omega L}{1/R_2 + i\omega C}} = \sqrt{\frac{L}{C}} \tag{60.32}$$

is a real quantity called the **characteristic resistance.**

The characteristic resistance depends on the shape and size of conductors, the distance between them and other factors which determine the capacitance

www.EngineeringBooksPDF.com

and inductance of a branch of the line. For example, the characteristic resistance of parallel cylindrical conductors of radius a, the distance between whose axes is equal to D, is given by

$$Z_1 = 276 \log (D/a). \tag{60.33}$$

It is assumed that the conductors are in a medium whose relative permittivity is close to unity (vacuum, air, and so on).

Velocity of propagation. We have considered the voltage and current distribution along a transmission line at a certain instant of time. If the input current and voltage vary periodically at a frequency ω, they will vary at the same frequency in all branches of the line. Under the conditions when the characteristic impedance (60.32) is a real quantity, the propagation constant α [see (60.24)] is a purely imaginary quantity:

$$\alpha = i\omega \sqrt{LC}. \tag{60.34}$$

Therefore, assuming that the time variation of quantities follows the law $\exp(i\omega t)$, we can write on the basis of (60.30)

$$\begin{aligned} U(x, t) &= U_0 \exp [i(\omega t - \omega \sqrt{LC}\, x)], \\ I(x, t) &= (U_0/\sqrt{L/C}) \exp [i(\omega t - \omega \sqrt{LC}\, x)]. \end{aligned} \tag{60.35}$$

This formula describes a wave having the frequency ω and propagating along the X-axis at a velocity

$$v = 1/\sqrt{LC}. \tag{60.36}$$

It should be recalled that L and C in this formula are the inductance and capacitance per metre of the transmission line. The capacitances and inductances per metre length of two thin cylindrical wires of radius a separated by a distance D in a vacuum are given by

$$C = \varepsilon_0/[2 \ln (D/a)], \quad L = 2\mu_0 \ln (D/a) \tag{60.37}$$

and hence the velocity of propagation of the wave is

$$v = 1/\sqrt{LC} = 1/\sqrt{\varepsilon_0 \mu_0} = c. \tag{60.38}$$

Reflection. If the resistance of the load is equal to the characteristic resistance, the entire power transmitted by the line is absorbed by the load. The load and the transmission line are said to be matched in this case. *If there is no such matching, a part of energy is reflected from the load and transmitted along the line in a direction opposite to the initial energy flow.*

Let us consider by way of an example a transmission line short-circuited at the end, when $U_1 = 0$. In this case, Eqs. (60.23) and (60.25) assume the form

$$0 = Ae^{-i\beta l} + Be^{i\beta l}, \tag{60.39}$$

$$I_1 = Ae^{-i\beta l}/\rho - Be^{i\beta l}/\rho, \tag{60.40}$$

www.EngineeringBooksPDF.com

where the notation $\rho = \sqrt{L/C}$ and $\beta = \omega\sqrt{LC}$ are introduced in order to simplify the formulas. Solving these equations for A and B, we obtain

$$A = I_1\rho e^{i\beta l}/2, \quad B = -I_1\rho e^{-i\beta l}/2. \tag{60.41}$$

Consequently, expressions (60.23) and (60.25) for voltage and current in the transmission line are written as follows:

$$U = I_0 \frac{\rho}{2}\left[e^{-i\beta(x-l)} - e^{i\beta(x-l)}\right], \tag{60.42}$$

$$I = \frac{I_0}{2}\left[e^{-i\beta(x-l)} + e^{i\beta(x-l)}\right]. \tag{60.43}$$

Since the time dependence of the quantities is characterized by the factor $\exp(i\omega t)$, we may conclude that the first terms on the right-hand sides of these formulas describe a wave propagating in the positive direction of the X-axis, while the second terms correspond to the negative direction (i.e. describe the wave reflected at the end of the line). Thus, we can draw the conclusion that the necessity of matching is dictated not only by the fact that in the absence of matching it is impossible to transmit power to the load completely. If signals are transmitted in the form of pulses, consecutive reflections from the load and then from the input distort the signal to such an extent that it becomes difficult to deal with it.

The energy transmitted with the help of electric current flows in the space surrounding the conductors. The conductors play the role of guides along which electromagnetic energy flows. Joule's heat is liberated in a conductor at the expense of electromagnetic energy entering the conductor through its surface from the surrounding space.

Give the definition of the characteristic impedance and the propagation constant of a transmission line.
Describe the physical processes leading to the reflection of energy from the load. Under which condition is the reflection absent and the entire energy transmitted through the line absorbed by the load?

Sec. 61. Electromagnetic Wave Radiation

The solution of the problem on radiation of a linear oscillator is given. The obtained equation is generalized for the case of an arbitrarily accelerated nonrelativistic electron. Radiation reaction is considered-

Equation for vector potential. The magnetic induction and electric field strength of varying fields are expressed in terms of the vector and scalar potentials through formulas (46.8) and (46.12). For this purpose, we must have appropriate equations.

www.EngineeringBooksPDF.com

We proceed from the first of Maxwell's equations (58.1a) which can be conveniently written in the form

$$\operatorname{curl} \mathbf{B} = \mu \mathbf{j} + \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t}, \tag{61.1}$$

where for the sake of simplicity it is assumed that μ and ε are independent of coordinates. Substituting (46.8) and (46.12) into this equation, we obtain

$$\operatorname{curl} \operatorname{curl} \mathbf{A} = \mu \mathbf{j} + \mu\varepsilon \frac{\partial}{\partial t} \left(-\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} \right). \tag{61.2}$$

Considering that curl curl $\mathbf{A}$ = grad div $\mathbf{A} - \nabla^2 \mathbf{A}$, we transform (61.2) as follows:

$$\nabla^2 \mathbf{A} - \mu\varepsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{j} + \operatorname{grad} \left(\operatorname{div} \mathbf{A} + \mu\varepsilon \frac{\partial \varphi}{\partial t} \right). \tag{61.3}$$

Using the ambiguity of potentials defined to within the gauge transformation (46.13), we can impose a certain condition on them. In order to simplify Eq. (61.3) as much as possible, this condition is chosen in the form

$$\boxed{\operatorname{div} \mathbf{A} + \mu\varepsilon \frac{\partial \varphi}{dt} = 0,} \tag{61.4}$$

which is called the **Lorentz equation.** As a result [see 61.3)], we obtain

$$\nabla^2 \mathbf{A} - \varepsilon\mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{j}. \tag{61.5}$$

This relation is called the **D'Alembert equation.**

The choice of gauge function χ. When the Lorentz equation (61.4) is imposed on potentials, the function χ with the help of which gauge transformation (46.13) of potentials is realized cannot be chosen arbitrarily: it is necessary that the Lorentz equation (61.4) should remain unchanged under gauge transformations. We have

$$\operatorname{div} \mathbf{A}' + \mu\varepsilon \frac{\partial \varphi'}{\partial t} = \operatorname{div} (\mathbf{A} + \operatorname{grad} \chi) + \mu\varepsilon \frac{\partial}{\partial t} (\varphi - \partial\chi/\partial t)$$

$$= \operatorname{div} \mathbf{A} + \mu\varepsilon \left[\frac{\partial \varphi}{\partial t} + \left(\nabla^2 \chi - \mu\varepsilon \frac{\partial^2 \chi}{\partial t^2} \right) \right].$$

Thus, *the Lorentz equation is invariant only under gauge transformations with the function χ satisfying the following equation*:

$$\nabla^2 \chi - \mu\varepsilon \frac{\partial^2 \chi}{\partial t^2} = 0. \tag{61.6}$$

An equation of this type is called the **wave equation, or the homogeneous D'Alembert's equation.**

Equation for scalar potential. Substituting (46.12) into Maxwell's equation (58.1a, IV), we obtain

$$\operatorname{div} \left(-\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} \right) = \frac{\rho}{\varepsilon}. \tag{61.7}$$

www.EngineeringBooksPDF.com

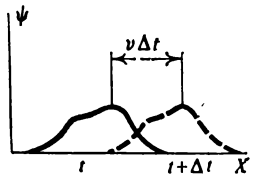


Fig. 251. Time variation of the solution of one-dimensional wave equation

Eliminating div A from this equation with the help of (61.4), we finally obtain the following equation for the scalar potential:

$$\nabla^2\varphi - \varepsilon\mu\frac{\partial^2\varphi}{\partial t^2} = -\frac{\rho}{\varepsilon}. \qquad (61.8)$$

Thus, we have obtained the same equation for the Cartesian projections of vector potential (61.5) and for the scalar potential:

$$\nabla^2\Phi - \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2} = -f(\mathbf{r}, t), \qquad (61.9)$$

where we can substitute A_x, A_y, A_z, φ for Φ and μj_x, μj_y, μj_z, ρ/ε for f, respectively. Let us elucidate the meaning of the expression $\varepsilon\mu = 1/v^2$.

Solution of the wave equation. Let us first consider solutions of Eq. (61.9) for $f = 0$, i.e. the solution of the corresponding homogeneous equation. We take the one-dimensional case $\Phi = \Phi(x)$. Equation (61.9) has the form

$$\frac{\partial^2\Phi}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2} = 0. \qquad (61.10)$$

It can be verified directly that any function Φ of the argument $t - x/v$ or $t + x/v$ is a solution of Eq. (61.10). Let us prove this, for example, for the function $\Phi(t - x/v)$:

$$\frac{\partial\Phi}{\partial t} = \Phi', \quad \frac{\partial^2\Phi}{\partial t^2} = \Phi'',$$
$$\frac{\partial\Phi}{\partial x} = -\frac{1}{v}\Phi', \qquad (61.11)$$
$$\frac{\partial^2\Phi}{\partial x^2} = \frac{1}{v^2}\Phi'',$$

where Φ' is the derivative with respect to the argument of the function. It follows from (61.11) that an arbitrary function $\Phi(t - x/v)$ indeed satisfies Eq. (61.10). Similarly, we can prove that the function $\Phi(t + x/v)$ also satisfies this equation.

These solutions have a very simple meaning. *The function $\Phi(t - x/v)$ is a wave moving in the direction of positive values of the X-axis at a velocity v.* Indeed,

$$t - x/v = t + \Delta t - (x + \Delta x)/v \qquad (61.12)$$

for $\Delta x/\Delta t = v$. This means that if a function $\Psi(t - x/v)$ is represented by a certain curve at the instant t (Fig. 251), at the moment $t + \Delta t$ it will be represented by the same curve but shifted in the direction of positive values of the X-axis by $v\,\Delta t$, i.e. it is a wave propagating in the direction of positive values of the X-axis at a velocity v. This is the reason behind the introduction of the notation $\varepsilon\mu = 1/v^2$.

Similarly it can be shown that *the function $\Phi(t + x/v)$ is a wave propagating at a velocity v in the direction of negative values of the X-axis.*

www.EngineeringBooksPDF.com

Let us consider the solution of the wave equation in spherically symmetric case, i.e. assuming that in (61.9) $f = 0$ and $\Phi = \Phi(r)$, where r is the distance from the origin to the point under consideration. In this case, Φ is independent of the angles, and the Laplace operator has the form

$$\nabla^2\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) = \frac{\partial^2\Phi}{\partial r^2} + \frac{2}{r}\frac{\partial\Phi}{\partial r} = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\Phi). \qquad (61.13)$$

Hence, the wave equation for Φ can be written in the form

$$\frac{\partial^2}{\partial r^2}(r\Phi) - \frac{1}{v^2}\frac{\partial^2(r\Phi)}{\partial t^2} = 0. \qquad (61.14)$$

As in the previous case, the solutions of this equation in $r\Phi$ are arbitrary functions of the arguments $t - r/v$ and $t + r/v$, i.e. the general expression for Φ is as follows:

$$\Phi(r, t) = \frac{\Psi_1(t - r/v)}{r} + \frac{\Psi_2(t + r/v)}{r}. \qquad (61.15)$$

The function $\Psi_1(t - r/v)$ is a wave propagating in the radial direction from the origin at a velocity v. This wave is divergent. Its shape does not change, while its amplitude decreases in proportion to $1/r$. The function $\Psi_2(t + r/v)$ represents a wave converging to the origin.

Returning to (61.5) and (61.8), we see that the field potentials, and hence the fields themselves, propagate in free space ($\rho = 0$) at a velocity $v = 1/\sqrt{\varepsilon\mu}$. In vacuum $\mu = \mu_0$ and $\varepsilon = \varepsilon_0$, and hence the velocity of propagation of fields is equal to the velocity of light $c = 1/\sqrt{\varepsilon_0\mu_0}$. Thus, *electromagnetic waves and all variations of the electric and magnetic fields propagate in vacuum with the velocity of light.* This means that electromagnetic interactions propagate at the velocity of light. For example, if two point charges are at rest at a distance r from each other, and if one of the charges is displaced relative to its initial position at a certain instant, the other charge will "perceive" this displacement only after a time $\tau = r/c$.

Retarded and advanced potentials. Considering the properties of solutions of the wave equation, it should be expected that the solution of Eqs. (61.5) and (61.8) for potentials of varying fields differs from the solutions of Eqs. (37.11a) and (14.35) for potentials of constant fields only in that the finite velocity of propagation of electromagnetic interactions should be taken into account in the former case. In other words, *a moving charge and an alternating current element create at each point of the surrounding space the same potential which would be created by the fixed charge and direct current, the only difference being that such a potential is created at each point after a lapse of the delay time, i.e. the time required for the electromagnetic field to propagate from the source to the point of observation.* Therefore, for charges and currents in a certain finite region of space, we obtain the following formulas instead of (37.11a) and (14.35):

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu}{4\pi}\int\frac{\mathbf{j}(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/v)}{|\mathbf{r} - \mathbf{r}'|}dV', \qquad (61.16)$$

$$\varphi(\mathbf{r}, t) = \frac{1}{4\pi\varepsilon}\int\frac{\rho(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/v)}{|\mathbf{r} - \mathbf{r}'|}dV', \qquad (61.17)$$

www.EngineeringBooksPDF.com

where $v = 1/\sqrt{\varepsilon\mu}$ and $|\mathbf{r} - \mathbf{r}'|$ is the distance between the point at which a potential is calculated and the element dV' of integration volume.

At a given instant and at a given point, the potential is determined not by the position and magnitude of charges and currents but by their positions and magnitudes at the previous instants of time which are determined by taking into account the velocity of propagation of electromagnetic field. Suppose, for example, that a certain electric charge is rapidly approaching a certain point. The scalar potential created by the charge at this point is determined not by the distance from the charge to the point at a given instant of time but by the distance at a previous moment of time, i.e. by a larger distance. *When the velocity of charge is close to the velocity of light, the difference in these distances may become significant.*

Here we do not formally verify that formulas (61.16) and (61.17) satisfy Eqs. (61.5) and (61.8). In principle, this can be done in the same way as for solutions (14.35) and (37.11a).

Potentials of the form (61.16) and (61.17) are called **retarded potentials** since they describe the potentials at a later instant t in comparison with the time $t - |\mathbf{r} - \mathbf{r}'|/v$ for charges and currents which created these potentials. Solutions similar to (61.16) and (61.17) are formally also the solutions of Eqs. (61.5) and (61.8) if we replace the time argument $t - |\mathbf{r} - \mathbf{r}'|/v$ by $t + |\mathbf{r} - \mathbf{r}'|/v$, which corresponds to two possible signs of the arguments in solutions (61.15) of the wave equation. The solution with the "plus" sign of the argument has no clear physical meaning since it formally corresponds to the situation when a potential is created first and then charges and currents corresponding to it appear, i.e. the potential leads charges and currents. For this reason, it is called the **advanced potential**. *Advanced potentials are used along with retarded potentials for solving boundary-value problems.* This can be explained as follows. Suppose that we have to find an electromagnetic field satisfying certain boundary conditions. At the points inside the volume, the field must obviously be such that when it reaches the boundary at a later moment, it has the values prescribed by the boundary conditions. Clearly, while solving such problems one should be guided not only by past events but also consider what is going to happen in future, i.e. advanced potentials should be used. *This, however, by no means signifies the violation of the causality principle* as can be directly seen from the example considered above. From the point of view of physics, it simply describes the events that occurred in the past in order to make the present as it is, taking into account the known laws of evolution.

Hertzian oscillator. It is an electrical dipole whose moment varies with time. A system of two metallic spheres (Fig. 252) connected by a conductor may serve as a real prototype of a Hertzian oscillator. If the spheres are supplied with equal and opposite charges and the system is left alone, an oscillatory process of charge exchange will occur in the system. The current oscillations will be damped. If the resistance of wires is small and the radiation losses per period are low, damping can be ignored for a sufficiently large number of periods. Then at distances considerably larger than l, the system can be treated as a dipole whose moment varies with time. Such an oscillator was used by Hertz who

www.EngineeringBooksPDF.com

was the first to obtain electromagnetic waves experimentally. This explains the term Hertzian oscillator.

The scalar potential of a dipole with a time-dependent moment. The potential of a dipole is defined by formula (61.17) which can be written in the following convenient form:

$$\varphi(\mathbf{r}, t) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho(\xi, t - r'/c)}{r'} \, dV_\xi, \tag{61.18}$$

where it is assumed that the dipole is in vacuum ($\varepsilon = \varepsilon_0$, $\mu = \mu_0$). While calculating the potential by this formula, it is expedient to place the origin in the charge distribution region. The position of the origin within the limits of this

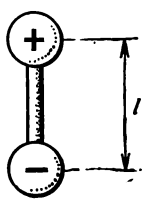


Fig. 252. Oscillator model

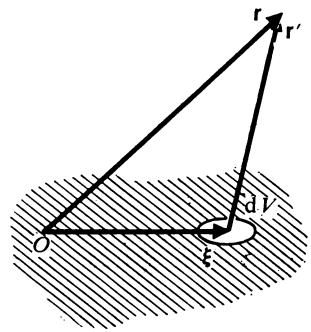


Fig. 253. To the computation of potential of a dipole

region is insignificant since the dipole size can be made as small as desired in comparison with the distance to the points at which the field is analyzed. The position of a point at which the field potential is being calculated is characterized by the radius vector $\mathbf{r}$; ξ is the radius vector of the volume element dV_ξ, and r' is the distance between the volume element dV_ξ and the point of observation (Fig. 253).

Let us consider the potential at large distances from the dipole ($\xi/r \ll 1$). Taking into account that

$$\mathbf{r}' = \mathbf{r} - \xi, \quad r' = \sqrt{r^2 - 2\mathbf{r}\cdot\xi + \xi^2}, \tag{61.19}$$

we can expand the expression for r' into a series in ξ/r and confine ourselves to the linear term of the expansion:

$$r' = r\left(1 - 2\frac{\mathbf{r}\cdot\xi}{r^2} - \frac{\xi^2}{r^2}\right)^{1/2} = r - \frac{\mathbf{r}\cdot\xi}{r} + \ldots \tag{61.20}$$

Using this formula, we expand the integrand in (61.18) into a Taylor series at the point r:

$$\frac{\rho(\xi, t - r'/c)}{r'} = \frac{\rho(\xi, t - r/c)}{r} - \frac{\mathbf{r}\cdot\xi}{r}\frac{\partial}{\partial r}\left[\frac{\rho(\xi, t - r/c)}{r}\right] + \ldots \tag{61.21}$$

www.EngineeringBooksPDF.com

Substituting this expression into (61.18), we obtain

$$\varphi = \frac{1}{4\pi\varepsilon_0}\frac{1}{r}\int \rho\, dV_\xi - \frac{1}{4\pi\varepsilon_0}\frac{\mathbf{r}}{r}\cdot\frac{\partial}{\partial r}\int \xi\rho\, dV_\xi, \tag{61.22}$$

where we took into account that r remains constant upon integration. The first integral on the right-hand side of (61.22) is equal to zero due to electrical neutrality of the system. The second integral is the dipole moment [see (17.2)]:

$$\int \xi\rho\,(t - r/c)\, dV_\xi = \mathbf{p}\,(t - r/c). \tag{61.23}$$

Hence we finally obtain the formula for a dipole with a time-dependent moment:

$$\varphi(\mathbf{r}, t) = -\frac{1}{4\pi\varepsilon_0}\frac{\mathbf{r}}{r}\cdot\frac{\partial}{\partial r}\left[\frac{\mathbf{p}(t-r/c)}{r}\right]. \tag{61.24}$$

Using the expression for divergence in spherical coordinates, we can represent this formula as follows:

$$\varphi(\mathbf{r}, t) = -\frac{1}{4\pi\varepsilon_0}\operatorname{div}\frac{\mathbf{p}(t-r/c)}{r}. \tag{61.25}$$

Vector potential. It is calculated by expanding the integrand in (61.16) into a series of the form (61.21):

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi}\frac{\partial}{\partial t}\left[\frac{\mathbf{p}(t-r/c)}{r}\right]. \tag{61.26}$$

Electric and magnetic fields. In order to simplify the formulas for further analysis, we introduce the following notation:

$$\mathbf{\Pi} = \frac{\mathbf{p}(t-r/c)}{r} = \mathbf{p}_0\Phi(t, r), \tag{61.27}$$

where $\mathbf{p}_0$ is a constant vector characterizing the direction of dipole oscillations. Proceeding from (61.25) and (61.26), we obtain

$$\mathbf{B} = \operatorname{curl}\mathbf{A} = \frac{\mu_0}{4\pi}\operatorname{curl}\frac{\partial\mathbf{\Pi}}{\partial t} = \frac{\mu_0}{4\pi}\frac{\partial}{\partial t}\operatorname{curl}\mathbf{\Pi}, \tag{61.28}$$

$$\mathbf{E} = -\operatorname{grad}\varphi - \frac{\partial\mathbf{A}}{\partial t} = \frac{1}{4\pi\varepsilon_0}\operatorname{grad}\operatorname{div}\mathbf{\Pi} - \frac{\mu_0}{4\pi}\frac{\partial^2\mathbf{\Pi}}{\partial t^2}$$

$$= \frac{1}{4\pi\varepsilon_0}\left(\operatorname{grad}\operatorname{div}\mathbf{\Pi} - \frac{1}{c^2}\frac{\partial^2\mathbf{\Pi}}{\partial t^2}\right) = \frac{1}{4\pi\varepsilon_0}\operatorname{curl}\operatorname{curl}\mathbf{\Pi}, \tag{61.29}$$

where we took into account that $\mu_0\varepsilon_0 = 1/c^2$ and used formula (A.10). Vector $\mathbf{\Pi}$ satisfies the wave equation

$$\nabla^2\mathbf{\Pi} - \frac{1}{c^2}\frac{\partial^2\mathbf{\Pi}}{\partial t^2} = 0. \tag{61.30}$$

The value of curl $\mathbf{\Pi}$ is calculated by formula (A.16):

$$\operatorname{curl}\mathbf{\Pi}\operatorname{curl}\mathbf{p}_0\Phi = \operatorname{grad}\Phi\times\mathbf{p}_0 = \frac{1}{r}\frac{\partial\Phi}{\partial r}\mathbf{r}\times\mathbf{p}_0. \tag{61.31}$$

www.EngineeringBooksPDF.com

It is more convenient to carry out subsequent calculations in a spherical system of coordinates. Let us direct the polar axis Z along the vector $\mathbf{p}_0$, placing the origin at the centre of the dipole. We denote the polar and azimuthal angles by θ and α respectively (Fig. 254). Obviously,

$$(\mathbf{r} \times \mathbf{p}_0)_r = (\mathbf{r} \times \mathbf{p}_0)_\theta = 0, \quad (\mathbf{r} \times \mathbf{p}_0)_\alpha = -rp_0 \sin\theta, \tag{61.32}$$

and hence

$$\operatorname{curl}_r \Pi = \operatorname{curl}_\theta \Pi = 0, \ \operatorname{curl}_\alpha \Pi = -\sin\theta \frac{\partial \Pi}{\partial t}. \tag{61.33}$$

Using (61.28), we obtain

$$B_r = B_\theta = 0, \ B_\alpha = \frac{\mu_0}{4\pi} \frac{\partial}{\partial t} \operatorname{curl}_\alpha \Pi = -\frac{\mu_0}{4\pi} \sin\theta \frac{\partial^2 \Pi}{\partial t \, \partial r}. \tag{61.34}$$

The projections of vector $\mathbf{E}$ are calculated with the help of the formulas for the curl in a spherical coordinate system:

$$\begin{aligned} E_r &= \frac{1}{4\pi\varepsilon_0} \frac{1}{r\sin\theta} \frac{\partial}{\partial\theta} (\sin\theta \operatorname{curl}_\alpha \Pi) = -\frac{1}{2\pi\varepsilon_0} \frac{\cos\theta}{r} \frac{\partial \Pi}{\partial r}, \\ E_\theta &= -\frac{1}{4\pi\varepsilon_0} \frac{1}{r} \frac{\partial}{\partial r} (r \operatorname{curl}_\alpha \Pi) = \frac{1}{4\pi\varepsilon_0} \frac{\sin\theta}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \Pi}{\partial r} \right). \end{aligned} \tag{61.35}$$

Formulas (61.34) and (61.35) show that the *electric field vector lies in meridional planes, while the magnetic induction vector is normal to the meridional plane passing through the corresponding point. The magnetic lines of force coincide with the parallels of the spherical coordinate system under consideration. The electric field and magnetic induction vectors are mutually perpendicular at each point.*

Formulas (61.34) and (61.35) are valid for an arbitrary time dependence of function $\Phi(r, t)$ in (61.27). Assuming that the dipole moment varies in accordance with the harmonic law

$$\mathbf{p} = \mathbf{p}_0 e^{i\omega t}, \tag{61.36}$$

we obtain

$$\Pi = \mathbf{p}_0 \frac{e^{i\omega(t - r/c)}}{r}. \tag{61.37}$$

Differentiating the expressions in formulas (61.34) and (61.35), we obtain the following expressions for the nonzero components of the field vectors:

$$\begin{aligned} B_\alpha &= \frac{\mu_0}{4\pi} i\omega \sin\theta \left(\frac{1}{r} + \frac{i\omega}{c} \right) \Pi, \\ E_r &= \frac{1}{2\pi\varepsilon_0} \cos\theta \left(\frac{1}{r^2} + \frac{i\omega}{cr} \right) \Pi, \\ E_\theta &= \frac{1}{4\pi\varepsilon_0} \sin\theta \left(\frac{1}{r^2} + \frac{i\omega}{cr} - \frac{\omega^2}{c^2} \right) \Pi. \end{aligned} \tag{61.38}$$

In the immediate vicinity of the oscillator, at distances shorter than the wavelength $\lambda = cT = 2\pi c/\omega$, *the field of the oscillator is identical to*

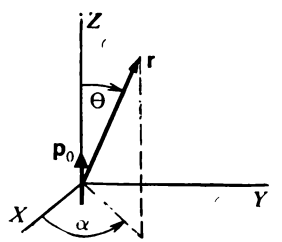


Fig. 254. Choice of a spherical system of coordinates for computing a dipole field

www.EngineeringBooksPDF.com

the field of a stationary dipole and current. At distances considerably exceeding the wavelength, the field of the oscillator differs in principle from the field of a permanent dipole and current. The corresponding region is called the wave zone.

The field of an oscillator in the wave zone. By definition, the distance r from the points of the wave zone satisfies the following inequality:

$$\frac{1}{r} \ll \frac{\omega}{c}. \tag{61.39}$$

Consequently, in formulas (61.38) we can ignore $1/r$ and $1/r^2$ in comparison with ω/c and ω^2/c^2. As a result, we obtain the following expressions for the components of the field vectors:

$$B_\alpha = -\frac{\mu_0}{4\pi}\frac{\omega^2}{c^2}\Pi \sin\theta, \quad B_r = B_\theta = 0; \tag{61.40}$$

$$E_\theta = -\frac{1}{4\pi\varepsilon_0}\frac{\omega^2}{c^2}\Pi \sin\theta, \quad E_r = E_\alpha = 0. \tag{61.41}$$

In these formulas, we can take either the real or the imaginary part of (61.37) for Π. For] example,

$$\Pi = \frac{p_0 \cos\omega(t - r/c)}{r}. \tag{61.42}$$

Consequently, the electric field and magnetic induction of the electromagnetic field in the wave zone of an oscillator can finally be represented in the following form:

$$\begin{aligned} &E_\theta = cB_\alpha = -\frac{1}{4\pi\varepsilon_0}\frac{\omega^2}{c^2}\frac{\sin\theta}{r} p_0 \cos\omega\left(t - \frac{r}{c}\right), \\ &E_r = E_\alpha = 0, \quad B_r = B_0 = 0. \end{aligned} \tag{61.43}$$

These formulas show that in the wave zone the electric field vector and the magnetic induction vector are perpendicular to each other and to the radius vector **r**. The vectors **E**, **B**, and **r** form a right-handed system of vectors at each point. *The electric field strength decreases in inverse proportion to the distance.* The wave described by formulas (61.43) is called a spherical wave. It propagates in the direction of the radius vector. The constant-phase surfaces of this wave are spheres. The (phase) velocity of the wave is equal to the velocity of light. Since $E_\theta = cB_\alpha$, small portions of the surface of a spherical wave can be treated as plane electromagnetic waves.

Power radiated by an oscillator. The flux density of electromagnetic energy is characterized by Poynting's vector (59.7). Consequently, the electromagnetic energy flux P through the surface S of the sphere of radius r, surrounding an

www.EngineeringBooksPDF.com

oscillator, is given by

$$P = \int_S \mathbf{E} \times \mathbf{H} \cdot d\mathbf{S} = \int_S E_\theta H_\alpha \, dS$$

$$= \frac{1}{16\pi^2 \varepsilon_0} \frac{\omega^4 p_0^2}{c^3} \cos^2 \omega \left(t - \frac{r}{c}\right) \int_0^\pi \sin^3 \theta \, d\theta \int_0^{2\pi} d\alpha$$

$$= \frac{1}{6\pi\varepsilon_0} \frac{\omega^4 p_0^2}{c^3} \cos^2 \omega \left(t - \frac{r}{c}\right). \tag{61.44}$$

This quantity is the rate of flow, i.e. the energy emitted by the oscillator in 1 s. The emissive power averaged over a period is

$$\langle P \rangle = \frac{1}{T} \int_0^T P \, dt = \frac{1}{12\pi\varepsilon_0} \frac{\omega^4 p_0^2}{c^3}. \tag{61.45}$$

This formula shows that the *emissive power of an oscillator strongly depends on frequency and is proportional to its fourth power*. This means that in order to increase the emissive power, it is expedient to operate on short waves.

Since Poynting's vector decreases in inverse proportion to the square of the distance, while the surface area of the sphere increases in direct proportion to the distance, the total energy flow crossing the surface of the sphere does not change with distance, and hence the energy is transferred from the oscillator to remote regions of space without losses in the form of electromagnetic waves. The radiant flux density decreases in inverse proportion to the square of the distance. The oscillations of an oscillator are damped due to loss of energy by radiation. In order to obtain undamped oscillations, it is necessary to continuously supply energy to the oscillator. An oscillator is the simplest emitter of electromagnetic waves.

Radiation of a current-carrying loop. Another example of a simple electromagnetic wave radiator is a current loop which is characterized by a magnetic moment $\mathbf{p}_m = I\mathbf{S}$ (Fig. 255). Its radiation is similar to that of a dipole. We shall give here just the result. The magnetic moment of a current loop varies according to the law

$$\mathbf{p}_m = \mathbf{p}_{m_0} \cos \omega t. \tag{61.46}$$

Let us place the origin of a spherical coordinate system at the centre of the loop and direct the Z-axis along the magnetic moment. In Fig. 254, the current is assumed to flow in the plane $z = 0$ and the magnetic moment $\mathbf{p}_m$ of the current is arranged as $\mathbf{p}$. The following formulas describe the radiation field of a current loop:

$$E_\alpha = -cB_\theta = \frac{\mu_0}{4\pi} \frac{\omega^2}{c} \frac{\sin\theta}{r} p_{m_0} \cos \omega \left(t - \frac{r}{c}\right),$$
$$E_r = E_\theta = 0, \quad B_r = B_\alpha = 0 \tag{61.47}$$

www.EngineeringBooksPDF.com

A comparison of formulas (61.47) and (61.43) shows that if the magnetic moment $\mathbf{p}_{m_0}$ of current and the dipole moment $\mathbf{p}_0$ are related through (Fig. 256)

$$\mathbf{p}_{m_0} = c\mathbf{p}_0, \tag{61.48}$$

the electric field strength and the magnetic induction of the dipole radiation are equal in magnitude to the corresponding values of current loop radiation but have different directions. The electric field of a dipole is directed along

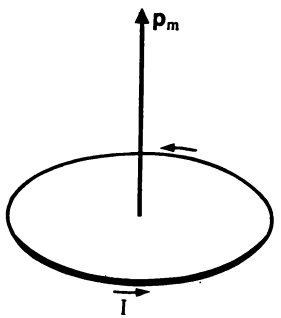


Fig. 255. Current loop

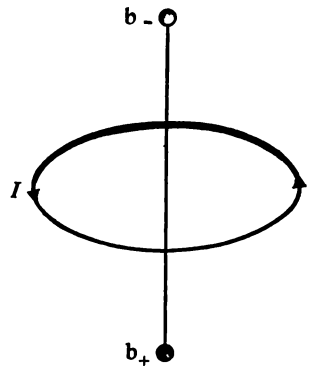


Fig. 256. Relation between the displacement of electric charges creating an electric dipole moment and the current in a loop producing a magnetic moment

meridians, while that of the current loop is directed perpendicularly to the meridional planes along the parallels. The orientation of magnetic induction vectors varies accordingly. It can be seen from (61.47) and (61.43) that there exists the following relation between the field vectors of dipole radiation and current loop radiation:

$$E_\alpha \text{ (loop)} = -cB_\alpha \text{ (dipole)}, \quad cB_\theta \text{ (loop)} = E_\theta \text{ (dipole)}. \tag{61.49}$$

The emissive power of a current loop can be defined by formulas (61.44) and (61.45) where the dipole moment is replaced by the magnetic moment in accordance with (61.48).

An oscillator and a current loop are elementary radiators of electromagnetic waves. Radiation of more complex systems can be reduced to that of elementary radiators with the help of the superposition principle.

Radiation of an electron moving with an acceleration. Let us imagine a positive charge equal in magnitude to the charge of an electron placed at the origin. It is fixed and according to Coulomb's law creates in the surrounding space a constant electric field, whose strength decreases in inverse proportion to the square of the distance. A system of a moving electron and a fixed charge is a dipole whose moment varies with time. The field vectors of the dipole radiation are varying quantities which decrease in inverse proportion to the

www.EngineeringBooksPDF.com

distance. Clearly, the constant electric field of the fixed charge is compensated by the electric field of the electron and has nothing in common with the radiation field, i.e. the radiation field is the field of radiation of the oscillating electron. The positive charge was placed at the origin just mentally, which makes it possible to use the above formulas for the radiation of a dipole with a time-dependent moment.

The dipole moment appearing when an electron deviates from the origin by $z(t)$ is given by

$$\mathbf{p}(t) = -|e|\, z(t)\, \mathbf{i}_z, \tag{61.50}$$

where $\mathbf{i}_z$ is the unit vector along the Z-axis. The minus sign appeared in view of the fact that the additional moment is directed from the negative charge to the positive one. Assuming that

$$z = b \cos \omega t, \tag{61.51}$$

where b is the amplitude of electron oscillation, we obtain the following formula for dipole moment (61.50):

$$\mathbf{p} = -\mathbf{i}_z |e| b \cos \omega t. \tag{61.52}$$

Comparing this formula with the real part of (61.36) for a dipole, we conclude that the moment $\mathbf{p}_0$ in formula (61.36) is connected with the quantities characterizing the motion of the electron through the relation

$$\mathbf{p}_0 = -\mathbf{i}_z |e| b, \quad p_0 = |e| b. \tag{61.53}$$

Formula (61.43) characterizing the radiation field vectors now assumes the form

$$\begin{aligned} &E_\theta = cB_\alpha = -\frac{\omega^2}{4\pi\varepsilon_0 c^2} \frac{\sin\theta}{r} |e| b \cos\omega\left(\tau - \frac{r}{c}\right), \\ &E_\alpha = E_r = 0, \quad B_r = B_\theta = 0, \end{aligned} \tag{61.54}$$

where τ is the time during which the wave arrives at the point of observation on a sphere of radius r. The variable $t = \tau - r/c$ is reserved for the time characterizing the motion of the electron. It follows from formula (61.51) that

$$\ddot{z} = -\omega^2 b \cos \omega t, \tag{61.55}$$

and hence Eq. (61.54) can be written in the form

$$\begin{aligned} E_\theta(r, \tau) = cB_\alpha(r, \tau) &= \frac{|e|}{4\pi\varepsilon_0 c^2} \frac{\sin\theta}{r} \ddot{z}\Big|_{t=\tau-r/c} \\ &= -\frac{|e|}{4\pi\varepsilon_0 c^2} \frac{\sin\theta}{r} \ddot{z}\Big|_{t=\tau-r/c}, \end{aligned} \tag{61.56}$$

where we took into account that the electron charge is negative. Formula (61.44) for the emissive power assumes the form

$$\boxed{P = \frac{1}{6\pi\varepsilon_0} \frac{e^2}{c^3} \ddot{z}^2,} \tag{61.57}$$

www.EngineeringBooksPDF.com

i.e. *the emissive power is proportional to the square of the electron acceleration. A uniformly moving charge does not radiate.*

Formulas (61.56) and (61.57) were obtained for the model of an oscillating electron. They depend, however, only on the electron acceleration at any instant of time. Consequently, the radiation field described by these formulas does not depend on the motion of the electron prior to a given moment and after that. Therefore, they can be applied in all cases and represent the electric field strength and the magnetic induction of the radiation field as well as the emissive power as functions of acceleration for any motion. In this case, however, the electron velocity should be small. Consequently, these formulas are applicable, strictly speaking, for an electron at rest but having an acceleration, which is obvious from the definition of a dipole which occupies an infinitely small spatial region and is at rest in it.

However, the generalization of these formulas to arbitrary velocities does not present any difficulty. For this purpose, we should go over to the coordinate system in which the electron moves at an arbitrary velocity and use the formulas for the transformation of fields and accelerations. As a result, we obtain formulas which are valid for arbitrary velocities and accelerations of a charge. We shall not present these formulas here.

Decelerative force due to radiation. An electron loses its energy due to radiation and is decelerated. In other words, a decelerative force acts on it. Let us find this force. The equation for electron oscillations, taking into account the decelerative force, has the form

$$m\ddot{z} + m\omega^2 z = F, \tag{61.58}$$

where ω is the natural frequency in the absence of a decelerative force due to radiation. Multiplying both sides of this equation by $\dot{z}$, we obtain

$$\frac{d}{dt}\left(\frac{m\dot{z}^2}{2} + \frac{m\omega^2}{2} z^2\right) = F\dot{z}. \tag{61.59}$$

The right-hand side of this equation is the work of the decelerative force due to radiation per unit time. By definition, it is equal to the emissive power [see (61.57)], and hence

$$F\dot{z} = -\frac{1}{6\pi\varepsilon_0}\frac{e^2}{c^3}\ddot{z}^2. \tag{61.60}$$

This equation expresses the law of energy conservation for radiation. It cannot be used in the general form for determining the force F as a function of z and its derivatives. This can be done only approximately, assuming that

(1) radiation, and hence the damping of oscillations, are not very strong so that the motion can be assumed to be practically periodic for several periods;

(2) from the law of energy conservation for average quantities pertaining to a small number of periods, we can conclude that the instantaneous values of corresponding quantities are equal.

We proceed from the obvious equality:

$$\ddot{z}^2 = -\dddot{z}\dot{z} + (\ddot{z}\dot{z})^{\cdot}. \tag{61.61}$$

www.EngineeringBooksPDF.com

By averaging $(\ddot{z}\dot{z})^{\cdot}$ over one period and using the first assumption, we have

$$\langle(\dddot{z}\,\dot{z})^{\cdot}\rangle = \frac{1}{T}\left[(\dddot{z}\,\dot{z})_{t=T} - (\dot{z}\,\dddot{z})_{t=0}\right] = 0. \tag{61.62}$$

Then Eq. (61.60) combined with (61.61) and (61.62) assumes the form

$$\langle\dot{F}z\rangle = \frac{1}{6\pi\varepsilon_0}\frac{e^2}{c^3}\langle\ddddot{z}\,\dot{z}\rangle. \tag{61.63}$$

On the basis of the second assumption, we obtain

$$F = \frac{1}{6\pi\varepsilon_0}\frac{e^2}{c^3}\dddot{z}. \tag{61.64}$$

This formula defines the decelerative force due to radiation. The equation for the electron oscillations, taking into account the decelerative force can be written as follows:

$$m\ddot{z} + m\omega^2 z - [e^2/(6\pi\varepsilon_0 c^3)]\dddot{z} = 0. \tag{61.65}$$

In electrodynamics, the expression for the decelerative force is generalized to an arbitrary motion. In this case too, this force is proportional to the third-order time derivative of the corresponding quantities characterizing the motion of an electron. The obtained equation is a relativistic invariant. For a long time it was assumed that it correctly describes the radiation reaction. However, computer calculations were recently carried out for a number of simple cases of motion and the obtained results proved to be quite meaningless. Therefore, the question of a relativistically invariant classical description of electron motion, on account of the radiation reaction, cannot be taken as solved so far.

The presence of a decelerative force was experimentally confirmed in particle accelerators. As was mentioned above, charged particles in accelerators undergo small harmonic oscillations about an equilibrium orbit, which are called the betatron oscillations (see Sec. 56). Besides, a moving charge radiates intensely. The decelerative force due to radiation causes damping of betatron oscillations.

Sec. 62. Propagation of Electromagnetic Waves in Dielectrics

Basic properties and peculiarities of propagation of electromagnetic waves in dielectrics are considered.

Plane waves. *An electromagnetic wave is called a plane wave if the field vectors of the wave have the same magnitude at all points of any plane perpendicular to the direction of wave propagation.* Naturally these vectors vary from plane to plane. It can be said that constant-phase surfaces in a plane wave are planes perpendicular to the direction of propagation. *A wave is called monochromatic if the time variation of the field vectors of the wave obeys a harmonic law with a certain*

www.EngineeringBooksPDF.com

constant frequency. For example, if a plane electromagnetic wave propagates along the Z-axis, the field vectors of the wave have the form

$$\mathbf{E}(z, t) = \mathbf{E}(z)\, e^{i\omega t}; \quad \mathbf{B}(z, t) = \mathbf{B}(z)\, e^{i\omega t}. \tag{62.1}$$

If the constant-phase surfaces coincide with the surfaces of constant amplitude, the wave is called homogeneous.

Equations for the field vectors of a wave. Unlike Sec. 61, we shall proceed directly from the field vectors rather than from the potentials. Let us consider a homogeneous unbounded medium for which $\varepsilon = \text{const}$ and $\mu = \text{const}$. The electrical conductivity of a dielectric is $\gamma = 0$. Maxwell's equations have the form

$$\operatorname{curl} \mathbf{B} = \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t}, \tag{62.2}$$

$$\operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{62.3}$$

Differentiating both sides of Eq. (62.2) with respect to time and substituting Eq. (62.3) for the derivative $\partial\mathbf{B}/\partial t$, we get

$$-\operatorname{curlcurl} \mathbf{E} = \varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{62.4}$$

Using formula (A.10) and considering that $\operatorname{div} \mathbf{E} = 0$, since there are no free charges, we write the equation for $\mathbf{E}$:

$$\nabla^2 \mathbf{E} - \varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \tag{62.5}$$

Similarly, we can find the equation for $\mathbf{B}$:

$$\nabla^2 \mathbf{B} - \varepsilon\mu \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0. \tag{62.6}$$

Thus, the field vectors satisfy the wave equation in which the velocity of wave propagation is

$$v = 1/\sqrt{\varepsilon\mu} = c/\sqrt{\varepsilon_r \mu_r}. \tag{62.7}$$

Formula (62.7) shows that the *velocity of wave propagation in a dielectric is smaller than in vacuum.*

Field vectors of a wave. Let us direct the Z-axis along the velocity of propagation of an electromagnetic wave. In this case, the field vectors are defined by the formulas of the form (62.1). Substituting into (62.5) the expression for $\mathbf{E}$ [see (62.1)] and cancelling $\exp(i\omega t)$ from both sides of the equation after differentiation, we write the following equation for $\mathbf{E}(z)$:

$$d^2\mathbf{E}(z)/dt^2 + k^2\mathbf{E}(z) = 0, \tag{62.8}$$

where $k = \omega\sqrt{\varepsilon\mu}$. The general solution of this equation is given by

$$\mathbf{E}(z) = \mathbf{E}_{01} e^{-ikz} + \mathbf{E}_{02} e^{ikz}, \tag{62.9}$$

where $\mathbf{E}_{01}$ and $\mathbf{E}_{02}$ are constants. Substituting (62.9) into (62.1), we obtain

$$\mathbf{E}(z, t) = \mathbf{E}_{01} e^{i(\omega t - kz)} + \mathbf{E}_{02} e^{i(\omega t + kz)}. \tag{62.10}$$

www.EngineeringBooksPDF.com

The first term on the right-hand side of this equation describes a wave propagating in the direction of positive values of the Z-axis, while the second term corresponds to a wave propagating in the negative direction [see (61.12)].

The solution for $\mathbf{B}$ is found in a similar way. Suppose that a wave propagates in the positive direction of the Z-axis. Then

$$\mathbf{E}(z, t) = \mathbf{E}_0 e^{i(\omega t - kz)}; \quad \mathbf{B}(z, t) = \mathbf{B}_0 e^{i(\omega t - kz)}. \quad (62.11)$$

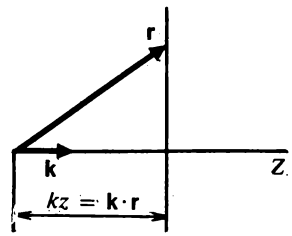


Fig. 257. Constant phase surface of a plane wave

This wave is plane, monochromatic and homogeneous.

Phase velocity. Formulas (62.11) indicate that plane waves propagate in a homogeneous dielectric without a change in their amplitude, i.e. without absorption. The velocity of motion of an equiphase surface is called the **phase velocity.** It can be found by differentiating the condition of constant phase

$$\omega t - kz = \text{const}, \quad (62.12)$$

with respect to time. This gives

$$\omega - k\frac{dz}{dt} = 0, \quad (62.13)$$

whence

$$v = \frac{dz}{dt} = \frac{\omega}{k} = \frac{1}{\sqrt{\varepsilon\mu}} = \frac{c}{\sqrt{\varepsilon_r \mu_r}}. \quad (62.14)$$

Formulas (62.11) are written for a special choice of the coordinate system, when the Z-axis coincides with the direction of wave propagation. This restriction can be eliminated by introducing the wave vector $\mathbf{k}$ directed along the vector of velocity of the wave. The magnitude of this vector is defined by (61.8). By definition of a plane wave propagating in the direction of the vector $\mathbf{k}$, vectors $\mathbf{E}$ and $\mathbf{B}$ are the same at any point of a plane perpendicular to this direction (in this case, to the Z-axis). Let $\mathbf{r}$ be the radius vector of a point on such an equiphase plane. Obviously, $\mathbf{k}\cdot\mathbf{r} = kz$ (Fig. 257), and we can write the following equations instead of (62.11):

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}; \quad \mathbf{B}(\mathbf{r}, t) = \mathbf{B}_0 e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}. \quad (62.15a)$$

Wavelength. By definition, it is the distance over which a point on an equiphase surface moves during a period of oscillations:

$$\lambda = vT = \omega T/k = 2\pi/k, \quad (62.15b)$$

where

$$k = 2\pi/\lambda \quad (62.15c)$$

is the wave number.

Properties of waves. In order to investigate the properties of plane waves, we substitute expressions (62.15a) into (62.2) and (62.3). For the sake of simplicity, it is expedient to use in calculations the symbolic representation of vector

www.EngineeringBooksPDF.com

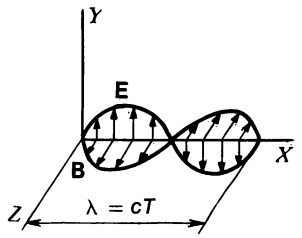


Fig. 258. Plane harmonic electromagnetic wave

operations. The initial operation is defined by operator nabla:

$$\nabla = \mathbf{i}_x \frac{\partial}{\partial x} + \mathbf{i}_y \frac{\partial}{\partial y} + \mathbf{i}_z \frac{\partial}{\partial z}, \qquad (62.16)$$

where $\mathbf{i}_x$, $\mathbf{i}_y$, $\mathbf{i}_z$ are unit vectors in the directions of coordinate axes.

It can be easily shown that the basic operations of vector calculus can be represented with the help of this operator as follows:

$$\text{grad } \varphi = \nabla\varphi, \quad \text{div } \mathbf{A} = \nabla\cdot\mathbf{A}, \quad \text{curl} = \nabla \times \mathbf{A}, \qquad (62.17)$$

where $\nabla\cdot\mathbf{A}$ and $\nabla \times \mathbf{A}$ are the scalar and vector products of operator ∇ by vector $\mathbf{A}$. Note that

$$\nabla e^{-i\mathbf{k}\cdot\mathbf{r}} = -i\mathbf{k}e^{-i\mathbf{k}\cdot\mathbf{r}}. \qquad (62.18)$$

Using Maxwell's equations and expressions (63.15a), we can investigate the properties of plane waves. Maxwell's equation div $\mathbf{E} = 0$ gives

$$\text{div } \mathbf{E} = \nabla\cdot\mathbf{E} = -i\mathbf{k}\cdot\mathbf{E} = 0. \qquad (62.19)$$

This means that the electric field vector $\mathbf{E}$ of the wave is perpendicular to $\mathbf{k}$, i.e. perpendicular to the direction of its propagation. Similarly, Maxwell's equation

$$\text{div } \mathbf{B} = \nabla\cdot\mathbf{B} = -i\mathbf{k}\cdot\mathbf{B} = 0 \qquad (62.20)$$

shows that vector $\mathbf{B}$ is also perpendicular to the direction of wave propagation. Substituting (62.15a) into (62.2) and (62.3), we obtain

$$-\mathbf{k} \times \mathbf{B} = \varepsilon\mu\omega\mathbf{E}, \qquad (62.21)$$

$$\mathbf{k} \times \mathbf{E} = \omega\mathbf{B}. \qquad (62.22)$$

Let $\mathbf{n}$ be a unit vector in the direction of wave propagation. Then, using (62.8), we can write

$$\mathbf{k} = \mathbf{n}\omega\sqrt{\varepsilon\mu} = \mathbf{n}\omega/v. \qquad (62.23)$$

Consequently [see (62.22)],

$$\mathbf{n} \times \mathbf{E} = v\mathbf{B}. \qquad (62.24)$$

It was shown by using (62.19) and (62.20) that vectors $\mathbf{E}$ and $\mathbf{B}$ are perpendicular to $\mathbf{n}$. Formulas (62.21), (62.22) and (62.24) indicate that these vectors are also mutually perpendicular. Writing Eq. (62.24) for the magnitudes of the quantities, we obtain

$$E = vB. \qquad (62.25)$$

We may conclude from relation (62.24) that in a homogeneous dielectric vectors $\mathbf{E}$ and $\mathbf{B}$ vary in phase. All formulas of this section are valid for vacuum as well if we put $\varepsilon = \varepsilon_0$, $\mu = \mu_0$ and $v = c$ (velocity of light). The spatial variation of the field vectors in a plane wave is shown in Fig. 258.

www.EngineeringBooksPDF.com

Energy flux density. This quantity is defined by Poynting's vector whose magnitude for a plane wave is given by

$$|\mathbf{S}| = |\mathbf{E}\times\mathbf{H}| = |\mathbf{E}|\,|\mathbf{H}| = \frac{1}{\sqrt{\varepsilon\mu}}\,\frac{1}{2}\left(\varepsilon E^2 + \frac{1}{\mu^2}B^2\right)$$

$$= \frac{1}{\sqrt{\varepsilon\mu}}\,\frac{1}{2}\,(\mathbf{E}\cdot\mathbf{D} + \mathbf{B}\cdot\mathbf{H}), \tag{62.26a}$$

where $1/\sqrt{\varepsilon\mu} = v$ is the velocity of wave propagation and

$$w = \frac{1}{2}\,(\mathbf{E}\cdot\mathbf{D} + \mathbf{B}\cdot\mathbf{H}) \tag{62.26b}$$

is the volume energy density in it. The expression for the energy flux can be written in the form

$$\mathbf{S} = w\mathbf{v}. \tag{62.27}$$

This means that the *rate of energy transfer by a plane wave in a homogeneous dielectric is equal to the phase velocity of the wave.*

Electromagnetic waves are emitted only by alternating currents and electric charges moving with acceleration. Direct currents and charges that are moving uniformly and rectilinearly do not emit waves.

Describe the physical processes which make the existence of electromagnetic waves possible.
What is the structure of a plane wave? At which velocity does it propagate in a vacuum?

Sec. 63. Propagation of Electromagnetic Waves in Conducting Media

Basic properties and peculiarities of the propagation of electromagnetic waves in conducting media are considered.

Complex permittivity. We consider a homogeneous medium: μ = const, ε = const, and γ = const ($\gamma \neq 0$, i.e. we have a conducting medium). In this case, Maxwell's equations have the form

$$\nabla\times\mathbf{B} = \mu\mathbf{j} + \mu\varepsilon\,\frac{\partial\mathbf{E}}{\partial t} = \mu\gamma\mathbf{E} + \mu\varepsilon\,\frac{\partial\mathbf{E}}{\partial t}, \tag{63.1}$$

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}, \tag{63.2}$$

where we used the symbolic notation for vector operations and took into account that $\mathbf{j} = \gamma\mathbf{E}$. Substituting into these equations expressions (62.15a) for the

www.EngineeringBooksPDF.com

field vectors, we obtain

$$-\mathbf{k}_\omega \times \mathbf{B} = \omega\mu\,[\varepsilon + \gamma/(i\omega)]\,\mathbf{E}, \tag{63.3}$$

$$\mathbf{k}_\omega \times \mathbf{E} = \omega\mathbf{B}, \tag{63.4}$$

where for $\mathbf{k}$ from (62.15a) we use the notation $\mathbf{k}_\omega = \mathbf{k}^{(0)}k_\omega$, $\mathbf{k}^{(0)}$ being the unit vector.

For $\gamma = 0$, Eq. (63.3) coincides with Eq. (62.21) for dielectrics. Equation (63.4) is the same as the corresponding equation for dielectrics. Thus, in mathematical sense, a conducting medium differs from a dielectric only in that the corresponding expression contains the complex permittivity

$$\varepsilon_\omega = \varepsilon + \gamma/(i\omega) = \varepsilon - i\gamma/\omega \tag{63.5}$$

instead of permittivity ε.

All subsequent calculations are identical to those for dielectrics where ε_ω is used instead of ε. Thus, instead of the real wave number, the complex-valued quantity k_ω appears, such that

$$k_\omega^2 = \omega^2\varepsilon_\omega\mu = \omega^2\varepsilon\mu - i\omega\gamma\mu. \tag{63.6}$$

Representing k_ω in the form of a complex number

$$k_\omega = k - is, \tag{63.7}$$

we can write Eq. (63.6) in the form

$$k^2 - 2iks - s^2 = \omega^2\varepsilon\mu - i\omega\gamma\mu. \tag{63.8}$$

Equating the real and imaginary parts of (63.8), we get

$$k^2 - s^2 = \omega^2\varepsilon\mu \equiv a, \tag{63.9}$$

$$2ks = \omega\gamma\mu \equiv b. \tag{63.10}$$

The solution of this algebraic system of equations has the form

$$k^2 = \frac{a}{2}\left(\sqrt{1 + \frac{b^2}{a^2}} + 1\right) = \frac{\omega^2\varepsilon\mu}{2}\left(\sqrt{1 + \left(\frac{\gamma}{\varepsilon\omega}\right)^2} + 1\right), \tag{63.11}$$

$$s^2 = \frac{a}{2}\left(\sqrt{1 + \frac{b^2}{a^2}} - 1\right) = \frac{\omega^2\varepsilon\mu}{2}\left(\sqrt{1 + \left(\frac{\gamma}{\varepsilon\mu}\right)^2} - 1\right). \tag{63.12}$$

Penetration depth. Let us investigate the amplitude of a plane wave propagating in the direction of positive values of the Z-axis:

$$E = E_0 e^{i(\omega t - k_\omega z)} = E_0 e^{-sz} e^{i(\omega t - kz)}. \tag{63.13}$$

It can be seen that the amplitude of the wave decreases during its propagation, i.e. the electromagnetic wave propagating in a conducting medium is damped. Along the path

$$\Delta = 1/s \tag{63.14}$$

www.EngineeringBooksPDF.com

the amplitude of the electric field of the wave decreases to 1/e of its initial value. The quantity Δ is called the **penetration depth** of a plane wave into a conducting medium.

Let us estimate the penetration depth for waves of different wavelengths. The wavelength of visible light is

$$\lambda = (0.4 \div 0.75)\, 10^{-6} \text{ m}, \tag{63.15}$$

which corresponds to the frequency ω of the order of $5 \times 10^{15}\ \text{s}^{-1}$. The electric conductivity of metals is of the order of $10^7\ \Omega^{-1} \cdot \text{m}^{-1}$ and the value of ε can be assumed to be equal to ε_0. Thus,

$$\gamma/(\varepsilon\omega) \simeq 2 \cdot 10^2 \gg 1. \tag{63.16}$$

For waves longer than light waves, this inequality becomes even more stringent. Consequently we may ignore unity in formula (63.12) in comparison with $\gamma/(\varepsilon\omega)$ and write the expression for s in the form

$$s = \sqrt{\omega\gamma\mu/2}. \tag{63.17}$$

Consequently, the penetration depth is given by

$$\Delta = 1/s = \sqrt{2/(\omega\gamma\mu)}. \tag{63.18}$$

Since the wavelength λ is connected with frequency ω through the relation $\omega = 2\pi/(\lambda\sqrt{\varepsilon\mu})$, formula (63.18) can be written in the form

$$\Delta = \sqrt{\frac{\lambda}{\pi\gamma}} \sqrt[4]{\frac{\varepsilon}{\mu}}, \tag{63.19}$$

where $\sqrt{\mu/\varepsilon}$ has the dimensions of resistance and is the characteristic resistance of the medium. For vacuum, it is equal to

$$\sqrt{\mu_0/\varepsilon_0} = 377\ \Omega. \tag{63.20}$$

Let us consider, for instance, copper for which $\gamma = 5 \times 10^7\ \Omega^{-1} \cdot \text{m}^{-1}$, $\mu \simeq \mu_0$ and $\varepsilon \simeq \varepsilon_0$. For $\lambda = 1$ m, the penetration depth is $\Delta \simeq 4 \times 10^{-6}$ m. This means that essentially we cannot speak of any penetration of the wave into the conducting medium since absorption takes place only in a very thin surface layer. This conclusion remains valid for very short waves. For example, for wavelengths of the order of that of light waves ($\lambda \simeq 10^{-6}$ m), the penetration depth Δ is about 4×10^{-9} m.

Physical grounds for absorption. The physical reason behind such a rapid attenuation of electromagnetic waves in a conducting medium is the conversion of the electromagnetic energy of the wave into Joule's heat: *the electric field of the wave induces in a conducting medium currents which, in accordance with Joule's law, heat the material of the medium.*

Interpretation of the skin effect. Now, we can present the interpretation of the skin effect. Formula (53.19) for the thickness of the skin depth coincides with formula (63.18) for the penetration depth of an electromagnetic wave into a conductor. This coincidence has deep physical grounds.

www.EngineeringBooksPDF.com

The energy transported by a current propagates in the space surrounding the conductors in the form of electromagnetic energy. This energy partially penetrates the conductor surface to sustain the motion of electrons. In the conductor it is converted into the kinetic energy of electrons which, in turn, is transformed into Joule's heat. For this reason, the current is maintained in the parts of the conductor which receive electromagnetic energy from the surrounding space. Since this energy can penetrate the conductor only to the depth Δ [see (63.18)], current may exist near the surface of the conductor only within the limits of this depth, i.e. Δ is the thickness of the skin layer.

Phase velocity and the wavelength in a conducting medium. Taking into account (63.13) and (63.11), formula (62.14) can be written in the form

$$v=\frac{\omega}{k}=\frac{1}{\sqrt{\mu\varepsilon}}\left\{\frac{2}{\{1+[\gamma/(\omega\varepsilon)]^2\}^{1/2}+1}\right\}^{1/2}. \tag{63.21}$$

This velocity is lower than the velocity of waves in a nonconducting medium with the same values of μ and ε, i.e. the conductivity of the medium reduces the phase velocity in it. The wavelength in a conducting medium is given by

$$\lambda=\frac{2\pi}{k}=\frac{2\pi}{\omega\sqrt{\mu\varepsilon}}\left\{\frac{2}{\{1+[\gamma/(\omega\varepsilon)]^2\}^{1/2}+1}\right\}^{1/2}, \tag{63.22}$$

i.e. is smaller in comparison with the wavelength in a nonconducting medium with the same values of μ and ε.

Formula (63.22) shows that the phase velocity in a conducting medium depends on frequency, i.e. dispersion takes place in this case. Therefore, a conducting medium is always dispersive. A most typical feature of propagation of signals in a dispersive media is the variation of their form during propagation.

Relation between the phases of field vectors. A complex quantity k_ω in (63.7) can be represented conveniently in the exponential form:

$$k_\omega=|k_\omega|\,e^{i\varphi}. \tag{63.23}$$

Formula (63.4) can be written as follows:

$$\mathbf{B}=\frac{|k_\omega|}{\omega}e^{i\varphi}\mathbf{k}^{(0)}\times\mathbf{E}, \tag{63.24}$$

where $\mathbf{k}^{(0)}$ is the unit vector in the direction of the wave propagation (in our case, the direction of the Z-axis). Vectors $\mathbf{E}$ and $\mathbf{B}$ are perpendicular to this axis.

Suppose that in accordance with formula (63.13) the electric field vector of the wave is given by

$$\mathbf{E}=\mathbf{E}_0e^{-sz}e^{i(\omega t-kz)}, \tag{63.25}$$

where, without loss of generality, we may assume that vector $\mathbf{E}_0$ is real since the choice of the reference point for time t is always arbitrary. Substituting (63.25) into (63.24), we obtain

$$\mathbf{B}=\frac{|k_\omega|}{\omega}\mathbf{k}^{(0)}\times\mathbf{E}_0e^{-sz}e^{i(\omega t-kz+\varphi)}. \tag{63.26}$$

www.EngineeringBooksPDF.com

Having determined the real parts of expressions (63.25) and (63.26), we can find the formulas for actual oscillations of the field vectors in a plane wave propagating in a conducting medium:

$$\mathbf{E} = \mathbf{E}_0 e^{-sz} \cos(\omega t - kz),$$
$$\mathbf{B} = \frac{|k_\omega|}{\omega} \mathbf{k}^{(0)} \times \mathbf{E}_0 e^{-sz} \cos(\omega t - kz + \varphi). \quad (63.27)$$

Consequently, the electric and magnetic vectors oscillate in a plane wave in different phases. Using (63.7) we find

$$\tan \varphi = -s/k = \sqrt{\varepsilon\mu/\gamma} - \sqrt{1 + (\varepsilon\mu/\gamma)^2}, \quad (63.28)$$

i.e. the angle φ is negative. This means that **B** attains a certain value later than **E**. This is manifested in two ways.

If we consider the oscillations of the field vectors of a wave at a fixed point which the wave travels by, **B** attains, say, its maximum value later than **E** does. In other words, **B** as a function of time lags behind **E**.

If we consider a wave at a certain fixed instant of time, **B** attains its maximum value for smaller values of z than **E** does, since **B** as a function of z leads **E**.

These statements complement one another and are united in the fact that a running electromagnetic wave moves in the direction of its propagation (in the case under consideration, in the positive direction of the Z-axis).

Relationship between the amplitudes of field vectors. It follows from (63.25) and (63.26) that

$$\frac{|\mathbf{B}|}{|\mathbf{E}|} = \frac{|k_\omega|}{\omega} = \sqrt{\mu\varepsilon}\,\{1 + [\gamma/(\varepsilon\mu)]^2\}^{1/4}. \quad (63.29)$$

Comparing this relation with (62.25) we conclude that the ratio between $|\mathbf{B}|$ and $|\mathbf{E}|$ in a conducting medium is larger than that in a nonconducting medium with the same values of μ and ε.

Sec. 64. Invariance of a Plane Wave

The invariants of electromagnetic field transformations are considered and the results of an analysis of the invariants are discussed.

Field transformation. The strengths of fields vary as we go over from one inertial reference system to another. Field transformations are described by formulas (11.15).

It may so happen that one inertial reference system contains electric and magnetic fields while another system contains only an electric field.

A plane electromagnetic wave is characterized by quite definite properties: vectors **E** and **B** are mutually perpendicular and their magnitudes are connected

www.EngineeringBooksPDF.com

through the relation $\mathbf{E} = cB$. Do these properties of the field vector remain invariant as we go over to another inertial reference system? If the answer is affirmative, the concept of a plane electromagnetic wave is a relativistic invariant reflecting the intrinsic properties of the electromagnetic field of the plane wave. Otherwise, this concept depends on a random choice of a certain inertial reference system and does not describe any actually existing physical object. Using formulas (11.15), it can be easily verified that the field vectors of an electromagnetic field satisfying the condition of plane wave in one coordinate system satisfy this condition in any other coordinate system, i.e. a *plane wave is a relativistically invariant concept that determines an actually existing physical object.* Instead of verifying directly a particular statement about the invariance of the plane wave, it is more expedient to analyze a wider problem on the invariants of electromagnetic field transformations and then substantiate the plane wave invariance as a particular conclusion along with which many other important conclusions, however, are obtained.

Invariants of electromagnetic field transformations. This term is applied to quantities composed of field vectors and remaining unchanged as we go over from one inertial reference system to another. The field vectors in different coordinate systems are related by formulas (11.15).

There are several ways of determining the invariants of transformations. Using formulas (11.15), it can be shown by direct calculations that the following quantities remain invariant upon a transition from one inertial reference system to another:

$$I_1 = c^2B^2 - E^2, \quad I_1' = H^2 - c^2D^2; \tag{64.1}$$

$$I_2 = \mathbf{B}\cdot\mathbf{E}, \quad I_2' = \mathbf{H}\cdot\mathbf{D}; \tag{64.2}$$

$$I_3 = \mathbf{H}\cdot\mathbf{B} - \mathbf{D}\cdot\mathbf{E}. \tag{64.3}$$

By way of an example, let us verify that the quantity I_2 is an invariant. In accordance with formulas (11.15), we have

$$\mathbf{B}'\cdot\mathbf{E}' = B_x'E_x' + B_y'E_y' + B_z'E_z' = B_xE_x + \frac{E_y - vB_z}{\sqrt{1-\beta^2}}\,\frac{B_y + (v/c)^2 E_z}{\sqrt{1-\beta^2}}$$

$$+ \frac{E_z + vB_y}{\sqrt{1-\beta^2}}\,\frac{B_z - (v/c^2)E_y}{\sqrt{1-\beta^2}} = B_xE_x + B_yE_y + B_zE_z = \mathbf{B}\cdot\mathbf{E}. \tag{64.4}$$

The invariance of other quantities can be proved in a similar way.

A plane wave is defined by the equality to zero of invariants I_1 and I_2, and its invariance does not require any further proof since I_1 and I_2 are invariants. However, the invariance of the quantities given by (64.1)-(64.3) allows us to draw some other important conclusions as well, concerning the behaviour of electromagnetic fields upon a transition from one reference system to another.

Analysis of field invariants. The invariance of quantities (64.1)-(64.3) lead to the following conclusions.

(1) If $c^2B^2 > E^2$ and $\mathbf{B} \perp \mathbf{E}$ in a certain inertial reference system, we can choose an inertial reference system in which there is no electric field, while

www.EngineeringBooksPDF.com

the magnetic field is other than zero. If, however, **B** is not perpendicular to **E**, such an inertial system does not exist.

(2) If $c^2B^2 < E^2$ and $\mathbf{B} \perp \mathbf{E}$ in a certain inertial reference system, we can choose an inertial reference system such that there is no magnetic field in it, while the electric field differs from zero. If **B** is not perpendicular to **E**, such an inertial reference system does not exist.

(3) If either an electric or a magnetic field exists in a certain inertial reference system, upon a transition to another inertial reference system we generally will have both an electric and a magnetic field which are perpendicular to each other.

(4) A plane wave for which $E = cB$ and $\mathbf{E} \perp \mathbf{B}$ remains a plane wave in all inertial reference systems.

Sec. 65. Pressure of Electromagnetic Waves. Photon Momentum

The emergence of electromagnetic wave pressure is described. The volume density of electromagnetic wave momentum is calculated and the photon momentum is defined.

Emergence of pressure. If a plane wave propagates in a conducting medium, in accordance with Ohm's law, its electric field induces in the medium a conduction current of the volume density

$$\mathbf{j} = \gamma \mathbf{E}. \tag{65.1}$$

The force exerted on the current element $\mathbf{j}\,dV$ by the magnetic field of the wave (Fig. 259) is given by

$$d\mathbf{F} = \mathbf{j} \times \mathbf{B}\,dV = \gamma \mathbf{E} \times \mathbf{B}\,dV. \tag{65.2}$$

This force acts along the vector $\mathbf{E} \times \mathbf{B}$, i.e. along the direction of wave propagation. Denoting the unit vector in the direction of wave propagation by **n**, we can write

$$\begin{aligned} d\mathbf{F} &= \gamma \mathbf{E} \times \mathbf{B}\,dV = \mathbf{n}\gamma EB\,dV \\ &= \mathbf{n}jE\,dV/v = \mathbf{n}\,dP/v, \end{aligned} \tag{65.3}$$

where we used the relation between the magnitudes of the field vectors in a plane wave ($E = vB$) and took into account Joule's law $dP = jE\,dV$. It is worth noting that the quantity dP in formula (65.3) is the energy absorbed per unit time.

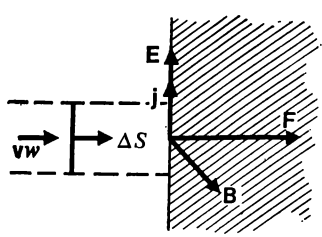


Fig. 259. Emergence of pressure of an electromagnetic wave

www.EngineeringBooksPDF.com

Pressure. Suppose that an energy flow of electromagnetic waves is incident from vacuum on a conducting medium and is entirely absorbed. In accordance with formula (62.27), the energy absorbed by the surface element dS in 1 s is

$$dP = vw\, dS. \tag{65.4}$$

This energy gives rise to force (65.3) directed along the normal and equal to

$$d\mathbf{F} = \mathbf{n}w\, dS. \tag{65.5}$$

Consequently, the pressure along the normal to the surface is given by

$$\mathbf{p} = \frac{d\mathbf{F}}{dS} = \mathbf{n}w. \tag{65.6}$$

The quantity

$$w = {}^1/_2\,(\mathbf{E}\cdot\mathbf{D} + \mathbf{B}\cdot\mathbf{H}) \tag{65.7}$$

is the electromagnetic wave energy per unit volume.

The momentum of an electromagnetic wave train. Suppose that the energy W contained in a certain volume in an electromagnetic wave train is absorbed by a certain volume of a conducting body during the time Δt. Then, in accordance with (65.3), the force exerted on this volume of the body is

$$\mathbf{F} = \mathbf{n}\frac{W}{\Delta t}\frac{1}{v}. \tag{65.8}$$

According to Newton's second law, the force acting on a volume is connected to the momentum acquired by the volume through the following relation:

$$\mathbf{F} = \mathbf{p}/\Delta t. \tag{65.9}$$

Substituting (65.9) into (65.8), we obtain

$$\mathbf{p} = \mathbf{n}\frac{W}{v}. \tag{65.10}$$

This formula contains a statement of fundamental importance: *an electromagnetic wave train, having an energy W and travelling at a velocity v, has the momentum* $\mathbf{p}$ *connected with the energy through relation* (65.10). *The momentum is directed along the line of wave propagation.*

Electromagnetic wave momentum per unit volume. Dividing both sides of (65.10) by the volume, in which the energy W is contained, we obtain the following formula for the electromagnetic wave momentum per unit volume:

$$\mathbf{G} = \mathbf{p}/V = \mathbf{n}w/v, \tag{65.11}$$

where $w = W/V$ is the electromagnetic energy of a plane wave per unit volume. Using (62.27), we can write Eq. (65.11) in the form

$$\mathbf{G} = \mathbf{S}/v^2, \tag{65.12}$$

where $\mathbf{S}$ is Poynting's vector and v is the velocity of wave propagation.

The electromagnetic wave pressure can be calculated from the change in the wave momentum. For example, if electromagnetic waves are incident on the

www.EngineeringBooksPDF.com

surface along the normal and are completely absorbed, then, according to (65.12), the pressure is

$$p = vG = S/v = w, \tag{65.13}$$

which naturally coincides with (65.6). On the other hand, if a wave is completely reflected, the momentum imparted to the body is twice as large, and the pressure is

$$p = 2vG = 2w. \tag{65.14}$$

Similarly we can calculate the pressure in the case of partial absorption, an oblique incidence on the surface, and so on.

The pressure of light waves was experimentally observed for the first time in 1900 by P. N. Lebedev (1866-1912). It can be seen from (65.14) that this pressure is very small. For instance, for a flux of 1.4 kW/m^2, which is approximately equal to the flux of solar radiation on the Earth orbit, the light pressure amounts to about 5 μPa. For this reason, very fine methods of measurements were required for detecting this pressure.

Momentum of a photon. According to quantum-mechanical concepts, light is a stream of energy quanta called photons. The energy of a photon is connected to the frequency of light through the Einstein relation:

$$\boxed{\varepsilon = \hbar\omega,} \tag{65.15}$$

where $\hbar$ is Planck's constant. The *existence of light pressure indicates that a photon has a momentum.* In accordance with (65.10), the photon momentum is given by

$$\mathbf{p} = \mathbf{n}\hbar\omega/c, \tag{65.16}$$

where c is the velocity of light in vacuum. Let us write this formula, taking into account (62.23):

$$\boxed{\mathbf{p} = \hbar\mathbf{k}.} \tag{65.17}$$

Along with (65.15), this relation is a fundamental equation in the quantum theory of light.

The electric field of a plane wave induces conduction currents in a conducting medium. As a result of interaction of these currents with the magnetic field of the wave, a Lorentz force appears, which is manifested in the form of pressure of the electromagnetic wave.

What in the classical model is the nature of force leading to the emergence of pressure when an electromagnetic wave is absorbed by a conducting medium?
What does the momentum density of an electromagnetic wave depend on?

Example 65.1. *Find the force with which photons having a flux density S of electromagnetic energy act on a perfectly reflecting sphere of radius r* (Fig. 260).

In view of the axial symmetry of pressure distribution, the only nonzero component of force will be that in the direction of the initial flow of photons. In accordance with formula (65.13), the force $dF = (2S/c)\cos\theta\, d\sigma$ exerted on a surface element $d\sigma$ (Fig. 260) is direct-

www.EngineeringBooksPDF.com

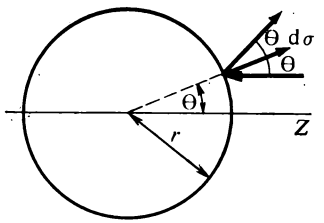


Fig. 260. Calculation of pressure of electromagnetic radiation on a perfectly reflecting sphere

ed towards the centre of the sphere. The component of this force along the Z-axis is given by

$$dF_z = -(2S/c)\cos^2\theta\, d\sigma.$$

The area of the surface element in the spherical system of coordinates is $d\sigma = r^2 \sin\theta\, d\theta\, d\alpha$, where α is the axial angle in the plane perpendicular to the Z-axis. Integrating this expression, we obtain the formula for the total force acting along the Z-axis:

$$F_z = -\left(\frac{2S}{c}\right) r^2 \int_0^{2\pi} d\alpha \int_0^{\pi/2} \cos^2\theta \sin\theta\, d\theta = \frac{4\pi}{3}\frac{S}{c} r^2,$$

i.e. the force is 4/3 times greater than in the case when the entire energy of the flow is absorbed by the sphere.

Sec. 66. Waveguides and Resonators

Basic characteristics of waveguides and the peculiarities of propagation of electromagnetic waves in them are considered. The classification of waves in waveguides is given. The principle of operation of resonators is discussed.

Subcircuit. Any subcircuit has an ohmic resistance, capacitance and inductance. Figure 261*a* presents an equivalent diagram of a subcircuit. Ohmic resistance R always differs from zero since wires have a resistance. Capacitance appears due to surface or volume charges which are always present in the subcircuit and due to electric fields in which the electric field energy is stored. When a current passes through a subcircuit, a magnetic field is generated and a certain amount of energy is stored in it. Consequently, a subcircuit also has an inductance. The role of R, C, and L depends on specific properties of the subcircuit and frequency.

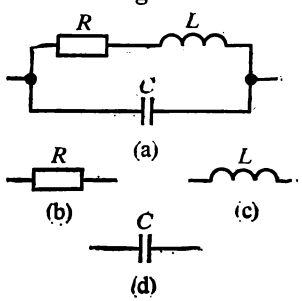


Fig. 261. Schematic diagram of a subcircuit at various frequencies

A section of a conductor. A small rectilinear section of a conductor contains a very small surface charge and a low magnetic field energy. This means that its capacitance and inductance are quite small. Therefore, at low frequencies the capacitive reactance of the section is larger than its ohmic resistance, while the inductive reactance is smaller than the ohmic resistance, i.e. the inequality $1/(\omega C) \gg R \gg \omega L$ is valid. Hence in the

www.EngineeringBooksPDF.com

circuit shown in Fig. 261*a*, the current mainly flows in the RL-branch as if the capacitor is disconnected. Since $\omega L \ll R$, the inductive reactance is insignificant, and for low frequencies the section of the conductor is depicted as shown in Fig. 261*b*.

The resistance R increases with frequency. Since the thickness of the skin layer decreases as $1/\sqrt{\omega}$, we can assume that the resistance grows in proportion to $\sqrt{\omega}$. The inductance L decreases only slightly with increasing frequency, and hence the inductive reactance ωL increases in proportion to ω. Consequently, with increasing frequency the relative role of the inductance in the conductor section increases, and it can no longer be treated as a section with a purely ohmic resistance. The capacitive reactance $1/(\omega C)$ decreases with increasing frequency. For this reason, at sufficiently high frequencies a considerable fraction of current is realized in the form of displacement currents. This means that for high frequencies the equivalent diagram of the section of the conductor has the form shown in Fig. 261*a*, where R as well as L and C should be taken into account. Their importance depends on frequency. At extremely high frequencies, the capacitance plays the decisive role.

Induction coil. At low frequencies, $1/\omega(C) \gg \omega L \gg R$ for a coil. The current mainly flows through R and L (see Fig. 261*a*), and since $R \ll \omega L$, the equivalent diagram for the coil has the form shown in Fig. 261*c*.

As the frequency increases, the inductive reactance of the coil grows while the capacitive reactance drops. Therefore, an increasingly large fraction of current passes in the form of displacement current through the capacitors connected between individual turns of the coil. Capacitance starts to play a significant role along with inductance and ohmic resistance. As a result, the equivalent diagram of the induction coil coincides with that shown in Fig. 261*a*, where the importance of R, L and C depends on frequency. At a very high frequency, almost the entire current flows in the form of displacement current, as if the induction coil is disconnected from the circuit.

Capacitor. At low frequencies, the capacitive reactance of a capacitor is smaller than the ohmic resistance and the inductive reactance [$1/(\omega C) \ll R$, $1/(\omega C) \ll \omega L$]. As a result, the RL-branch of the circuit shown in Fig. 261*a* is as if disconnected. The equivalent diagram of the capacitor has the form shown in Fig. 261*d*.

As the frequency increases, the situation changes. In order to analyze this situation, let us consider a parallel-plate capacitor.

An increase in frequency leads to a deviation of the electric field in a parallel-plate capacitor from the uniform field. The reason behind this is the interaction between the electromagnetic induction and the displacement currents. It seems at first sight that the pattern should be the same as that leading to the skin effect (see Fig. 223), but this is not so. The difference is due to some other phase relations between the field vectors.

Let us consider the vector diagram of fields and currents in the case of skin effect (see Fig. 223). The magnetic induction is in phase with the current and the strength of the electric field generating the magnetic field. The derivative of the magnetic induction leads them by $\pi/2$, while an additional electric field $\Delta \mathbf{E}$

www.EngineeringBooksPDF.com

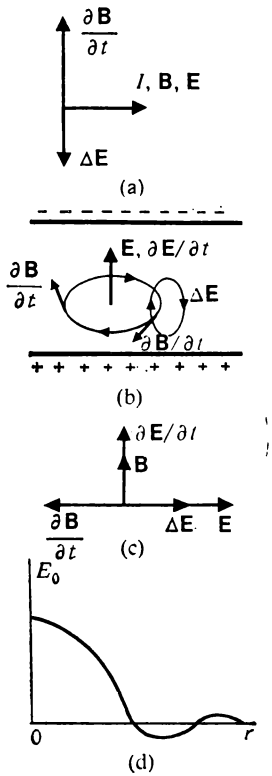


Fig. 262. Relation between field strengths in a capacitor at high frequencies

generated by the change in magnetic field and leading to the skin effect lags behind the electric field **E** by $\pi/2$. Consequently, with a more stringent approach, we should take into account in Fig. 223 not only the spatial distribution of the fields but also the phases of the electric field variation.

The vector diagram describing the appearance of the skin effect is shown in Fig. 262*a*.

Computational formulas automatically take into account the relation between the phases of the vectors.

A different relation exists between the phases of the field vectors in the case of a capacitor (Fig. 262*b*). Since the magnetic field is generated by displacement currents in accordance with the law

$$\operatorname{curl} \mathbf{B} = \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t},$$

the magnetic induction is in phase with $\partial\mathbf{E}/\partial t$, and hence leads the electric field strength **E** by $\pi/2$ (Fig. 262*c*). Consequently, the field $\Delta\mathbf{E}$, appearing according to the law of electromagnetic induction and leading to a redistribution of the electric field **E** in the capacitor, is in phase with the electric field **E** (Fig. 262*c*). The principal difference between these phenomena and the skin effect consists in the different relations between the phases of **E** and **B**: in the skin effect their phases coincide while in a capacitor the magnetic induction leads the electric field strength by $\pi/2$. Consequently, if, for example, the electric field strength is equal to zero, the magnetic induction is also equal to zero in the case of skin effect, while it has its maximum value for a capacitor. When the electric field strength increases from the zero value in the skin effect, the magnetic induction also increases and the vector $\partial\mathbf{B}/\partial t$ forms a right-handed system with vector **E** (see Fig. 223). In a capacitor, the magnetic induction decreases, and hence vector $\partial\mathbf{B}/\partial t$ forms a left-handed system with **E** (see Fig. 262*b*). Consequently, the strength ΔE of the vortex electric field is directed so that it enhances the electric field at the centre of the capacitor and weakens this field at the periphery. In other words, the field in a capacitor decreases from the centre to the periphery. At a certain distance from the centre, the electric field vanishes and then reverses its direction (Fig. 262*d*).

We can obtain a quantitative characteristic of this phenomenon by proceeding from (62.5) and solving the equation for the electric field strength **E**. In the case under consideration, we have only one component of vector **E** and the prob-

www.EngineeringBooksPDF.com

lem is axisymmetric, i.e. $E = E(r)$, where r is the distance from the capacitor axis to the point at which the electric field is determined. Putting, as usual,

$$E(r, t) = E_0(r)\, e^{i\omega t}$$

and assuming, for the sake of definiteness, that $\varepsilon = \varepsilon_0$ and $\mu = \mu_0$ between the capacitor plates, we obtain the following equation for $E_0(r)$:

$$\frac{d^2E_0}{dr^2} + \frac{1}{r}\frac{dE_0}{dr} + \frac{\omega^2}{c^2}E_0 = 0,$$

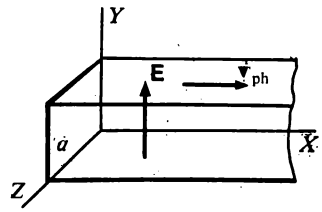


Fig. 263. Rectangular waveguide

which is written in cylindrical coordinates. This equation is called the zero-order Bessel equation, and its solution is written in the form $J_0(\omega r/c)$. Bessel functions are well known. Figure 262*c* shows the function $J_0(\omega r/c)$. The smallest roots of the zero-order function are $\xi_1 = 2.40$, $\xi_2 = 5.52$, $\xi_3 = 8.65$, ... We take into account the fact that $\omega/c = 2\pi/\lambda$, where λ is the wavelength of an electromagnetic wave of frequency ω in vacuum. Consequently, the distances at which the electric field in a capacitor vanishes are given by

$$r_i = \lambda\xi_i/(2\pi).$$

In particular, the field vanishes for the first time at a distance $r_1 = \lambda\xi_1/(2\pi) = 0.38\lambda$. On account of such a behaviour of the electric field, the capacitor no longer plays the role of a pure capacitance. Clearly, the magnetic fields in the capacitor become significant, which means that inductance comes into play. In other words, the capacitor also loses the property of capacitance at high frequencies.

Radiation. It was shown in Sec. 61 that the emissive power of an oscillator increases in proportion to the fourth power of frequency ($\sim\omega^4$), i.e. very rapidly. This means that when high-frequency currents pass through the wires, strong electromagnetic radiation appears. At high frequencies, the losses become so significant that transmission through the wires becomes inexpedient and it is necessary to find some other method of transmission of electromagnetic energy since the methods of generation and transmission of electromagnetic oscillations worked out for low frequencies are inapplicable at very high frequencies.

Waveguides. The main purpose of a waveguide is to direct electromagnetic waves along a certain channel, reducing possible losses during propagation to the minimum. Obviously, for this purpose we must avoid, if possible, the excitation of conduction currents and exclude penetration of electromagnetic energy behind the channel walls. A simple model of a waveguide is a hollow tube inside which electromagnetic waves propagate. We shall consider the basic properties of such electromagnetic waves by using a rectangular rectilinear waveguide as a simple example.

Rectangular waveguide. The walls of the waveguide are assumed to be perfectly conducting. The dimensions of the waveguide and the arrangement of the coordinate axes is shown in Fig. 263. Generally speaking, many types of waves may propagate through waveguides. Let us consider one of them.

www.EngineeringBooksPDF.com

Let us suppose that the electric vector of the wave is directed along the Y-axis. To simplify the situation, we assume that the length of the waveguide along the Y-axis is infinite. This allows us to disregard the boundary conditions for the vector $\mathbf{E}$ on the surfaces of the waveguide which are parallel to the XZ-plane and considerably simplifies the solution of the problem. Besides, in the case of an infinitely long waveguide oriented along the Y-axis, the problem can be solved by the image method, which allows us to clarify the physical situation and the essence of processes occurring during the propagation of waves in the waveguide.

The problem is thus reduced to two dimensions. The wave equation for the electric field strength has the form

$$\frac{\partial^2 E}{\partial x^2}+\frac{\partial^2 E}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}=0, \tag{66.1}$$

where $E = E_y(x, z, t)$.

Since the walls of the waveguide are perfect conductors, the boundary condition for E has the form

$$E(x, 0, t) = 0, \quad E(x, a, t) = 0. \tag{66.2}$$

We shall seek the solution of the equation in the form

$$E = E_0 \sin k_z z e^{i(\omega t - k_x x)}, \tag{66.3}$$

where in order to satisfy the boundary conditions (66.2), we must put

$$k_z a = n\pi \quad (n = 1, 2, \ldots). \tag{66.4}$$

It is also clear that solution (66.3) satisfies the condition of absence of free charges in the waveguide: $\operatorname{div}\mathbf{E} = \partial E_y/\partial y = 0$, $E_x = E_z = 0$. Substituting (66.3) into (66.1), we obtain

$$(-k_x^2 - k_z^2 + \omega^2/c^2)\, E = 0. \tag{66.5}$$

This equation can be satisfied only if

$$-k_x^2 - k_z^2 + \omega^2/c^2 = 0, \tag{66.6}$$

from which it follows that

$$k_x = \sqrt{\omega^2/c^2 - \pi^2 n^2/a^2}. \tag{66.7}$$

Cut-off frequency. An electromagnetic wave propagates in a waveguide without attenuation if the quantity k_x in (66.3) is real. This means that the radicand in (66.7) should not be negative. Hence we can obtain the condition for waves propagation in a waveguide:

$$\frac{\omega^2}{c^2}-\frac{\pi^2 n^2}{a^2}\geqslant 0 \tag{66.8}$$

or

$$\omega \geqslant \frac{\pi c}{a} n. \tag{66.9}$$

www.EngineeringBooksPDF.com

Thus, there exists a **cut-off frequency** for a given value of n characterizing the shape of the wave. Electromagnetic waves having a lower frequency cannot propagate in the waveguide. The value of this frequency is obtained from (66.9) for $n = 1$:

$$\omega_0 = \pi c/a. \tag{66.10}$$

In other words, the existence of the cut-off frequency indicates the existence of the maximum wavelength of a wave which can propagate in the waveguide. Considering that $\lambda = cT = 2\pi c/\omega$, we obtain the following expression for the boundary wavelength:

$$\lambda_0 = 2\pi c/\omega_0 = 2a. \tag{66.11}$$

This expression has a clear geometrical meaning: *in a waveguide under consideration, only waves with a wavelength smaller than twice the transverse dimension of the waveguide can propagate.*

The existence of the cut-off frequency is a typical feature of all waveguides, although its value is different for different waveguides.

Phase velocity. According to (66.3), this velocity can be found from the condition

$$\omega t - k_x x = \text{const}, \tag{66.12}$$

whence

$$v_{\text{ph}} = \frac{dx}{dt} = \frac{\omega}{k_x} = \frac{\omega}{\sqrt{\omega^2/c^2 - \pi^2/a^2}} = c\frac{\omega}{\sqrt{\omega^2 - \pi^2 c^2/a^2}} > c, \tag{66.13}$$

This means that the phase velocity of electromagnetic waves in a waveguide is higher than the velocity of light. This is another typical feature of waveguides, although the concrete value of phase velocity depends on the properties of a waveguide and the type of waves.

Considering Eqs. (66.10) and (66.11), formula (66.13) can conveniently be written in the form

$$v_{\text{ph}} = \frac{c}{\sqrt{1 - (\omega_0/\omega)^2}} = \frac{c}{\sqrt{1 - (\lambda/\lambda_0)^2}}. \tag{66.14}$$

Consequently, $\omega \geqslant \omega_0$, and $\lambda \leqslant \lambda_0$, since otherwise the phase velocity becomes imaginary, i.e. the propagation of waves becomes impossible.

Wavelength in a waveguide. By definition of a wavelength, we have

$$\lambda_{\text{w}} = v_{\text{ph}} T = \frac{\lambda}{\sqrt{1 - (\lambda/\lambda_0)^2}} > \lambda, \tag{66.15}$$

where $\lambda = cT$. The wavelength in a waveguide is always larger than the wavelength in free space. Squaring both sides of (66.15) and taking the reciprocals, we obtain

$$1/\lambda_{\text{w}}^2 = 1/\lambda^2 - 1/\lambda_0^2. \tag{66.16}$$

This relation is valid for waveguides of any shape in spite of the fact that it was derived here for a particular case.

www.EngineeringBooksPDF.com

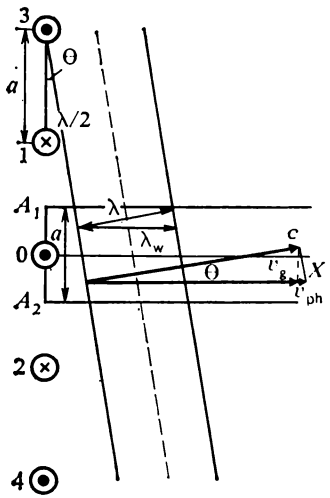


Fig. 264. Analysis of a rectangular waveguide by the method of images

Application of the method of images to the analysis of waveguides. To clarify the physical nature of propagation of waves in a waveguide and the meaning of the relations obtained above, let us analyze the above example by the method of images. For an elementary radiator we shall take a very long straight conductor carrying an alternating current of frequency ω. Like a Hertzian dipole, this radiator emits waves whose electric vector is parallel to the conductor. In the case of a very long conductor the waves will obviously be cylindrical. However, at a sufficiently large distance from the radiator, they can be assumed to be plane.

Figure 264 shows the projections of waveguide walls onto the XY-plane, the electric vector of the waves being perpendicular to the plane of the figure. We arrange the first radiator at the middle of the waveguide at a distance $a/2$ from its wall perpendicular to the plane of the figure. The phase of oscillations of the radiator is shown by a dot, i.e. at a given instant of time the current flows towards us. The oscillator emits waves in all directions, and hence the field strength on the waveguide walls differs from zero. The problem consists in selecting a system of radiators such that the total strength of their fields at the waveguide walls is always equal to zero. The field satisfying this condition will be the required field in the waveguide. Naturally, when the waves propagate from imaginary radiators, waveguide walls are also assumed to be imaginary, and imaginary waves freely pass through them.

In order to nullify the field generated by radiator *0* at the wall A_1 of the waveguide, we must place radiator *1* at a distance $a/2$ from this wall, which oscillates with a phase shift of half a period relative to radiator *0*. Consequently, the phase of oscillations of radiator *1* should be opposite to that for oscillator *0*, which is shown by the sign "+" (current flows away from us). The waves emitted by radiator *1* arrive at points of wall A_1 after the same period of time as from radiator *0*. Since the phases of the waves from *0* and *1* at wall A_1 differ by π, the sum of field strengths of these waves is equal to zero. In a similar way, radiator *2* neutralizes the radiation from radiator *0* at wall A_2.

However, radiator *1* creates a field at wall A_2, while radiator *2* creates the field at wall A_1. We must add some more radiators that would quench these fields. In order to neutralize the radiation emitted by *1* at wall A_2, it is necessary to use radiator *4*, radiator *3* serves to quench radiation from *2* at wall A_1, and so on and so forth. The field strength created by an infinite system of these radiators at walls A_1 and A_2 is equal to zero. Consequently the obtained field satisfies Maxwell's equations since it is a superposition of the fields each of which satisfies these equations, and is the required electromagnetic wave in the waveguide.

www.EngineeringBooksPDF.com

The field outside the waveguide has an auxiliary meaning and is of no interest to us.

Discrete nature of directions of propagation of plane waves from system oscillators. Plane waves from an individual radiator propagate in all directions. However, plane waves emitted by a system of radiators may propagate only in quite definite directions. These are the directions in which the plane waves from individual radiators enhance each other. This is possible only when the path lengths differ for the waves emitted by neighbouring radiators by an integral number of wavelengths and a half, since neighbouring radiators emit in antiphase. As a result, it turns out that in a certain selected direction the waves emitted by different radiators have a phase difference equal to an integral number of periods, and hence these waves enhance each other. In Fig. 264 the direction of propagation of waves is characterized by an angle θ. The condition of mutual augmentation of waves has the form

$$a \sin \theta = \lambda (m + 1/2) \quad (m = 0, 1, 2, \ldots). \tag{66.17}$$

A similar condition can be written for waves propagating in the opposite direction along the waveguide axis, i.e. for negative values of θ.

Boundary wavelength. Condition (66.17) indicates that for each wavelength there exists a minimum angle between the direction of propagation and the waveguide axis, which is attained for $m = 0$. There also exists a maximum value of m for which the angle θ is equal to $\pi/2$, i.e. the wave propagates at right angles to the waveguide length. For a sufficiently large wavelength, even at $m = 0$ we arrive at the condition $\sin \theta = 1$, i.e. this wave can propagate only in a direction perpendicular to the waveguide axis. This means that the waves with such and larger wavelengths cannot propagate in the waveguide. This is the boundary wavelength λ_0 determined from (66.17) for $\sin \theta = 1$, $m = 0$:

$$a = \lambda_0/2, \tag{66.18a}$$

which coincides with (66.11). This wavelength corresponds to the boundary frequency (66.10).

Wavelength and phase velocity in a waveguide. Phase velocity is the velocity of points on an equiphase surface of a wave propagating along a waveguide, i.e. the velocity of the point of intersection of the waveguide walls with the front of the plane wave. Figure 264 shows that this velocity is equal to

$$v_{ph} = c/\cos \theta. \tag{66.18b}$$

Taking in (66.17) a wave with $m = 0$, we obtain $\sin \theta = \lambda/(2a)$ and represent formula (66.18b) in the form

$$v_{ph} = \frac{c}{\sqrt{1-\sin^2\theta}} = \frac{c}{\sqrt{1-[\lambda/(2a)]^2}} = \frac{c}{\sqrt{1-(\lambda/\lambda_0)^2}} = \frac{c}{\sqrt{1-(\omega_0/\omega)^2}}, \tag{66.19}$$

which coincides with (66.14). Thus, *the phase velocity is not associated with the motion of any physical object and energy in space.* We can imagine that the X-axis in Fig. 264 depicts the edge of a desk, while the line depicting the wave-front is a ruler. Then for an angle θ sufficiently close to $\pi/2$, low velocities of motion of the ruler perpendicular to its length lead to the velocities of the point of

www.EngineeringBooksPDF.com

contact between the ruler and the edge of the desk exceeding the velocity of light. Obviously, the existence of this velocity does not contradict the limitations imposed by the theory of relativity on the velocity of motion of physical objects and on the velocity of propagation of interactions.

The wavelength λ is also determined with the help of geometrical construction in Fig. 264:

$$\lambda_w = \frac{\lambda}{\cos\theta} = \frac{\lambda}{\sqrt{1-(\lambda/\lambda_0)^2}}, \tag{66.20}$$

which coincides with (66.15). Equation (66.16) also follows from this equation.

Group velocity. It is clear that the phase velocity does not represent the velocity of propagation of the wave energy along a waveguide. The energy of a plane wave propagates in vacuum at the velocity c at right angles to the wave-front. The velocity of propagation of energy in the direction of the waveguide axis is determined by the projection of the velocity c onto this axis. This quantity is called the **group velocity**. It can be seen from Fig. 264 that it is given by

$$v_g = c\cos\theta = c\sqrt{1-(\lambda/\lambda_0)^2} \tag{66.21}$$

The group velocity is always smaller than the velocity of light. The term used for this quantity is explained by the fact that it is equal to the velocity of the peak value of the total amplitude of a group of waves having close frequencies and propagating with different phase velocities which depend on frequency. An aggregate of waves with different frequencies in a waveguide forms such a group of waves for which the frequency dependence of phase velocities is determined by formula (66.14). The most important physical property of the group velocity has already been formulated: this is the velocity of propagation of the energy associated with waves.

Relation between the group velocity and phase velocity. Multiplying Eqs. (66.21) and (66.19) termwise, we obtain

$$\boxed{v_{ph}v_g = c^2.} \tag{66.22}$$

This relation is of fundamental importance in the theory of propagation of waves and is of universal nature, although is was obtained for a particular case by a special method.

Magnetic field. The magnetic induction of the field of a plane wave is perpendicular to its electric field strength. Consequently, the vectors of magnetic induction lie in planes parallel to the plane of Fig. 264. Since plane waves propagate at an angle to the waveguide axis, the magnetic induction of the field of each plane wave has components directed along the waveguide axis and at right angles to it.

The same applies to the magnetic induction of the field of superposition of plane waves that form a wave in the waveguide. This means that electromagnetic waves propagating in the waveguide are not purely transverse waves. They also have magnetic induction components in the direction of their propagation. In other cases, there may exist waves with an electric field component along the

www.EngineeringBooksPDF.com

direction of their propagation, and so on. It should also be noted that waves in a waveguide are generally inhomogeneous.

Classification of waves in waveguides. The following classification of waves in waveguides is generally adopted.

1. Transverse magnetic waves (TM-waves) defined by the condition $H_x = 0$, i.e. the magnetic field strength component along the direction of propagation of waves is equal to zero. It can be shown that in this case all characteristics of waves are expressed only in terms of E_x.

2. Transverse electric waves (TE-waves) defined by the condition $E_x = 0$. In this case, the solutions are expressed only in terms of H_x.

3. Transverse electromagnetic waves (TEM-waves) defined by the requirements $E_x = 0$ and $H_x = 0$.

4. Hybrid waves for which $H_x \neq 0$ and $E_x \neq 0$ simultaneously. They appear when the boundary conditions require that E_x and H_x are simultaneously different from zero, which can be realized in real waveguides with a finite electric conductivity of walls.

Resonators. Let us consider a capacitor whose electric field strength varies with distance at high frequencies as shown in Fig. 262*d*. There is no electric field on the cylindrical surface of radius r_1. This means that Poynting's vector on this surface is equal to zero, and hence the electromagnetic energy flux through this surface is equal to zero. We shall assume that this cylindrical surface is a perfect conductor connecting the capacitor plates. The electric field on its surface remains equal to zero. The magnetic field differs from zero, and its lines of force are circles concentric with the points on the cylinder axis. The currents flow along the cylindrical conductor from one capacitor plate to the other, as it follows from the boundary condition (38.35) for the tangential component of vector **H**. Now the entire closed cylindrical volume bounded by perfectly conducting walls can be insulated and left alone. The electric field in it will oscillate at a frequency ω which is equal to the rate of recharging the capacitor plates. A closed volume in which electromagnetic field oscillations occur is called a resonator. The frequency of field oscillations in the absence of electromagnetic energy losses is called the natural frequency of the resonator. The resonator of this type is called the cylindrical resonator. In resonators, as well as in waveguides, different types of oscillations and standing waves may exist. They have different resonance frequencies. For the type of oscillations in a cylindrical resonator considered above the resonance frequencies ω_i of oscillations are given by $\omega_i = \xi_i c/r_0$, where ξ_i are the roots of the zero-order Bessel's functions. Thus, for this type of oscillations the resonator has a countless number of resonance frequencies rather than one resonance frequency. For other possible types of oscillations other resonance frequencies are obtained. In a real resonator there are always energy losses and oscillations are damped. The terms and concepts associated with oscillations in resonators are completely identical to those used while considering mechanical oscillations.

A characteristic feature of any wavguide is the existence of the cut-off frequency. In any waveguide, the phase velocity of electromagnetic waves is higher than the velocity of light.

www.EngineeringBooksPDF.com

Problems

9.1. Find the average emissive power of a loop with current $I = I_0 \cos \omega t$. The area of the loop is equal to σ. Assume that $I_0 = 10$ A, $\sigma = 100$ cm^2 and $\omega = 10^8$ s^{-1}.

9.2. Using the data of Problem 9.1, calculate the maximum radiant flux density in the plane of a current loop at a distance of 200 m from it.

9.3. Find the dipole arm if its emissive power is the same as that of the current loop in Problem 9.1. The frequency of dipole oscillations is equal to the frequency of current oscillations in the loop and each dipole charge is equal to $|q| = 10^{-4}$ C.

9.4. Electrical breakdown in air occurs at an electric field strength $E \simeq 30$ kV/cm. What flux density of plane electromagnetic wave energy corresponds to this breakdown at a moderate frequency?

9.5. A plane polarized electromagnetic wave with a cyclic frequency $\omega = 10^6$ s^{-1} is incident from the edge of a conducting loop so that the magnetic induction vector of the wave is perpendicular to the plane of the loop. The linear dimensions of the loop are small in comparison with the wavelength. The area of the loop is $\sigma = 100$ cm^2 and the average energy flux density in the wave is $\langle S \rangle = 1$ W/m^2. Find the maximum e.m.f induced in the loop.

9.6. The solar energy flux on the Earth orbit is approximately equal to $S = 1.4$ kW/m^2. Find the radius of a black spherical particle of density $\rho = 5$ g/cm^3 for which the light pressure in the interplanetary space is equal to solar attraction. The mass of the Sun is $m_S = 2 \times 10^{30}$ kg, the gravitational constant $G = 6.7 \times 10^{-11}$ N·m^2/kg^2. The distance between the Earth and the Sun is $R = 150 \times 10^6$ km.

9.7. A parallel-plate capacitor with circular plates of radius a is connected to a source of constant extraneous e.m.f.s $\mathscr{E}_{ext}$. The distance between the plates varies slowly according to the harmonic law $d = d_0 + \Delta \sin \omega t$. Find the magnetic field generated between the plates by displacement currents.

9.8. A loop of n turns having an area S lies in the plane XZ. A plane electromagnetic wave propagates along the X-axis. The electric field vector of this wave is parallel to the Y-axis: $E_y = E_0 \cos(\omega t - kx)$. Find the electromotive force induced in the loop. The wavelength is much larger than linear dimensions of the loop.

9.9. The solar energy flux on the Earth orbit is equal to $S = 1340$ W/m^2. Find the amplitudes E_0 and B_0 of a plane electromagnetic wave having such an energy flux density.

9.10. It follows from formula (65.14) that the pressure exerted by an electromagnetic wave on a perfectly reflecting surface at the angle of incidence θ is $p_0 = 2w \cos^2 \theta$, where w is the electromagnetic energy density of the incident wave. Suppose that an isotropic radiation is incident on the surface, i.e. the density of energy fluxes arriving from different directions is the same. Find the wave pressure on the surface.

9.11. Find the amplitude of the electric field strength of an electric dipole radiation in the plane passing through a dipole at right angles to its direction at a distance of 10 km from the dipole for the dipole emissive power of 10 kW.

9.12. The medium between the plates of a parallel-plate capacitor has a permittivity ε and a low electric conductivity γ (nonideal dielectric). The capacitance of the capacitor is C. A potential difference U is applied to the capacitor plates, after which they are insulated. Find the law of time-variation of charge on each capacitor plate and the displacement current flowing through the capacitor.

Answers

9.1. $\langle P \rangle = \mu_0 \omega^4 I_0^2 \sigma^2 / (12\pi c^3) = 0.124$ W. **9.2.** $S_{max} = \mu_0 \omega^4 I_0^2 \sigma^2 / (16\pi^2 c^3 r^2) = 0.47 \cdot 10^{-5}$ W/m^2. **9.3.** $l = I_0 \sigma / (|q| c) = 3.3 \cdot 10^{-4}$ m $= 0.33$ mm. **9.4.** $\langle S \rangle = [\varepsilon_0/(4\mu_0)]^{1/2} E_0^2 = 1.2 \cdot 10^3$ kW/cm^2 = 12 GW/m^2. **9.5.** $\mathscr{E}_{max}^{ind} = \sqrt{2 \langle S \rangle \mu_0} (\varepsilon_0 \mu_0)^{1/4} \sigma \omega = 9$ mW. **9.6.** $r = \dfrac{3SR^2}{4Gm_C\rho} = 0.5 \cdot 10^{-7}$ m.

9.7. $H_\varphi = -\varepsilon_0 \mathscr{E}_{ext} \omega \Delta r \cos \omega t / [2(d_0 + \Delta \sin \omega t)^2]$. **9.8.** $\mathscr{E}^{ind} = nkSE_0 \sin \omega t$. **9.9.** $E_0 = 1005$ W/m, $B_0 = 3.35 \cdot 10^{-6}$ T. **9.10.** $p = w_{tot}/3$. **9.11.** $E_0 = 0.095$ W/m. **9.12.** $Q = CUe^{-\gamma t/\varepsilon}$, $I_{dis} = -(\gamma/\varepsilon) CUe^{-\gamma t/\varepsilon}$.

www.EngineeringBooksPDF.com

CHAPTER 10

Fluctuations and Noises

Noises in a current loop are due to a discrete nature of charge carriers and current fluctuations. Noises cannot be completely eliminated in principle, but they can be reduced. Under certain conditions, it is possible to detect signals below the noise level.

Sec. 67. Fluctuations in a Current-Carrying Loop. Resistance Noise

The physical reasons behind the existence of noise are discussed. Quantitative characteristics of noise in current-carrying circuits are considered.

Theorem on equipartition of energy. The conclusion that each degree of freedom of a system in thermodynamic equilibrium corresponds to the same energy equal to $kT/2$ (k is the Boltzmann constant and T is the thermodynamic temperature) plays an important role in statistical physics. The Brownian movement illustrates the validity of this statement. The mean kinetic energy $\langle (mv^2/2) \rangle$ of translational motion of a Brownian particle satisfies the relation $\langle mv^2/2 \rangle = 3\,kT/2$, since there are three degrees of freedom of translational motion.

Application of the equipartition theorem to a freely suspended mirror. If a small mirror is freely suspended from an elastic string, it cannot be completely at rest in accordance with the equipartition theorem. As a result of interaction of the mirror with thermal motion of the air molecules, torsional vibrations are excited and each degree of freedom in this case must have an energy $kT/2$. It should be recalled that the theorem on equipartition of energy is applicable not only to the kinetic energy of an oscillator, but to its potential energy as well.

Let us denote the torsion modulus of the string by D, and the angle of its deflection from the equilibrium position (Fig. 265) by φ. The equation for the torsional vibrations can be written in the form

$$J\ddot{\varphi} = -D\varphi, \tag{67.1}$$

where J is the moment of inertia of the mirror with respect to the torsion axis. Multiplying both sides of (67.1) by $\dot{\varphi}$ and integrating the expression thus obtained, we can find the law of conservation of energy:

$$1/2\,J\dot{\varphi}^2 + 1/2D\,\varphi^2 = \text{const}. \tag{67.2}$$

www.EngineeringBooksPDF.com

Since the energy corresponding to each degree of freedom is equal to $kT/2$, we obtain from (67.2)

$$\langle 1/2\, J\dot{\varphi}^2\rangle = \langle 1/2\, D\varphi^2\rangle = 1/2\, kT. \tag{67.3}$$

Consequently,

$$\langle \varphi^2\rangle = kT/D. \tag{67.4}$$

This means that the mirror cannot be in its equilibrium position, but oscillates about it with a mean square deflection given by (67.4). Thus, (67.4) characterizes the deviation of the angle from the mean value, i.e. describes the fluc-

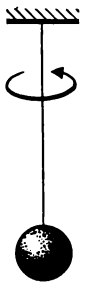


Fig. 265. Fluctuations of torsional vibrations

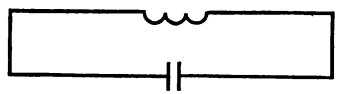


Fig. 266. Fluctuations in an oscillatory circuit

tuations. Obviously, if there is a certain torsional vibration, we can conclude in accordance with the superposition principle that Eq. (67.4) characterizes the fluctuation of the square of the amplitude.

Fluctuations in an oscillatory circuit. Oscillations with a frequency $\omega = 1/\sqrt{LC}$ are generated in the oscillatory circuit shown in Fig. 266. Physically, these oscillations cause the mutual conversion of the electric energy in the capacitor and the magnetic energy in the induction coil. The law of conservation of energy has the following form in this case:

$$Q^2/(2C) + LI^2/2 = \text{const}. \tag{67.5}$$

where Q is the charge on the capacitor plates and I is the current in the circuit.

It is impossible to imagine a circuit in which there is absolutely no current at all and the capacitor plates carry no charge at all. To be more precise, such a situation can be imagined only at 0 K. At any other temperature, the thermal motion of electrons leads to the appearance of a charge on the capacitor plates and a current in the circuit. From the theorem on equipartition of energy, we have

$$\langle Q^2/(2C)\rangle = \langle LI^2/2\rangle = kT/2. \tag{67.6}$$

Consequently, the mean square charge on the capacitor plates and the mean square current are given by the formulas

$$\boxed{\langle (Q)^2\rangle = kTC,\ \langle (I)^2\rangle = kT/L.} \tag{67.7}$$

www.EngineeringBooksPDF.com

On the basis of the principle of superposition, it can be stated that (67.7) represents the mean square fluctuations of charge and current in an oscillatory circuit.

Frequency distribution of fluctuations. Formula (67.7) describes only the total mean square value of fluctuations and does not give any idea of its frequency distribution. In order to solve this problem, we must solve the equation describing oscillations in an oscillatory circuit subjected to random forces which are expressed in the form of a Fourier series (integral) in frequencies:

$$U = \sum_{\omega} U_{\omega} e^{i\omega t}. \tag{67.8}$$

Equation (50.10) for the oscillations of the capacitor charge in this case assumes the form

$$L\ddot{Q} + R\dot{Q} + Q/C = \sum_{\omega} U_{\omega} e^{i\omega t}, \tag{67.9}$$

whence

$$Q = \sum_{\omega} \frac{U_{\omega} e^{i\omega t}}{-L\omega^2 + iR\omega + 1/C}, \tag{67.10}$$

which can be verified by differentiation. For mean square amplitude, we have $\langle | QQ^* | \rangle = \langle | Q |^2 \rangle$ whence

$$\langle |Q|^2 \rangle = \langle QQ^* \rangle = \left\langle \sum_{\omega,\ \omega'} \frac{U_{\omega} U^*_{\omega'} e^{i\omega t} e^{-i\omega' t}}{(-L\omega^2 + iR\omega + 1/C)(-L\omega^2 - iR\omega + 1/C)} \right\rangle. \tag{67.11}$$

The electromotive forces generating oscillations of different frequencies are independent and not mutually correlated. Hence terms with $\omega \neq \omega'$ are not considered while averaging in (67.11) and we get

$$\langle Q^2 \rangle = \langle |Q|^2 \rangle = \sum_{\omega} \frac{\langle U^2_{\omega} \rangle}{(L\omega^2 - 1/C)^2 + R^2\omega^2}, \tag{67.12}$$

where $\langle Q^2 \rangle$ and $\langle U^2_{\omega} \rangle$ are the average values of the real squares of amplitudes of the respective quantities.

Let us now turn to the continuous frequency spectrum, since the above calculations were carried out for a discrete spectrum just for the sake of simplicity. Actually, the spectrum of frequencies is continuous. We must go over from the mean square values of frequencies in a discrete spectrum to the densities of the corresponding quantities.

The mean square value of the total charge is the sum of the contributions from individual frequencies. Hence

$$\langle Q^2 \rangle = \int_0^{\infty} \frac{d \langle Q^2_{\omega} \rangle}{d\omega} d\omega, \tag{67.13a}$$

www.EngineeringBooksPDF.com

where $d\langle Q_\omega^2\rangle/d\omega$ is the density of the squares of charge oscillations amplitudes and $d\langle Q_\omega^2\rangle$ is the mean square value of the amplitude of charge oscillations in the frequency interval $(\omega, \omega + d\omega)$. We make the following substitution on the right-hand side of (67.12):

$$\langle U_\omega^2\rangle \to \frac{d\langle U_\omega^2\rangle}{d\omega}\, d\omega, \tag{67.13b}$$

where by $d\langle U_\omega^2\rangle/d\omega$ we mean the density of frequency distribution of the squares of voltage amplitudes. Having made such a substitution, we can go over from summation to integration in (67.12). This gives

$$\langle Q^2\rangle = \int_0^\infty \frac{[d\langle U_\omega^2\rangle/d\omega]\, d\omega}{(L\omega^2 - 1/C)^2 + R^2\omega^2} = \int_0^\infty \frac{d\langle Q_\omega^2\rangle}{d\omega}\, d\omega, \tag{67.14}$$

whence

$$d\langle Q_\omega^2\rangle = \frac{d\langle U_\omega^2\rangle}{(L\omega^2 - 1/C)^2 + R^2\omega^2}. \tag{67.15}$$

Resistance noise. The mean energy of harmonic oscillations is proportional to the square of the amplitude. Hence the *density of the mean square amplitude of oscillations characterizes their energy density*. Subsequent analysis is based on the assumption that the mean density of the squares of amplitudes

$$\left(\frac{d\langle U_\omega^2\rangle}{d\omega} = A\right)$$

is independent of frequency. The validity of this assumption stems from the random nature of electromotive forces. Consequently, we can write (67.14) in the following form:

$$Q^2 = A\int_0^\infty \frac{d\omega}{(L\omega^2 - 1/C)^2 + R^2\omega^2}, \quad A = \frac{d\langle U_\omega^2\rangle}{d\omega}. \tag{67.16}$$

The integral is evaluated by elementary methods and leads to the equality

$$\int_0^\infty \frac{d\omega}{(L\omega^2 - 1/C)^2 + R^2\omega^2} = \frac{\pi C}{2R}. \tag{67.17}$$

Taking into account (67.16) and (67.17), we obtain from (67.7)

$$d\langle U_\omega^2\rangle = (2/\pi)\, kRT\, d\omega. \tag{67.18}$$

On the basis of (67.15) we get from (67.18)

$$d\langle Q_\omega^2\rangle = \frac{(2/\pi)\, kTR\, d\omega}{(L\omega^2 - 1/C)^2 + R^2\omega^2}. \tag{67.19}$$

It should be noted that $d\langle U_\omega^2\rangle/d\omega$ defines the density of the mean square value of the amplitude, divided by the interval of cyclic frequencies ω. Quite

www.EngineeringBooksPDF.com

frequently, we use the density of mean square value of the amplitude divided by the frequency $\nu = 1/T$ instead of the cyclic frequency $\omega = 2\pi/T$, i.e. by the quantity $d\langle U_\nu^2\rangle/d\nu$. Considering that $\omega = 2\pi\nu$ and $d\omega = 2\pi\, d\nu$, we get

$$\frac{d\langle U_\omega^2\rangle}{d\omega} = \frac{1}{2\pi}\frac{d\langle U_\nu^2\rangle}{d\nu}. \tag{67.20}$$

In this case (see (67.18)), we get

$$d\langle U_\nu^2\rangle = 4kTR\, d\nu. \tag{67.21}$$

This relation is called the Nyquist formula: the *mean square value of the noise voltage amplitude is proportional to the frequency interval and depends only on the resistance of the circuit and on temperature.* The existence of this noise was experimentally discovered by Johnson. This noise is called the noise of a resistor or the Johnson noise.

Equivalent noise generator. Noise due to the resistance R, whose mean square voltage is determined by formula (67.21), can be represented as a result of the action of the source of e.m.f. U_ν and the internal resistance R. The equivalent current generator is shunted by resistance R and is characterized (in accordance with Ohm's law) by the mean square value of current:

$$d\langle I_\nu^2\rangle = 4kT\, d\nu/R. \tag{67.22}$$

Power of noise generator. An aerial which receives radio signals and passes them on to a receiver is equivalent to a generator having a corresponding internal impedance. Its matching with the receiver means that the sum of the reactive components of the impedance of the aerial and the receiver is equal to zero, while their resistances are equal (see Sec. 49). The maximum power which a generator (aerial) can transmit to the receiver [see (59.35)] is

$$P_{max} = \langle U^2\rangle/(4R), \tag{67.23}$$

where $\langle U^2\rangle$ is the mean square value of the e.m.f. of the aerial and R is its internal resistance, equal to the load resistance.

Suppose that the load resistance R does not produce any noise by itself and is, say, an ohmic resistance maintained at 0 K. We can also imagine the load in the form of an ideal receiver which does not have any internal noise. Nevertheless, the signal received from the aerial will contain noise whose power, in accordance with (67.23) and (67.21), is

$$dP_n = \frac{d\langle U_\nu^2\rangle}{4R} = kT\, d\nu. \tag{67.24}$$

Upon a sufficiently large amplification, this noise will be audible in the headphones and cannot be eliminated by any modifications of the receiver. It can also be seen on the screen of an oscilloscope. An increase in the amplification factor of the receiver will proportionally amplify the legitimate signal at the output and the noise (67.24) supplied to the input, without altering their ratio.

Maximum sensitivity. A signal can be detected if its power is higher than that of the noise. Hence, we obtain the following expression from (67.24) for the

www.EngineeringBooksPDF.com

minimum power of a detectable signal:

$$dP_0 = kT\,d\nu, \tag{67.25}$$

This relation is valid for an ideal receiver, and the power dP_0 represents the sensitivity threshold of the receiver.

The only way to increase the sensitivity (at a fixed temperature) is to decrease the bandwidth $d\nu$ of the frequencies being used. This, however, reduces the information carried by the signal, and in every case there is a limit to which the bandwidth can be reduced. For example, a band of the order $d\nu = 10$ kHz is necessary for transmitting a speech over radio with the help of amplitude modulation without significant distortions. This gives the following value of the minimum detectable signal at room temperature ($T = 290$ K):

$$dP_0 = 1.38 \times 10^{-23} \times 290 \times 10^4 \text{ W} = 4 \times 10^{-17} \text{ W}. \tag{67.26}$$

In order to transmit television pictures, the minimum bandwidth must be of the order of 4 MHz since the information required to reproduce an image is much more than that required to reproduce a speech. In this case, the minimum power of the signal that can be detected by an ideal receiver is 1.6×10^{-14} W.

Equivalent noise temperature of a receiver. In actual practice, a receiver itself is a source of additional noises which are superimposed on the aerial noise. Hence the power dP_1 of the minimum signal that can be detected is higher than dP_0 by an amount dP_r corresponding to the internal noise of the receiver:

$$dP_1 = dP_0 + dP_r. \tag{67.27}$$

The power dP_r of the internal noise of the receiver is usually expressed through formula (67.25) in terms of equivalent noise temperature $T_{eq.n}$ in the following form:

$$dP_r = kT_{eq.n}\,d\nu \tag{67.28}$$

In an ideal receiver, $T_{eq.n} = 0$ K. However, there is no need in practice to approach this limit very closely. It is sufficient to make the equivalent noise temperature about one-tenth of the temperature of the generator (aerial). In this case, additional noise generated by the receiver is practically insignificant.

Noise factor of a receiver. In accordance with (67.26), the power corresponding to the frequency interval $d\nu = 1$ Hz at room temperature is $dP_{01} = 4 \times 10^{-21}$ W. The noise characteristic of a receiver is described by the noise factor

$$F = \frac{dP_1}{dP_{01}} \tag{67.29}$$

This factor is usually expressed in decibels.

Signal-to-noise ratio. The reliability of signal detection depends on the extent to which the signal exceeds the noise level. This is especially important, for example, for the quality of transmission and reception of musical compositions. This characteristic of receivers and reproduction equipment is defined by the

www.EngineeringBooksPDF.com

ratio of the signal voltage to the noise voltage. Since this ratio is usually very large under normal conditions, it is expressed in decibels according to the formula

$$N = 20 \log \frac{U_s}{U_n} = 10 \log \frac{U_s^2}{U_n^2}, \quad (67.30)$$

where U_s and U_n represent the signal and noise voltage respectively.

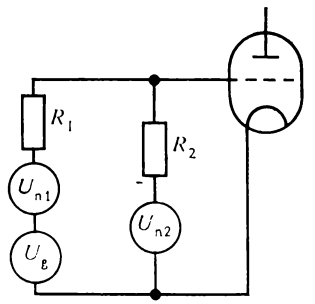


Fig. 267. Calculation of noise on the grid of a triode

As an example, let us consider the signal-to-noise ratio in a triode (Fig. 267). The signal is supplied to the circuit input between the grid and the cathode. The signal generator is characterized by an e.m.f. U_g and an internal resistance R_1. The noise power due to the generator resistance can be written on the basis of (67.21) as follows:

$$4kT\,d\nu = U_{n1}^2/R_1, \quad (67.31)$$

where U_{n1} is the e.m.f. of the equivalent noise generator connected in series with R_1 and the generator U_g.

Another source of noise is the resistance R across which the voltage is measured. The noise power of this source is equal to

$$4kT\,d\nu = U_{n2}^2/R, \quad (67.32)$$

where U_{n2} is the e.m.f. of the equivalent noise generator.

To calculate the noise at the grid, we consider that the resistance R_2 is the load for the noise generator U_{n1}, while R_1 is the load for the noise generator U_{n2}. Obviously, the noise generators operate independently, and hence the mean square value of the total noise voltage is equal to the sum of the mean square voltages of the noises produced by each generator.

Hence, we obtain the expression for the mean square value of noise voltage at the grid.

$$U_n^2 = \left(\frac{U_{n1}}{R_1+R_2} R_2\right)^2 + \left(\frac{U_{n2}}{R_1+R_2} R_1\right)$$
$$= 4kT\,d\nu \left[\frac{R_1R_2^2}{(R_1+R_2)^2} + \frac{R_2R_1^2}{(R_1+R_2)^2}\right] = 4kT\,d\nu \frac{R_1R_2}{R_1+R_2} \quad (67.33)$$

It should be noted that the mean square value of amplitude of the signal at the grid is equal to

$$U_s^2 = \left(\frac{U_g}{R_1+R_2} R_2\right)^2. \quad (67.34)$$

From (67.33) and (67.34), we obtain the ratio of the mean square signal voltage to the mean square noise voltage at the grid:

$$\frac{U_s^2}{U_n^2} = \frac{U_g^2}{4kT\,d\nu} \frac{R_2}{R_1+R_2} \frac{1}{R_1} = \frac{P}{kT\,d\nu} \frac{R_2}{R_1+R_2}, \quad (67.35)$$

www.EngineeringBooksPDF.com

where $P = U_g^2/(4R_1)$ is the maximum signal power supplied by the generator to the external circuit [see (67.23)]. It can be seen from formula (67.35) that *for a perfect matching of the load and the generator* ($R_2 = R_1$), *the signal-to-noise ratio is not the best possible. On the contrary, a mismatching attained by increasing the load resistance* R_2 *can double the signal-to-noise ratio.*

The same conclusion can also be drawn by estimating the sensitivity. The minimum power of the generator signal at the grid which can be distinguished from noise is obtained from (67.35) if $U_s^2/U_n^2 = 1$:

$$P_1 = kT\,d\nu\frac{R_1+R_2}{R_2}. \qquad (67\ 36)$$

Obviously, the minimum detectable power in the case of a matching of the load with the generator ($R_2 = R_1$) is equal to $2kT\,d\nu$, while in the case of mismatching ($R_2 \gg R_1$), this value is equal to $kT\,d\nu$. In other words, the *sensitivity increases when the load is mismatched with the generator.*

If an aerial is the generator in the above case, the above reasoning is valid for the aerial-receiver system as well.

When the load and the generator are perfectly matched, the signal-to-noise ratio is not the best. This ratio can be nearly doubled upon a mismatching of the load and the generator by increasing the load resistance R_2. The same conclusion can be drawn by estimating the sensitivity: a mismatch between the load and the generator, attained by increasing the load resistance R_2, leads to an increased sensitivity.

Sec. 68. Schottky Noise and Current Noise

The physical reasons behind the emergence of Schottky noise are considered and its frequency distribution is analyzed. Basic properties of current noise are discussed.

Source of Schottky noise. Electric current is the motion of discrete elementary charges and not a continuous flow of charge. Hence it gives a sequence of current pulses, each of which is associated with the arrival of an individual electron at the point under consideration. The current through an area element is identical to the passage of shots which are emitted through it from a certain device and have a random distribution in time. It is obvious that the number of shots crossing the surface in identical successive small time intervals will experience a considerable fluctuation. Similarly, the current will also fluctuate in view of the random nature of charges. These fluctuations are called the Schottky noise.

Frequency distribution of noise. The arrival of an electron at a point is equivalent to a current pulse of extremely small duration. If an electron is treated as a point charge, this duration is taken equal to zero and the current pulse is

www.EngineeringBooksPDF.com

assumed to be infinite, i.e. the pulse is treated as a delta-function. Since the charge in the current pulse is equal to the electron charge e, we can express the current associated with the arrival of an electron at the instant t_i in the form

$$i(t) = e\delta(t - t_i). \tag{68.1}$$

Let T be a large interval of time during which N electrons arrive on the average. The mean current generated by an electron arriving during this interval of time is equal to $\langle i \rangle = e/T$, while the mean current due to N electrons is given by the expression $\langle I \rangle = N \langle i \rangle = Ne/T$. However, electrons arrive at random, thus leading to current fluctuations which are responsible for the noise. In order to determine the spectral composition of noise, we express the current $i(t)$ in the form of a Fourier series in the interval $(-T/2, T/2)$:

$$i(t) = a_0/2 + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \quad (\omega = 2\pi/T), \tag{68.2}$$

where

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} i(t) \cos n\omega t \, dt \quad (n = 0, 1, 2, \ldots), \tag{68.3a}$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} i(t) \sin n\omega t \, dt \quad (n = 1, 2, \ldots). \tag{68.3b}$$

Taking into consideration the rule for integration involving a delta-fuction

$$\int f(t)\,\delta(t - t_i)\,dt = f(t_i),$$

and (68.1) we obtain from (68.3a) and (68.3b)

$$a_n = \frac{2e}{T} \cos n\omega t_i, \quad b_n = \frac{2e}{T} \sin n\omega t_i. \tag{68.4}$$

In this case [see (68.2)],

$$i(t) = \frac{e}{T} + \frac{2e}{T} \sum_{n=1}^{\infty} \cos n\omega (t - t_i). \tag{68.5}$$

The mean square value of current for the nth component is equal to

$$\langle i_n^2 \rangle = \frac{4e^2}{T^2} \left\langle \cos^2 \frac{2\pi n}{T} t \right\rangle = \frac{2e^2}{T^2}. \tag{68.6}$$

Since individual electrons move at random and are not mutually correlated, their contributions to the Fourier expansion for current will differ in phase. In calculating the square of current fluctuations, the phase averaging will make all terms with different frequencies vanish, and the series will contain only terms with identical frequencies. Hence, for the mean square fluctuations of

www.EngineeringBooksPDF.com

the nth Fourier component of the current due to N electrons arriving during time T, we obtain

$$\langle T_n^2 \rangle = N \langle i_n^2 \rangle = 2e^2 N/T^2 = 2eI_0/T, \tag{68.7}$$

where $I_0 = eN/T$ is the mean current.

The number of components in the Fourier series whose frequencies lie between ν and $\nu + d\nu$ is equal to $Td\nu$, since successive components are separated from each other on the frequency scale by $1/T$. The interval T can be assumed to be quite large and hence the interval between neighbouring frequencies $[(n + 1)/T] - (n/T) = 1/T$ will be quite small.

Summing the contributions of these components in the frequency interval $d\nu$, we obtain the following expression for the mean square values of current fluctuations on the basis of formula (68.7):

$$d\langle I^2 \rangle = \langle I_n^2 \rangle T\, d\nu = 2eI_0 \nu. \tag{68.8}$$

This formula describes the Schottky noise and is called the **Schottky formula.** If negative values are included in the spectral interval of frequencies ν the factor 2 in formula (68.8) disappears. This approach is usually followed when the exponential form of Fourier series or integrals is used.

Current noise. At very low frequencies, noise is generated due to various inhomogeneities of resistances. The mean square voltage of this noise decreases in inverse proportion to frequency.

An experimental investigation of this noise, called the **current noise**, leads to the following formula

$$\langle (\Delta U)^2 \rangle = \alpha I_n^2/\nu \tag{68.9}$$

where α is an empirical constant which depends on the geometry of the resistor and its material. In large metallic conductors, there is practically no noise. In various types of compound resistors, however, the noise is quite high.

The reason behind this noise is not clear so far. However, its role becomes negligibly small in practically all cases with increasing frequency.

Methods of reducing noise. Noise distorts the shape of a true signal and should therefore be suppressed. The signal-to-noise ratio provides a quantitative measure of the relation between a signal and noise. Our task is to find the ways in which this ratio can be increased.

The amplification of a signal does not lead to the desired result, since the amplifier increases both the signal and the noise supplied at its input to the same extent and, besides, the internal noise of the amplifier is also added to the signal as it passes through the amplifier. Hence the amplification decreases the signal-to-noise ratio, i.e. deteriorates this characteristic and cannot be used as a method for decreasing noise.

Resistance noise can be decreased by decreasing the temperature at which the equipment operates. This method is widely used, but has its own limitations. Firstly, it considerably complicates the operation and, secondly, circuit elements change their electrical properties upon intense cooling, sometimes irreversibly.

www.EngineeringBooksPDF.com

Schottky noise and current noise can be decreased by decreasing the current, while the current noise can also be decreased by increasing the signal frequency. The increase in the signal frequency is limited by the high-frequency characteristics of the circuits and circuit elements.

All kinds of noise are reduced upon decreasing the bandwidth. The bandwidth, however, is limited by the signal properties, since every signal has a finite width and a decrease in the bandwidth below the signal width considerably distorts the signal, i.e. introduces a new noise.

Thus, the signal-to-noise ratio of devices can be improved by improving their technical characteristics, although this approach is fraught with its own

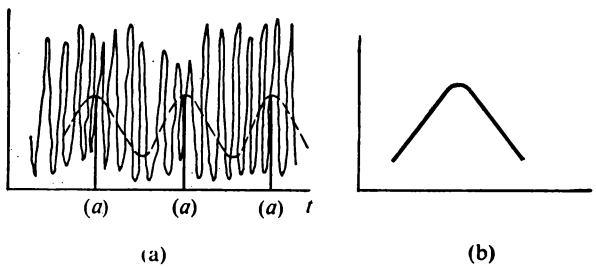


Fig. 268. Isolation of a signal from a strong background noise

limitations. Hence methods have been worked out to receive the signal in such a way as to overcome these limitations. One of the most frequently used methods consists in the following.

Suppose that we have a periodically repeating signal which is strongly distorted by the background noise (Fig. 268*a*). The signal period can be determined quite accurately, since noise does not distort the period. We can then synchronize the moment of measurement of the signal with the periodicity of its variation, i.e. to measure the signal several times at the same point of its period, say, at point *a* in Fig. 268*a*. Due to superposition of the noise, each measurement gives a different result, but the mean value of a large number of measurements yields the value of the signal at this point with an appropriate degree of accuracy. In principle this accuracy can be improved indefinitely by correspondingly increasing the number of measurements. Carrying out such measurements for different points of the period, we obtain the shape of the signal for one period without any noise distortions (Fig. 268*b*).

www.EngineeringBooksPDF.com

Appendices

I. SI Units Used in This Book

Quantity			Unit	
name	notation	dimensions	name	notation
Basic Units				
Length	l	L	metre	m
Mass	m	M	kilogram	kg
Time	t	T	second	s
Current	I	I	ampere	A
Temperature	T	Θ	kelvin	K
Amount of substance	ν	N	mole	mole
Luminous intensity	I	J	candela	cd
Derived Units				
Velocity	v, u	LT^{-1}	metre per second	m/s
Acceleration	a	LT^{-2}	metre per second per second	m/s^2
Force	F	LMT^{-2}	newton	N
Pressure	p	$L^{-1}MT^{-2}$	pascal	Pa
Momentum	p	LMT^{-1}	kilogram-metre per second	kg·m/s
Energy	W, U, E	L^2MT^{-2}	joule	J
Power	P	L^2MT^{-3}	watt	W
Moment of inertia	J	L^2M	kilogram-metre-squared	$kg \cdot m^2$
Moment of force	M	L^2MT^{-2}	newton-metre	N·m
Angular momentum	L	L^2MT^{-1}	kilogram-metre squared per second	$kg \cdot m^2/s$
Electric charge	Q, q	TI	coulomb	C
Volume charge density	ρ	$L^{-3}TI$	coulomb per cubic metre	C/m^3
Surface charge density	σ	$L^{-2}TI$	coulomb per metre squared	C/m^2
Linear charge density	τ	$L^{-1}TI$	coulomb per metre	C/m
Absolute permittivity	ε	$L^{-3}M^{-1}T^4I^2$	farad per metre	F/m
Dielectric constant	ε_0	$L^{-3}M^{-1}T^4I^2$	farad per metre	F/m
Relative permittivity	ε_r		dimensionless quantity	
Electric field strength	E	$LMT^{-3}I^{-1}$	volt per metre	V/m
Flux of electric field	N	$L^3MT^{-3}I^{-1}$	volt-metre	V·m
Electric potential	φ	$L^2MT^{-3}I^{-1}$	volt	V
Dipole electric moment	p	LTI	coulomb-metre	C·m
Polarization	P	$L^{-2}TI$	coulomb per metre squared	C/m^2
Electric displacement	D	$L^{-2}TI$	coulomb per metre squared	C/m^2
Displacement flux (electric flux)	Ψ	TI	coulomb	C
Electric capacitance	C	$L^{-2}M^{-1}T^4I^2$	farad	F

www.EngineeringBooksPDF.com

(*Continued*)

Quantity			Unit	
name	notation	dimensions	name	notation
Volume energy density of electric and magnetic fields	w	$L^{-1}MT^{-2}$	joule per cubic metre	J/m^3
Voltage	U	$L^2MT^{-3}I^{-1}$	volt	V
Electric resistance	R	$L^2MT^{-3}I^{-2}$	ohm	Ω
Mobility of charge carriers	b	$M^{-1}T^2I$	square metre per volt-second	$m^2/V \cdot s$
Volume current density	j	$L^{-2}I$	ampere per metre squared	A/m^2
Magnetic moment of current	p_m	L^2I	ampere-metre squared	$A \cdot m^2$
Magnetic induction	B	$MT^{-2}I^{-1}$	tesla	T
Magnetic flux	Φ	$L^2MT^{-2}I^{-1}$	weber	Wb
Magnetic field strength	H	$L^{-1}I$	ampere per metre	A/m
Inductance	L	$L^2MT^{-2}I^{-2}$	henry	H
Absolute magnetic permeability	μ	$LMT^{-2}I^{-2}$	henry per metre	H/m
Magnetic constant	μ_0	$LMT^{-2}I^{-2}$	henry per metre	H/m
Relative magnetic permeability	μ_r		dimensionless quantity	
Magnetization	J	$L^{-1}I$	ampere per metre	A/m
Oscillation frequency	ν	T^{-1}	hertz	Hz
Cyclic frequency of oscillations	ω	T^{-1}	second inverse	s^{-1}
Electromagnetic field energy flux density	S	MT^{-3}	watt per metre squared	W/m^2

II. Relation between Formulas in SI and Gaussian System of Units

Although SI system of units is introduced almost everywhere, sometimes a transition from one system of units to another is still required. This table is used for the conversion of formulas from SI to the Gaussian system of units

Quantity	SI	Gaussian system	Quantity	SI	Gaussian system
Current	I	$(4\pi\varepsilon_0)^{1/2}I$	Electric field strength	E	$(4\pi\varepsilon_0)^{-1/2}E$
Current density	j	$(4\pi\varepsilon_0)^{1/2}j$	Electric displacement	D	$(\varepsilon_0/4\pi)^{1/2}D$
Electric charge	Q	$(4\pi\varepsilon_0)^{1/2}Q$	Magnetic field strength	H	$(4\pi\mu_0)^{-1/2}H$
Charge density	ρ	$(4\pi\varepsilon_0)^{1/2}\rho$	Magnetic induction	B	$[\mu_0/(4\pi)]^{1/2}B$
Conductance	γ	$4\pi\varepsilon_0\gamma$			
Capacitance	C	$4\pi\varepsilon_0 C$			

www.EngineeringBooksPDF.com

Continued

Quantity	SI	Gaussian system	Quantity	SI	Gaussian system
Magnetic flux	Φ	$[\mu_0/4\pi)]^{1/2}\Phi$	Velocity of light	c	$(\mu_0/\varepsilon_0)^{-1/2}$
Inductance	L	$(4\pi\varepsilon_0)^{-1}L$	Magnetic susceptibility	χ	$4\pi\chi$
Polarization	P	$4\pi\varepsilon_0 P$	Dielectric susceptibility	$\varkappa$	$4\pi\varkappa$
Magnetization	J	$(4\pi/\mu_0)^{1/2}J$	Permittivity	ε	$\varepsilon\varepsilon_0$
Electric resistance	R	$(4\pi\varepsilon_0)^{-1}R$	Permeability	μ	$\mu\mu_0$
Electric dipole moment	p	$(4\pi\varepsilon_0)^{1/2}p$	Relative permittivity	ε_r	$\varepsilon/\varepsilon_0$
Magnetic moment of current	p_m	$(4\pi/\mu_0)^{1/2}p_m$	Relative permeability	μ_r	μ/μ_0
Scalar potential	φ	$(4\pi\varepsilon_0)^{-1/2}\varphi$			
Vector potential	A	$[\mu_0/(4\pi)]^{1/2}A$			

How to use the table. In order to convert a relation written in SI into the corresponding formula in the Gaussian system, each symbol from the "SI" column should be replaced by the symbol from the column "Gaussian system". Using this rule in the reverse order, we can convert formulas written in the Gaussian system into those in SI. Upon these transitions, mechanical and other nonelectric and nonmagnetic quantities remain unchanged, as well as the derivatives with respect to coordinates and time.

Examples illustrating the use of the table.

1. Write the Maxwell equation

$$\text{curl } \mathbf{H} = \mathbf{j} + \partial\mathbf{D}/\partial t \quad \text{(SI)}$$

in the Gaussian system. We have

$$\text{curl}\,[(4\pi\mu_0)^{-1/2}\mathbf{H}] = (4\pi\varepsilon_0)^{1/2}\mathbf{j} + \frac{\partial}{\partial t}\left[\left(\frac{\varepsilon_0}{4\pi}\right)^{1/2}\mathbf{D}\right],$$

i.e.

$$\text{curl } \mathbf{H} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}$$

2. Write Poynting's vector

$$\mathbf{S} = [c/(4\pi)]\ \mathbf{E}\times\mathbf{H} \quad \text{(Gaussian system)}$$

in SI. We have

$$\mathbf{S} = \frac{(\mu_0\varepsilon_0)^{-1/2}}{4\pi}\,[(4\pi\varepsilon_0)^{1/2}\mathbf{E}\times(4\pi\mu_0)^{1/2}\mathbf{H}] = \mathbf{E}\times\mathbf{H}.$$

Remark. A transition from SI to the Gaussian system always leads to correct results. In the reverse transition (from the Gaussian system to SI), errors are possible if the formula in the Gaussian system is written for vacuum. In this case, $D = E$, $B = H$, and one of the quantities in the formula may turn out to be replaced by the other quantity, and the conversion factors for these quantities are different. Therefore, before converting a formula from the Gaussian system into SI, we should take care and write it so that it is valid for a medium as well as for vacuum.

www.EngineeringBooksPDF.com

The conversion of numerical values of quantities from one system of units to another is made with the help of tables contained in the books on the systems of units.

III. Formulas of Vector Algebra and Calculus

1. The property of the scalar triple product of vectors:

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = (\mathbf{A}\times\mathbf{B})\cdot\mathbf{C}, \tag{A.1}$$

2. Decomposition of the vector triple product:

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}) \tag{A.2}$$

3. The definition of the vector nabla operator:

$$\nabla = \mathbf{i}_x\frac{\partial}{\partial x} + \mathbf{i}_y\frac{\partial}{\partial y} + \mathbf{i}_z\frac{\partial}{\partial z} \tag{A.3}$$

where $\mathbf{i}_x$, $\mathbf{i}_y$, $\mathbf{i}_z$ ar the unit vectors of the Cartesian system of coordinates.

4. The definition of the gradient operation:

$$\operatorname{grad}\varphi = \nabla\varphi. \tag{A.4}$$

5. The definition of the divergence operation:

$$\operatorname{div}\mathbf{A} = \nabla\cdot\mathbf{A}. \tag{A.5}$$

6. The definition of the curl operation:

$$\operatorname{curl}\mathbf{A} = \nabla\times\mathbf{A}. \tag{A.6}$$

7. Vector identities:

$$\nabla\cdot\nabla_\varphi = \nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} \tag{A.7}$$

$$\nabla\times\nabla_\varphi = 0, \tag{A.8}$$

$$\nabla\cdot(\nabla\times\mathbf{A}) = 0, \tag{A.9}$$

$$\nabla\times(\nabla\times\mathbf{A}) = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}, \tag{A.10}$$

$$\nabla(\varphi\psi) = \varphi\nabla\psi + \psi\nabla\varphi \tag{A.11}$$

$$\nabla\cdot(\varphi\mathbf{A}) = \varphi(\nabla\cdot\mathbf{A}) + \mathbf{A}\cdot\nabla\varphi \tag{A.12}$$

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{A} + (\mathbf{A}\cdot\nabla)\mathbf{B} + \mathbf{B}\times(\nabla\times\mathbf{A}) + \mathbf{A}\times(\nabla\times\mathbf{B}), \tag{A.13}$$

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = \mathbf{B}(\nabla\cdot\mathbf{A}) + \mathbf{A}(\nabla\cdot\mathbf{B}) + (\mathbf{B}\times\nabla)\times\mathbf{A} + (\mathbf{A}\times\nabla)\times\mathbf{B}, \tag{A.14}$$

$$\nabla\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\nabla\times\mathbf{A}) - \mathbf{A}\cdot(\nabla\times\mathbf{B}), \tag{A.15}$$

$$\nabla\times(\varphi\mathbf{A}) = \varphi(\nabla\times\mathbf{A}) + (\nabla\varphi)\times\mathbf{A}, \tag{A.16}$$

$$\nabla\times(\mathbf{A}\times\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{A} - (\mathbf{A}\cdot\nabla)\mathbf{B} + \mathbf{A}(\nabla\cdot\mathbf{B}) - \mathbf{B}(\nabla\cdot\mathbf{A}) \tag{A.17}$$

8. Gauss' theorems: the closed surface S envelopes a volume V. The vector dS of the surface area element is directed along the outward normal to the surface:

$$\int_V (\nabla\cdot\mathbf{A})\,dV = \oint_S \mathbf{A}\cdot d\mathbf{S}, \tag{A.18}$$

$$\int_V (\nabla\varphi)\,dV = \oint_S \varphi\,d\mathbf{S}, \tag{A.19}$$

$$\int_V (\nabla\times\mathbf{A})\,dV = \oint_S d\mathbf{S}\times\mathbf{A}. \tag{A.20}$$

www.EngineeringBooksPDF.com

9. Stokes' theorems: the closed contour L bounds a surface S. The vector $\mathbf{dl}$ of the element of the contour L coincides in direction with the positive circumvention which is connected with the direction of the positive normal to the surface S through the right-hand rule:

$$\int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = \oint_L \mathbf{A} \cdot d\mathbf{l}, \tag{A.21}$$

$$\int_S d\mathbf{S} \times \nabla \varphi = \oint_L \varphi \, d\mathbf{l}, \tag{A.22}$$

$$\int_S (d\mathbf{S} \times \nabla) \times \mathbf{A} = \oint_L d\mathbf{l} \times \mathbf{A}. \tag{A.23}$$

10. Green's theorems:

$$\int_V (\varphi \nabla^2 \psi - \psi \nabla^2 \varphi) \, dV = \oint_S (\varphi \nabla \psi - \psi \nabla \varphi) \cdot d\mathbf{S}, \tag{A.24}$$

$$\int_V (\nabla \varphi \times \nabla \psi) \, dV = \frac{1}{2} \oint_S d\mathbf{S} \times (\varphi \nabla \psi - \psi \nabla \varphi), \tag{A.25}$$

$$\int_S (\nabla \varphi \times \nabla \psi) \cdot d\mathbf{S} = \frac{1}{2} \oint_L (\varphi \nabla \psi - \psi \nabla \varphi) \cdot d\mathbf{l}. \tag{A.26}$$

www.EngineeringBooksPDF.com

Subject Index

Accumulator, 201
Antiferroelectric(s), 186
Antiferromagnetic(s), 186
Antiferromagnetism, 292f
Autotransformer, 352

Band, conduction, 220
Betatron, 369ff
  stability, 371ff
Bohr magneton, 299
Boundary conditions for field vectors, 263f

Capacitance, 112, 118f, 127, 140, 152, 168, 170, 213, 333
Capacitor(s), 118, 183, 419
  cylindrical, 119, 170
    two-layer, 155
  parallel-plate, 119, 140, 143, 163, 167, 169f, 183, 377
  spherical, 119
Cell(s),
  Daniell, 200
  galvanic, 197
  polarization of, 200
  voltaic, 198
Charge(s),
  bound, 135f, 138, 147, 173
    density of, 135
  mobility of, 231
  polarization, 135
  space, 107, 193
  surface, 193
Charge concentration, 32
Charge density,
  surface, 32, 106, 129
  volume, 31, 121, 238
Charge invariance, 81
Circuit,
  branched, 207
  linear, 206
  LCR, 327, 345
  mutual inductance, 347
  oscillatory, 346
  quasistationary a.c., 323ff
  RC, 327
  RL, 325
Coefficient(s),
  capacitance, 115ff, 127
  dissociation, 226f
  potential, 113f
  transverse magnetoresistance, 222
Condition(s),
  betatron, 370
  boundary, 140f, 143
Constant,
  Avogadro, 176
  dielectric, 130
  Hall, 221
  Planck, 20
  propagation, 389
Crystals, ionic, 182
Current(s),
  alternating quasistationary, 300ff
  density of, 215
  direct, 191ff
  displacement, 376ff
  elementary, 249f
  Foucault (eddy), 343
  molecular,
    surface density of, 259
    volume density of, 256f
  non-self-sustained, 228, 230
  saturation, 233, 238
    density of, 230, 235
  self-sustained, 228, 230
  single-phase, 353
  three-phase, 353
  unit of, 57
  volume, 62
Curie point, 184
  lower, 184
  upper, 184, 187
Curl of vector, 83f

Damping factor, 347
Diamagnetics, 256, 273f, 277ff
Diamagnetic susceptibility, 281
Diamagnetism, 280
Dielectric(s), 101, 172
  anisotropic, 134
  linear, 134
  nonpolar, 175ff
  polar, 177ff, 180
Dielectric domains, 185
Dielectric polarization, 132
Dielectric susceptibility, 134, 175, 184
  tensor of, 185
Dipole, 123,
  arm of, 123
  field of, 123
Dipole moment, 123, 129f, 132, 153, 158, 170, 174, 182, 185
  induced, 182
  of molecules, 175, 178f, 181
  permanent, 182
Distribution, Boltzmann, 178
Divergence, 36
Domain, 291
  boundaries of, 292

Effect,
  Barnett, 298
  galvanomagnetic, 222
  gyromagnetic, 295,
  Hall, 220ff
    anomalous, 222
  Meissner, 224
  photoelectric, 23
  piezoelectric, 188
    inverse, 188f
    longitudinal, 188
    transverse, 188
  pyroelectric, 189
    direct, 189
    inverse, 189
  skin, 357ff, 411f
    anomalous, 360f
Electric charge, 15ff
Electric conductivity, 101, 218ff
  of gases, 228ff
  of liquids, 225
    calculation of, 226
Electric current, 30, 34
  in vacuum, 232ff

www.EngineeringBooksPDF.com

Electric displacement, 139, 144, 147
Electric field strength, 121, 129, 131
Electric motor, 337
  asynchronous d.c., 340f
  synchronous, 339, 341
Electrolyte(s), 228
Electromagnetic induction, 300ff
Electrometer, 108
Electron(s), 15ff
  conduction, 285
  energy spectrum of, 20
  free, 285
Electroscope, 108
Electrostatic charging, 25
Electrostatic induction, 103, 115, 125, 130
Elementary charge, 26, 28f
Energy,
  of capacitor, 168
  of charged conductors, 152
    interaction, 149, 151f, 155
    intrinsic, 151, 155
    kinetic, 151, 155
    potential, 151, 155, 158
  of dielectric in external field, 153
  of dipole in external field, 152
  of electrostatic field, 148ff
  exchange interaction, 289
  Fermi, 22f
  free, 162, 164
  of magnetic dipole, 321
  of magnetic in external field, 316f
  magnetic, density of, 313
  of magnetic field, 309f
    in presence of magnetic, 313
  total, 153, 155
Energy density, 150f
Energy flux, 383f
  density of, 409
Energy levels, 20f
Equation(s),
  continuity, 41, 143, 212
  D'Alembert, 393
  for field vectors of electromagnetic wave, 406
  Laplace, 96, 143f
  Lorentz, 393
  Maxwell, 47, 80, 245ff, 377ff, 381ff
  Newton, 277
  Poisson, 96f, 144, 237
  Richardson-Dushman, 235
  for scalar potential, 393
  for vector potential, 392
  transmission line, 389
  wave, solution of, 394
Experiment,
  Einstein-de Hass, 296f
  Millikan oil-drop, 26
  Tolman and Stewart, 218

Fermi level, 23
Ferrimagnetic(s) (ferrites), 293
Ferroelectric(s), 134, 187
Ferromagnetic(s), 259, 287ff
Ferromagnetism, 293
Field,
  inside conductor, 191
  outside conductor, 193
  Coulomb, 83, 94, 123
  critical, 223
  electric, 48
    constant, 76ff
    strength of, 48
  electrostatic, 82
    in dielectric, 129
    in presence of conductors, 100
    in vacuum, 94ff
Field,
  external, 172, 176, 181
  Hall, 222
  local, 172, 176, 181
  magnetic, 53ff, 426
    in presence of magnetics, 254ff
    of rectilinear current, 62
    relativistic nature of, 56
    stationary, 240ff
  in magnetics, 261
  saturation, 179, 181
  of solenoid, 315
Field lines, 81
Filter(s), 366ff
  high-pass, 368f
  iterative, 368
  low-pass, 366ff
Fluctuation(s), 427
  in oscillatory circuit, 430
    frequency distribution of, 431
Forbidden bandwidth, 22
Force(s),
  acting on conductor, 160
  Ampère, 58ff, 61, 161, 340
  body, 272
    acting on compressible magnetic, 320
  coercive, 288
  elastic, body (volume), 161
  electromotive, 191
  in electrostatic field, 156ff
  extraneous, 194ff
  of interaction
    between current-carrying conductors, 56
    between rectilinear currents, 66
    mechanical, 195
  long-range, 182
  Lorentz, 58ff, 61, 69, 338f
  magnetic, 56
  in magnetic field, 270ff
  surface, 161
  surface density of, 164
  volume, 157, 165
Formula,
  Clausius-Mosotti, 177
  Stokes, 27
  Thomson, 346
Frequency,
  cut-off, 422
  Larmor, 278
  resonance, 344
Function, Langevin, 179

Gain factor, 365
Generator(s), 354ff
  a.c., 302
  equivalent noise, 433
Gradient, 87

Hall angle, 222
Helmholtz ring, 68
Hysteresis loop, 183f

Impedance, 329, 364f
  characteristic, 365, 389f
  input, 364
  output, 364
Inductance, 310, 314, 321
  mutual, 311, 347
Induction coil, 419

www.EngineeringBooksPDF.com

Interaction,
Coulomb, 91
long-range, 47
short-range, 47, 130
spin-orbital, 283

Joule heat, 197, 228, 337, 385

Larmor precession, 279ff
Law(s),
Ampère, 61, 64
Ampère circuital, 240, 323
differential form, 242f
experimental verification of, 244
integral form, 240
Biot-Savart, 63ff, 66, 249
of conservation, 303
of angular momentum, 297
of charge, 35, 103, 193, 199
differential form, 41
integral form, 36
of electromagnetic energy, 383ff
of energy, 151, 155, 166
of mass, 165
Coulomb, 42ff, 50, 64, 72, 81, 94, 98, 194, 381
differential form, 77ff, 80
field form, 47
Curie, 283
Curie-Weiss, 185, 283, 291
Faraday, 340
of electrolysis, 227
of electromagnetic induction, 195, 304ff
differential form, 306f
Joule, differential form of, 202ff
Kirchhoff, 206, 208ff, 330ff
Lenz, 325, 340
Newton,
gravitational, 46
second, 203
third, 65
Ohm, 100ff, 127, 202ff, 216, 222, 227, 340
in differential form, 190, 195
in local form, 191
three-halves power, 236, 238

Magnet(s), permanent, 262
Magnetic dipole, 270ff
Magnetic field strength, 260
Magnetic induction, 61, 67
measurement of, 264
Magnetic moment, 118, 251, 284
Magnetic shielding, 268
Magnetics, 277
Magnetization, 255f, 260, 282
Magnetization curve, 287
Magnetoresistance, 222
Method(s),
Cavendish, 43, 45
depolarization, 200
of image charges, 123ff, 142, 170
mesh-current, 331
resonance, 27
Model, Onsager, 182
Molecules,
nonpolar, 132
polar, 132
Moment of force, 157, 169
Momentum of electromagnetic wave train, 416

Network, four-terminal, 361ff
Neutron, 17f
Noise(s), 429ff
current, 436, 438
resistance, 432
Schottky, 436
Noise factor, 434

Oscillator, Hertzian, 396

Paramagnetic(s), 255, 273f, 282
Permeability, 266, 273
Permittivity, 139, 142, 147, 153f, 158ff
complex, 409
relative, 139, 177, 180ff, 184f, 189
Photon, momentum of, 415, 417
Piezoelectric(s), 187ff
polar axis of, 187
transducers, 188
Polarization, 146, 158, 166, 172f, 175, 180ff, 185 188
of dielectric, 129, 131ff
dielectric, 137f
ionic lattice, 133
residual, 184
spontaneous, 184
temperature dependence of, 177
Potential, 144f, 148f, 152, 164, 166, 194
advanced, 395f
calculation of, 94
of conductor, 111
electrochemical, 198
absolute, 198
of field of point charge, 89
of field of surface charges, 90
normalization of, 88
retarded, 395
scalar, 88, 397
of system of point charges, 89
thermodynamic, 161
vector, 247ff, 252, 256
gauging of, 248
potential difference (drop), 199, 209, 212f, 221
contact, 23f
Power of alternating current, 334
mean value, 336
Power factor, 337
Poynting, vector, 384
Principle of superposition, 50, 81
field form, 51f
Process, isothermal, 161
Proton, 15ff
Pyroelectrics, 189

Q-factor, 297, 347
Quark(s), 28

Radiation, electromagnetic wave, 392
Ratio,
gyromagnetic, 295
signal-to-noise, 434
Resistance, 213
characteristic, 390
internal, 199
ohmic, 195, 198f, 202
temperature dependence of, 220
Resonance,
in a.c. circuit, 343ff

www.EngineeringBooksPDF.com

Resonance
current, 344ff
ferromagnetic, 294
paramagnetic, 279, 285
voltage, 343ff
Resonator(s), 418, 427

Schumann resonances, 46
Self-energy, 149, 155
of point charge, 150
Self-inductance, 315, 325, 347
Semiconductor(s), 101
Skin depth, 359
Spin, 18
Stability,
radial, 371
vertical, 372
Statistics,
Bose-Einstein, 224f
Fermi-Dirac, 20, 224
Step voltage, 215
Superconductivity, 223
Superconductor(s), 224
hard, 224
soft, 224
Surface, equipotential, 124ff
Susceptibility,
diamagnetic, 277, 280
temperature dependence of, 281
dielectric,
atomic, 177
molecular, 177

Temperature,
critical, 223
Curie-Weiss, 184, 187, 291
Nèel, 293
Theorem,
Earnshaw, 82, 91
of equipartition of energy, 429
Gauss, 39ff, 77, 79, 81f, 91, 95, 98, 103f, 110, 121, 137, 139f, 147, 150, 155, 165, 210
reciprocity, 114
Stokes integral, 86f, 92
uniqueness, 124, 143

Theory,
band, 219
of electric conductivity, 180
classical, 205
quantum, 180, 205, 219
of ferromagnetism, 289
potential, 83
of superconductivity, 225
Thermionic emission, 232
Thermionic work function, 19, 23
Time,
mean free, 206
relaxation, 102
Transformation,
field, 413
invariants, 414
gauge, 308f.
Lorentz, 72
Transformer,
real, 348, 352
vector diagram, 350ff
Transmission line, 337, 385ff
Triboelectrification, 25

Vector diagram, 329
Velocity,
drift, 226f
group, 426
phase, 407, 412, 423, 426

Wave(s),
electromagnetic, 376ff
in conducting media, 409
in dielectric, 405
pressure of, 415
plane, invariance of, 413ff
Waveguides, 418ff, 421
regular, 421
Wavelength, 407
boundary, 425
in waveguide, 425
Wave number, 407
Wimshurst machine, 196
Work,
of alternating current, 334ff
in electric field, 82ff
in terms of potential, 89

To the Reader

Mir Publishers would be grateful for your comments on the content, translation and design of this book. We would also be pleased to receive any other suggestions you may wish to make.
Our address is:
Mir Publishers
2 Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR

Printed in the Union of Soviet Socialist Republics

www.EngineeringBooksPDF.com