SYSTEM DYNAMICS

System Dynamics: Modeling, Simulation, and Control of Mechatronic Systems, Fifth Edition
Dean C. Karnopp, Donald L. Margolis and Ronald C. Rosenberg
Copyright © 2012 John Wiley & Sons, Inc.

SYSTEM DYNAMICS

Modeling, Simulation, and Control of Mechatronic Systems

Fifth Edition

DEAN C. KARNOPP
Department of Mechanical and Aerospace Engineering
University of California
Davis, California

DONALD L. MARGOLIS
Department of Mechanical and Aerospace Engineering
University of California
Davis, California

RONALD C. ROSENBERG
Department of Mechanical Engineering
Michigan State University
East Lansing, Michigan



WILEY

JOHN WILEY & SONS, INC.

This book is printed on acid-free paper. ♾

Copyright © 2012 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and the author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor the author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information about our other products and services, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at http://booksupport.wiley.com. For more information about Wiley products, visit www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Karnopp, Dean.
System dynamics : modeling, simulation, and control of mechatronic systems / Dean C. Karnopp, Donald L. Margolis, Ronald C. Rosenberg. – 5th ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-88908-4 (cloth); ISBN 978-1-118-15281-2 (ebk); ISBN 978-1-118-15282-9 (ebk); ISBN 978-1-118-15283-6 (ebk); ISBN 978-1-118-15982-8 (ebk); ISBN 978-1-118-16007-7 (ebk); ISBN 978-1-118-16008-4 (ebk)
1. Systems engineering. 2. System analysis. 3. Bond graphs. 4. Mechatronics. I. Margolis, Donald L. II. Rosenberg, Ronald C. III. Title.
TA168.K362 2012
620.001′1–dc23

2011026141

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTENTS

Preface **xi**

1 Introduction **1**

1.1 Models of Systems, 4
1.2 Systems, Subsystems, and Components, 7
1.3 State-Determined Systems, 9
1.4 Uses of Dynamic Models, 10
1.5 Linear and Nonlinear Systems, 11
1.6 Automated Simulation, 12
References, 13
Problems, 14

2 Multiport Systems and Bond Graphs **17**

2.1 Engineering Multiports, 17
2.2 Ports, Bonds, and Power, 24
2.3 Bond Graphs, 27
2.4 Inputs, Outputs, and Signals, 30
Problems, 33

3 Basic Bond Graph Elements **37**

3.1 Basic 1-Port Elements, 37
3.2 Basic 2-Port Elements, 50
3.3 The 3-Port Junction Elements, 57
3.4 Causality Considerations for the Basic Elements, 63

3.4.1 Causality for Basic 1-Ports, 64
3.4.2 Causality for Basic 2-Ports, 65
3.4.3 Causality for Basic 3-Ports, 66
3.5 Causality and Block Diagrams, 67
Reference, 71
Problems, 71

4 System Models **77**

4.1 Electrical Systems, 78
4.1.1 Electrical Circuits, 78
4.1.2 Electrical Networks, 84
4.2 Mechanical Systems, 91
4.2.1 Mechanics of Translation, 91
4.2.2 Fixed-Axis Rotation, 100
4.2.3 Plane Motion, 106
4.3 Hydraulic and Acoustic Circuits, 121
4.3.1 Fluid Resistance, 122
4.3.2 Fluid Capacitance, 125
4.3.3 Fluid Inertia, 130
4.3.4 Fluid Circuit Construction, 132
4.3.5 An Acoustic Circuit Example, 135
4.4 Transducers and Multi-Energy-Domain Models, 136
4.4.1 Transformer Transducers, 137
4.4.2 Gyrator Transducers, 139
4.4.3 Multi-Energy-Domain Models, 142
References, 144
Problems, 144

5 State-Space Equations and Automated Simulation **162**

5.1 Standard Form for System Equations, 165
5.2 Augmenting the Bond Graph, 168
5.3 Basic Formulation and Reduction, 175
5.4 Extended Formulation Methods—Algebraic Loops, 183
5.4.1 Extended Formulation Methods—Derivative Causality, 188
5.5 Output Variable Formulation, 196
5.6 Nonlinear and Automated Simulation, 198
5.6.1 Nonlinear Simulation, 198
5.6.2 Automated Simulation, 202
Reference, 207
Problems, 207

6 Analysis and Control of Linear Systems 218

6.1 Introduction, 218
6.2 Solution Techniques for Ordinary Differential Equations, 219
6.3 Free Response and Eigenvalues, 222
6.3.1 A First-Order Example, 223
6.3.2 Second-Order Systems, 225
6.3.3 Example: The Undamped Oscillator, 230
6.3.4 Example: The Damped Oscillator, 232
6.3.5 The General Case, 236
6.4 Transfer Functions, 239
6.4.1 The General Case for Transfer Functions, 241
6.5 Frequency Response, 244
6.5.1 Example Transfer Functions and Frequency Responses, 249
6.5.2 Block Diagrams, 255
6.6 Introduction to Automatic Control, 258
6.6.1 Basic Control Actions, 259
6.6.2 Root Locus Concept, 273
6.6.3 General Control Considerations, 285
6.7 Summary, 310
References, 311
Problems, 311

7 Multiport Fields and Junction Structures 326

7.1 Energy-Storing Fields, 327
7.1.1 *C*-Fields, 327
7.1.2 Causal Considerations for *C*-Fields, 333
7.1.3 *I*-Fields, 340
7.1.4 Mixed Energy-Storing Fields, 348
7.2 Resistive Fields, 350
7.3 Modulated 2-Port Elements, 354
7.4 Junction Structures, 357
7.5 Multiport Transformers, 359
References, 364
Problems, 365

8 Transducers, Amplifiers, and Instruments 371

8.1 Power Transducers, 372
8.2 Energy-Storing Transducers, 380

8.3 Amplifiers and Instruments, 385
8.4 Bond Graphs and Block Diagrams for Controlled Systems, 392
References, 397
Problems, 397

9 Mechanical Systems with Nonlinear Geometry 411

9.1 Multidimensional Dynamics, 412
9.1.1 Coordinate Transformations, 416
9.2 Kinematic Nonlinearities in Mechanical Dynamics, 420
9.2.1 The Basic Modeling Procedure, 422
9.2.2 Multibody Systems, 433
9.2.3 Lagrangian or Hamiltonian *IC*-Field Representations, 440
9.3 Application to Vehicle Dynamics, 445
9.4 Summary, 452
References, 452
Problems, 453

10 Distributed-Parameter Systems 470

10.1 Simple Lumping Techniques for Distributed Systems, 471
10.1.1 Longitudinal Motions of a Bar, 471
10.1.2 Transverse Beam Motion, 476
10.2 Lumped Models of Continua through Separation of Variables, 482
10.2.1 The Bar Revisited, 483
10.2.2 Bernoulli–Euler Beam Revisited, 491
10.3 General Considerations of Finite-Mode Bond Graphs, 499
10.3.1 How Many Modes Should Be Retained?, 499
10.3.2 How to Include Damping, 503
10.3.3 Causality Consideration for Modal Bond Graphs, 503
10.4 Assembling Overall System Models, 508
10.5 Summary, 512
References, 512
Problems, 512

11 Magnetic Circuits and Devices 519

11.1 Magnetic Effort and Flow Variables, 519
11.2 Magnetic Energy Storage and Loss, 524
11.3 Magnetic Circuit Elements, 528
11.4 Magnetomechanical Elements, 532
11.5 Device Models, 534
References, 543
Problems, 544

12 Thermofluid Systems 548

12.1 Pseudo-Bond Graphs for Heat Transfer, 548
12.2 Basic Thermodynamics in True Bond Graph Form, 551
12.3 True Bond Graphs for Heat Transfer, 558
12.3.1 A Simple Example of a True Bond Graph Model, 561
12.3.2 An Electrothermal Resistor, 563
12.4 Fluid Dynamic Systems Revisited, 565
12.4.1 One-Dimensional Incompressible Flow, 569
12.4.2 Representation of Compressibility Effects in True Bond Graphs, 573
12.4.3 Inertial and Compressibility Effects in One-Dimensional Flow, 576
12.5 Pseudo-Bond Graphs for Compressible Gas Dynamics, 578
12.5.1 The Thermodynamic Accumulator—A Pseudo-Bond Graph Element, 579
12.5.2 The Thermodynamic Restrictor—A Pseudo-Bond Graph Element, 584
12.5.3 Constructing Models with Accumulators and Restrictors, 587
12.5.4 Summary, 590
References, 592
Problems, 592

13 Nonlinear System Simulation 600

13.1 Explicit First-Order Differential Equations, 601
13.2 Differential Algebraic Equations Caused by Algebraic Loops, 604
13.3 Implicit Equations Caused by Derivative Causality, 608
13.4 Automated Simulation of Dynamic Systems, 612
13.4.1 Sorting of Equations, 613
13.4.2 Implicit and Differential Algebraic Equation Solvers, 614
13.4.3 Icon-Based Automated Simulation, 614
13.5 Example Nonlinear Simulation, 616
13.5.1 Some Simulation Results, 620
13.6 Summary, 623
References, 624
Problems, 624

Appendix: Typical Material Property Values Useful in Modeling Mechanical, Acoustic, and Hydraulic Elements 630

Index 633

PREFACE

This is the fifth edition of a textbook originally titled *system Dynamics: A Unified Approach,* which in subsequent editions acquired the title *System Dynamics: Modeling and Simulation of Mechatronic Systems.* As you can see, the subtitle has now expanded to be *Modeling, Simulation, and Control of Mechatronic Systems.*

The addition of the term *control* indicates the major change from previous editions. In older editions, the first six chapters of the book typically have been used as an undergraduate text and the last seven chapters have been used for more advanced courses. Now the latter part of Chapter Six can be used to introduce undergraduate students to a major use of mathematic models; namely, as a basis for the design control systems. In this case we are not trying to replace the many excellent books dealing with the design of automatic control systems. Rather we are trying to provide a contrasting approach to such books that often have a single chapter devoted to the construction of mathematical models from physical principles, while the rest of the book is devoted to discussing the dynamics of control systems given a model of the control system in the form of state equations, transfer functions, or frequency response functions. It is our contention that while the design of control systems is very important, the skills of modeling and computer simulation for a wide variety of physical systems are of fundamental importance even if an automatic control system is not involved.

Furthermore, we contend that the bond graph method is uniquely suited to the understanding of physical system dynamics. The basis of bond graphs lies in the study of energy storage and power flow in physical systems of almost any type. In this edition, we have tried to simplify the earlier chapters to focus on mechanical, electrical, and hydraulic systems that are relatively easy to model using bond graphs, leaving the more complex types of systems to the later chapters. It would be easy for an instructor to choose some topics of particular interest from these chapters to supplement the types of systems studied in the earlier part of the book if desired.

It has been gratifying to see that over the years, instructors and researchers world wide have learned that the bond graph technique is uniquely suited to the description of physical systems of engineering importance. It is our hope that this book will continue to provide useful information for engineers dealing with the analysis, simulation, and control of the devices of the future.

1

INTRODUCTION

This book is concerned with the development of an understanding of the dynamic physical systems that engineers are called upon to design. The type of systems to be studied can be described by the term *mechatronic*, which implies that while the elements of the system are mechanical in a general sense, electronic control will also be involved. For the design of a computer-controlled system, it is crucial that the dynamics of systems that exchange power and energy in various forms be thoroughly understood. Methods for the mathematical modeling of real systems will be presented, ways of analyzing systems in order to shed light on system behavior will be shown, and techniques for using computers to simulate the dynamic response of systems to external stimuli will be developed. In addition, methods of using mathematical models of dynamic systems to design automatic control systems will be introduced. Before beginning the study of physical systems, it is worthwhile to reflect a moment on the nature of the discipline that is usually called *system dynamics* in engineering.

The word *system* is used so often and so loosely to describe a variety of concepts that it is hard to give a meaningful definition of the word or even to see the basic concept that unites its diverse meanings. When the word *system* is used in this book, two basic assumptions are being made:

1. A system is assumed to be an entity separable from the rest of the universe (the environment of the system) by means of a physical or conceptual boundary. An air conditioning system, for example, can be thought of as a system that reacts to its environment (the temperature of the outside air, for example) and that interchanges energy and information with its environment. In this case the boundary is physical or spatial. An air traffic control system, however, is a complex system, the environment of which is not only the physical surroundings but also the fluctuating demands for air traffic, which ultimately come from human decisions about travel and

Copyright © 2012 John Wiley & Sons, Inc.

the shipping of goods. The unifying element in these two disparate systems is the conceptual boundary between what is considered to be part of the system and what represents an external disturbance or command originating from outside the system.

2. A system is composed of interacting parts. In an air conditioning system, we recognize devices with specific functions, such as compressors and fans, sensors that transmit information, and actuators that act on information, and so on. The air traffic control system is composed of people and machines with communication links between them. Clearly, the *reticulation* of a system into its component parts is something that requires skill and art, since most systems could be broken up into so many parts that any analysis would be swamped with largely irrelevant detail. The art and science of system modeling has to do with the construction of a model complex enough to represent the relevant aspects of the real system but not so complex as to be unwieldy.

These two aspects of systems can be recognized in everyday situations as well as in the more specific and technical applications that form the subject matter of most of this book. For example, when one hears a complaint that the transportation system in this country does not work well, one may see that there is some logic in using the word *system*. First of all, the transportation system is roughly identifiable as an entity. It consists of air, land, and sea vehicles and the human beings, machines, and decision rules by which they are operated. In addition, many parts of the system can be identified—cars, planes, ships, baggage handling equipment, computers, and the like. Each part of the transportation system could be further reticulated into parts (i.e., each component part is itself a system), but for obvious reasons we must exercise restraint in this division process.

The essence of what may be called the "systems viewpoint" is to concern oneself with the operation of a complete system rather than with just the operation of the component parts. Complaints about the transportation system are often real "system" complaints. It is possible to start a trip in a private car that functions just as its designers had hoped it would, transfer to an airplane that can fly at its design speed with no failures, and end in a taxi that does what a taxi is supposed to do, and yet have a terrible trip because of traffic jams, air traffic delays, and the like. Perfectly good components can be assembled into an unsatisfactory system.

In engineering, as indeed in virtually all other types of human endeavor, tasks associated with the design or operation of a system are broken up into parts that can be worked on in isolation to some extent. In a power plant, for example, the generator, turbine, boiler, and feed water pumps typically will be designed by separate groups. Furthermore, heat transfer, stress analysis, fluid dynamics, and electrical studies will be undertaken by subsets of these groups. In the same way, the bureaucracy of the federal government represents a splitting up of the various functions of government. All the separate groups working on an overall task must interact in some manner to make sure that not only will the parts of the system work, but also the system as a whole will perform its intended function. Many

times, however, oversimplified assumptions about how a system will operate are made by those working on a small part of the system. When this happens, the results can be disappointing. The power plant may undergo damage during a full load rejection, or the economy of a country may collapse because of the unfavorable interaction of segments of government, each of which assiduously pursues seemingly reasonable policies.

In this book, the main emphasis will be on studying system aspects of behavior as distinct from component aspects. This requires knowledge of the component parts of the systems of interest and hence some knowledge in certain areas of engineering that are taught and sometimes even practiced in splendid isolation from other areas. In the engineering systems of primary interest in this book, topics from vibrations, strength of materials, dynamics, fluid mechanics, thermodynamics, automatic control, and electrical circuits will be used. It is possible, and perhaps even common, for an engineer to spend a major part of his or her professional career in just one of these disciplines, despite the fact that few significant engineering projects concern a single discipline. Systems engineers, however, must have a reasonable command of several of the engineering sciences as well as knowledge pertinent to the study of systems per se.

Although many systems may be successfully designed by careful attention to static or steady-state operation in which the system variables are assumed to remain constant in time, in this book the main concern will be with *dynamic* systems, that is, those systems whose behavior as a function of time is important. For a transport aircraft that will spend most of its flight time at a nearly steady speed, the fuel economy at constant speed is important. For the same plane, the stress in the wing spars during steady flight is probably less important than the time-varying stress during flight through turbulent air, during emergency maneuvers, or during hard landings. In studying the fuel economy of the aircraft, a static system analysis might suffice. For stress prediction, a dynamic system analysis would be required.

Generally, of course, no system can operate in a truly static or steady state, and both slow evolutionary changes in the system and shorter time transient effects associated, for example, with startup and shutdown are important. In this book, despite the importance of steady-state analysis in design studies, the emphasis will be on dynamic systems. Dynamic system analysis is more complex than static analysis but is extremely important, since decisions based on static analyses can be misleading. Systems may never actually achieve a possible steady state due to external disturbances or instabilities that appear when the system dynamics are taken into account.

Moreover, systems of all kinds can exhibit counterintuitive behavior when considered statically. A change in a system or a control policy may appear beneficial in the short run from static considerations but may have long-run repercussions opposite to the initial effect. The history of social systems abounds with sometimes tragic examples, and there is hope that dynamic system analysis can help avoid some of the errors in "static thinking" [1]. Even in engineering with rather

simple systems, one must have some understanding of the dynamic response of a system before one can reasonably study the system on a static basis.

A simple example of a counterintuitive system in engineering is the case of a hydraulic power generating plant. To reduce power, wicket gates just before the turbine are moved toward the closed position. Temporarily, however, the power actually increases as the inertia of the water in the penstock forces the flow through the gates to remain almost constant, resulting in a higher velocity of flow through the smaller gate area. Ultimately, the water in the penstock slows down and power is reduced. Without an understanding of the dynamics of this system, one would be led to open the gates to *reduce* power. If this were done, the immediate result would be a gratifying decrease of power followed by a surprising and inevitable increase. Clearly, a good understanding of dynamic response is crucial to the design of a controller for mechatronic systems.

1.1 MODELS OF SYSTEMS

A central idea involved in the study of the dynamics of real systems is the idea of a *model* of a system. Models of systems are simplified, abstracted constructs used to predict their behavior. Scaled physical models are well known in engineering. In this category fall the wind tunnel models of aircraft, ship hull models used in towing tanks, structural models used in civil engineering, plastic models of metal parts used in photoelastic stress analysis, and the "breadboard" models used in the design of electric circuits.

The characteristic feature of these models is that some, but not all, of the features of the real system are reflected in the model. In a wind tunnel aircraft model, for example, no attempt is made to reproduce the color or interior seating arrangement of the real aircraft. Aeronautical engineers assume that some aspects of a real craft are unimportant in determining the aerodynamic forces on it, and thus the model contains only those aspects of the real system that are supposed to be important to the characteristics under study.

In this book, another type of model, often called a *mathematical model*, is considered. Although this type of model may seem much more abstract than the physical model, there are strong similarities between physical and mathematical models. The mathematical model also is used to predict only certain aspects of the system response to inputs. For example, a mathematical model might be used to predict how a proposed aircraft would respond to pilot input command signals during test maneuvers. But such a model would not have the capability of predicting every aspect of the real aircraft response. The model might not contain any information on changes in aerodynamic heating during maneuvers or about high-frequency vibrations of the aircraft structure, for example.

Because a model must be a simplification of reality, there is a great deal of art in the construction of models. An unduly complex and detailed model may contain parameters virtually impossible to estimate, may be practically impossible to analyze, and may cloud important results in a welter of irrelevant detail if it can

be analyzed. An oversimplified model will not be capable of exhibiting important effects. It is important, then, to realize that *no system can be modeled exactly* and that any competent system designer needs to have a procedure for constructing a variety of system models of varying complexity so as to find the simplest model capable of answering the questions about the system under study.

The rest of this book deals with models of systems and with the procedures for constructing models and for extracting system characteristics from models. The models will be mathematical models in the usual meaning of the term even though they may be represented by stylized graphs and computer printouts rather than the more conventional sets of differential equations.

System models will be constructed using a uniform notation for all types of physical systems. It is a remarkable fact that models based on apparently diverse branches of engineering science can all be expressed using the notation of *bond graphs* based on energy and information flow. This allows one to study the *structure* of the system model. The nature of the parts of the model and the manner in which the parts interact can be made evident in a graphical format. In this way, analogies between various types of systems are made evident, and experience in one field can be extended to other fields.

Using the language of bond graphs, one may construct models of electrical, magnetic, mechanical, hydraulic, pneumatic, thermal, and other systems using only a rather small set of ideal elements. Standard techniques allow the models to be translated into differential equations or computer simulation schemes. Historically, diagrams for representing dynamic system models developed separately for each type of system. For example, parts *a*, *b*, and *c* of Figure 1.1 each represent a diagram of a typical model. Note that in each case the elements in the diagram seem to have evolved from sketches of devices, but in fact a photograph of the real system would not resemble the diagram at all. Figure 1.1*a* might well represent the dynamics of the heave motion of an automobile, but the masses, springs, and dampers of the model are not directly related to the parts of an automobile visible in a photograph. Similarly, symbols for resistors and inductors in diagrams such as Figure 1.1*b* may not correspond to separate physical elements called resistors and chokes but instead may correspond to the resistance and inductance effects present in a single physical device. Thus, even semipictorial diagrams are often a good deal more abstract than they might at first appear.

When mixed systems such as that shown in Figure 1.1*d* are to be studied, the conventional means of displaying the system model are less well developed. Indeed, few such diagrams are very explicit about just what effects are to be included in the model. The basic structure of the model may not be evident from the diagram. A bond graph is more abstract than the type of diagrams shown in Figure 1.1, but it is explicit and has the great advantage that all the models shown in Figure 1.1 would be represented using exactly the same set of symbols. For mixed systems such as that shown in Figure 1.1*d*, a universal language such as bond graphs provide is required in order to display the essential structure of the system model.

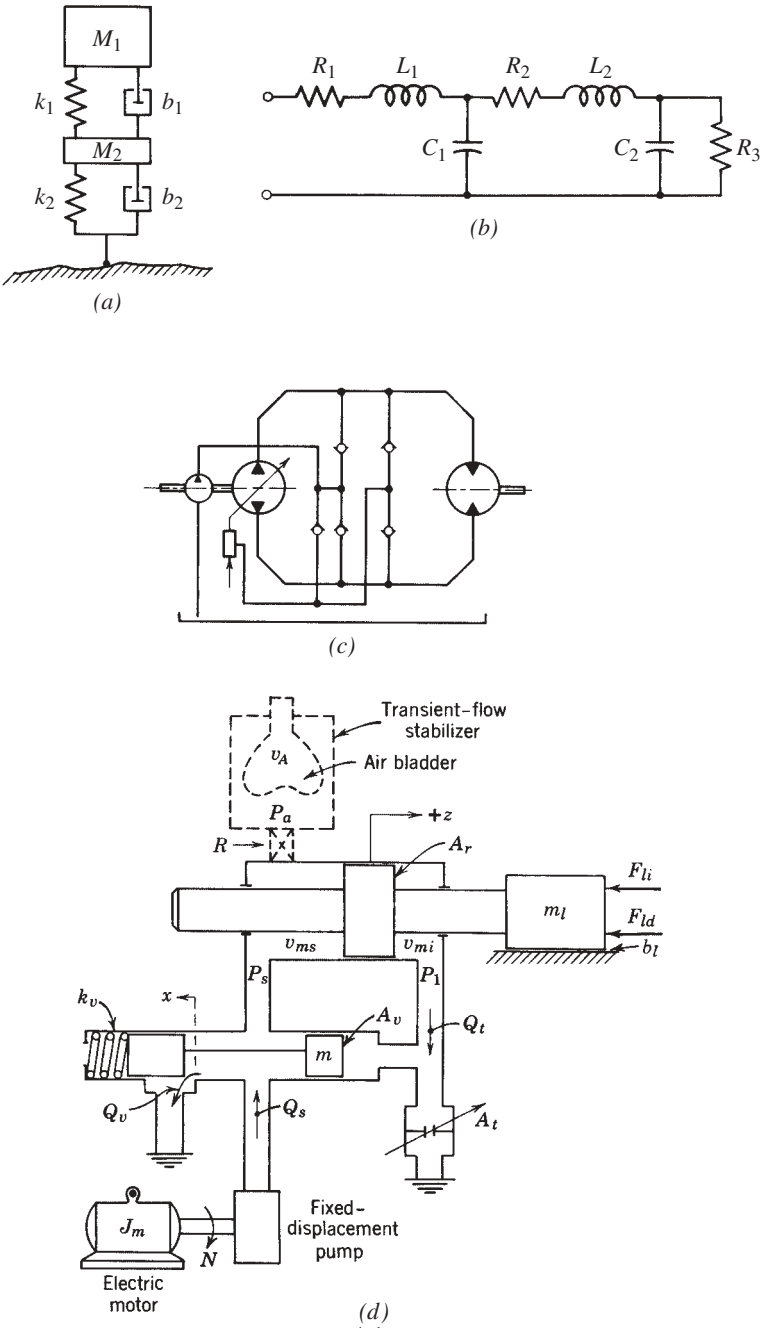


FIGURE 1.1. (*a*) Typical schematic diagram; (*b*) typical electric circuit diagram; (*c*) typical hydraulic diagram; (*d*) schematic diagram of system containing mechanical, electrical, and hydraulic components.

1.2 SYSTEMS, SUBSYSTEMS, AND COMPONENTS

To model a system, it is usually necessary first to break it up into smaller parts that can be modeled and perhaps studied experimentally and then to assemble the system model from the parts. Often, the breaking up of the system is conveniently accomplished in several stages. In this book major parts of a system will be called *subsystems* and primitive parts of subsystems will be called *components* or, at the most primitive level, *elements*. Of course, the hierarchy of components, subsystems, and systems can never be absolute, since even the most primitive part of a system could be modeled in such detail that it would be a complex subsystem. Yet in many engineering applications, the subsystem and component categories are fairly obvious.

Basically, a subsystem is a part of a system that will be modeled as a system itself; that is, the subsystem will be broken into interacting component parts. A component, however, is modeled as a unit and is not thought of as composed of simpler parts. One needs to know how the component interacts with other components and one must have a characterization of the component, but otherwise a component is treated as a "black box" without any need to know in detail what causes it to act as it does.

To illustrate these ideas, consider the vibration test system shown in Figure 1.2. The system is intended to subject a test structure to a vibration environment specified by a signal generator. For example, if the signal generator delivers a random-noise signal, it may be desired that the acceleration of the shaker table be a faithful reproduction of the electrical noise signal waveform. In a system that is assembled from physically separate pieces, it is natural to consider the parts that are assembled by connecting wires and hydraulic lines or by mechanical fasteners as subsystems. Certainly, the electronic boxes labeled signal generator, controller, and electrical amplifier are subsystems, as are the electrohydraulic valve, the

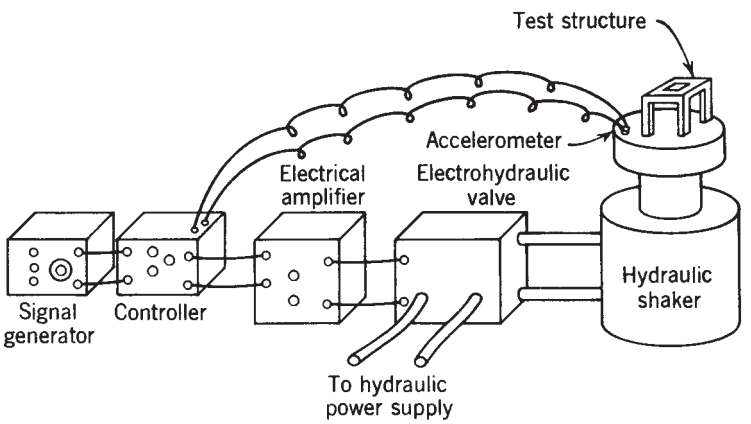


FIGURE 1.2. Vibration test system.

hydraulic shaker, and the test structure. It may be possible to treat some of these subsystems as components if their interactions with the rest of the system can be specified without knowledge of the internal construction of the subsystem. The electrical amplifier is obviously composed of many components, such as resistors, capacitors, transistors, and the like, but if the amplifier is sized correctly so that it is not overloaded, then it may be possible to treat the amplifier as a component specified by the manufacturer's input–output data. Other subsystems may require a subsystem analysis in order to achieve a dynamic description suitable for the overall system study.

Consider, for example, the electrohydraulic valve. A typical servo valve is shown in Figure 1.3. Clearly, the valve is composed of a variety of electrical, mechanical, and hydraulic parts that work together to produce the dynamic response of the valve. For this subsystem the components might be the torque motor, the hydraulic amplifier, mechanical springs, hydraulic passages, and the spool valve. A subsystem dynamic analysis can reveal weaknesses in the subsystem design that may necessitate the substitution of another subsystem or a reconfiguration of the overall system. Yet such an analysis may indicate that, from the point of view of the overall system, the subsystem may be adequately characterized as a simple component. A skilled and experienced system designer often makes a judgment on the appropriate level of detail for the modeling of a subsystem on an intuitive basis. A major purpose of the methods presented in this book is to show how system models can be assembled conveniently from component models. It is then possible to experiment with subsystem models of varying degrees of sophistication in order to verify or disprove initial modeling decisions.

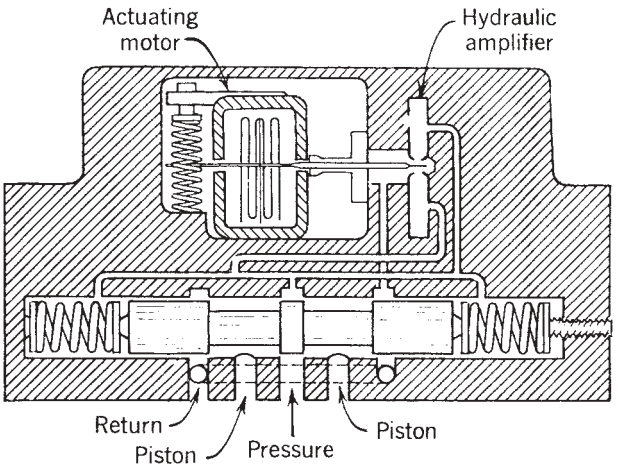


FIGURE 1.3. Electrohydraulic valve.

1.3 STATE-DETERMINED SYSTEMS

The goal of this book is to describe means for setting up mathematical models for systems. The type of model that will be found is often described as a *state-determined system.* In mathematical notation, such a system model is often described by a set of ordinary differential equations in terms of so-called *state variables* and a set of algebraic equations that relate other system variables of interest to the state variables. In succeeding chapters an orderly procedure, beginning with physical effects to be modeled and ending with state differential equations, will be demonstrated. Even though some techniques of analysis and computer simulation do not require that the state equations be written explictly, from a mathematical point of view all the system models that will be discussed are state-determined systems.

The future of all the variables associated with a state-determined system can be predicted if (1) the state variables are known at some initial time and (2) the future time history of the input quantities from the external environment is known.

Such models, which are virtually the only types used in engineering, have some built-in philosophical implications. For example, events in the future do not affect the present state of the system. This implication is correlated with the assumption that time runs only in one direction—from past to future. That models should have these properties probably seems plausible, if not obvious, yet it is remarkably difficult to conceive of a demonstration that real systems always have these properties.

Clearly, past history can have an effect on a system; yet the influence of the past is exhibited in a special way in state-determined systems. All the past history of a state-determined system is summed up in the present values of its state variables. This means that many past histories could have resulted in the same present value of state variables and hence the same future behavior of the system. It also means that if one can condition the system to bring the state variables to some particular values, then the future system response is determined by the future inputs and nothing is important about the past except that the state variables were brought to those values.

Scientific experiments are run as if the systems to be studied were state determined. The system is always started from controlled conditions that are expressed in terms of carefully monitored variables. If the experiment is repeatable, then the assumption is that the state variables are properly initialized by the operations used to set up the experiment. If the experiment is not repeatable, then the assumption is that some important influence has not been controlled. This influence can be either a state variable that was not monitored and initialized properly or an unrecognized input quantity through which the environment influences the system.

State-determined system models have proved useful over centuries of scientific and technical work. For the usual macroscopic systems encountered in engineering, state-determined system models are nearly universal, and there is continuing interest in developing such models for social and economic systems.

This book can be regarded as a textbook devoted to the establishment and study of state-determined system models using well-defined physical systems of interest to engineers as examples.

1.4 USES OF DYNAMIC MODELS

In Figure 1.4 a general dynamic system model is shown schematically. The system S is characterized by a set of state variables, indicated by X, that are influenced by a set of input variables U that represent the action of the system's environment on the system or variables that could be manipulated by a control system. The set of output variables Y are either back effects from the system onto the environment or variables that can be sensed and used by a control system. This type of dynamic system model may be used in four quite distinct ways:

1. *Analysis*. Given U for the future, X at the present, and the model S, predict the future of Y. Assuming that the system model is an accurate representation of the real system, analysis techniques allow one to predict system behavior. In some cases, this analysis will indicate the need of a control system in order for the system to respond in a desirable way.
2. *Identification*. Given time histories of U and Y, usually by experimentation on real systems, find a model S and state variables X that are consistent with U and Y. This is the essence of scientific experimentation. Clearly, a "good" model is one that is consistent with a great variety of sets U and Y.
3. *Synthesis*. Given U and some desired Y, find S such that U acting on S will produce Y. Most of engineering deals with synthesis, but only in limited contexts are there direct synthesis methods. Often we must be content to accomplish synthesis of systems via a trial-and-error process of repetitive analysis of a series of candidate systems. In this regard, dynamic models pay a vital role, since progress would be slow indeed if one had to construct each candidate system "in the metal" in order to discover its properties.
4. *Control*. Given the system model S, design a control system that uses those variables in the output Y available from sensors to produce the some of the inputs in U in order to make the system respond properly.

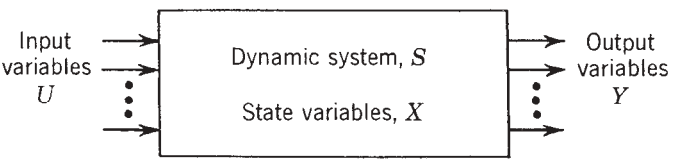


FIGURE 1.4. General dynamic system model.

In this book we will concentrate on setting up system models and predicting the behavior of the systems using analytical or computational techniques. Thus, we will concentrate on analysis, but it is important to remember that the techniques are useful for identification problems and that the major challenge to a systems engineer is to synthesize desirable systems. It may not be too much to say that analysis, except in the service of synthesis, is a rather sterile pursuit for an engineer. Also, a brief introduction to automatic control techniques will be given.

1.5 LINEAR AND NONLINEAR SYSTEMS

An overall system model, consisting of subsystems and their components, requires modeling decisions as to what dynamic effects must be included in order to use the model for its intended purpose. The result of these modeling decisions is typically a system schematic that indicates the important dynamic effects. Figures 1.1 and 1.2 are examples of system schematics where modeling decisions have been made. In Figure 1.1*d*, the important dynamic effects at the component level are indicated by labeling inertial, compliance, and resistance effects. In Figure 1.2, modeling decisions are indicated at the subsystem level, while the detail modeling of each subsystem remains to be done. A very important aspect of the modeling process is whether components of subsystems behave linearly or nonlinearly. As we progress through the chapters, it will become very clear as to what is meant by linear or nonlinear behavior. For now it is simply stated that linear systems are represented by sets of linear, first-order differential state equations, and nonlinear systems, while still state determined, are represented by sets of nonlinear, first-order differential equations.

If it is justified to assume that an overall system can be represented as linear, then there exist an abundance of analytical tools for obtaining exact analytical solutions to the linear equations and for extracting detailed information about the response of the system. Some of the analytical information that is covered in later chapters includes eigenvalues, transfer functions, and frequency response. If the systems have large numbers of state variables, then pencil and paper analysis may be virtually impossible, and one must resort to numerical computation to obtain the properties of the system.

If a single component in a system model is represented as a nonlinear element, then the system is nonlinear, and linear analysis tools will not work. Sliding friction is an example of a common nonlinear phenomenon. If friction is important in a system, eigenvalues, transfer functions, and frequency response concepts do not apply as they would to a strictly linear system. To extract information about the response of nonlinear systems, one must resort to time step simulation. Fortunately, there is an abundance of commercial programs to simulate nonlinear systems so that the use of nonlinear models is not the stumbling block it was before the advent of computers.

The fact is that there are no real physical systems that are truly linear. However, in order to introduce the concepts of constructing overall system models of interacting electrical, mechanical, hydraulic, and thermal components, it is easier to start with linear systems and then extend the procedures to nonlinear systems. In addition, it is hard to understand the dynamics of system without first concentrating of the special case of linear systems. In the following five chapters, the emphasis is on linear system models, but whenever possible, the reader is reminded that real physical systems are nonlinear and simulation tools must be used to obtain system responses. The more advanced topics dealt with in Chapters 7–13 are largely concerned with nonlinear phenomena.

1.6 AUTOMATED SIMULATION

Mathematical models of dynamic physical systems have been made ever since the invention of differential equations. But until the development of powerful computers, there were severe limitations on the analysis of these models. Practically speaking, dynamic system behavior could generally be predicted only for low-order linear models that often were not very accurate representatives of real systems.

There is a lot to be said for the study of low-order linear models in order to gain an appreciation of system dynamics, and the first six chapters of this book deal primarily with just such system models. However, computer simulation can now be used to gain experience with system dynamics even when the system models become large and when they contain nonlinear elements. Chapter 13 discusses some of the issues that arise when dealing with complex but realistic models.

The next Chapters, 2, 3, and 4, present techniques for representing elements of mechanical, electrical, and fluid systems (and combination systems) in the abstract form of bond graphs instead of the schematic diagrams usually used to show vibratory systems, electric circuits, or hydraulic systems. For some, this may seem to be an unnecessary step away from physical reality, but it has useful consequences.

First of all, a bond graph is a precise way to represent a mathematical model. Often schematic diagrams are not entirely clear about whether certain effects are to be included or neglected in the model. Second, for many systems involving two or more forms of energy, such as mechanical, electrical, and hydraulic, there are no standard schematic diagrams that clearly indicate assumptions made in the modeling process. Finally, it is much easier to communicate a bond graph model unambiguously to a computer than a schematic diagram.

The bond graph uses only a few standard symbols, whereas typical schematic diagrams for the same system model drawn by different people are almost never identical. Just as computers more easily read bar codes than handwriting, they more easily interpret bond graphs than schematic diagrams.

Computer programs have been developed that recognize bond graphs and can process them in the same manner that a human would in order to extract differential equations for analysis or simulation. In the process, useful facts about the mathematics of the model are discovered even before any numerical parameters or laws have been supplied. Furthermore, when the parameters of the elements and the forcing functions for the system have been specified, the programs can then simulate the response of the system. In this process, only a minimum of human intervention is required.

Although it is important for a system engineer to understand the entire process of modeling and simulation, the use of bond graphs and bond graph simulation programs allows a beginner to start developing the skills associated with computer simulation even before all the bond graph modeling techniques have been learned. This has proved to be very effective in teaching. From the first day, a student given a bond graph model and a bond graph simulation program can see how the model reacts to various input forcing functions and to variations in system parameters. Simple design studies on dynamic systems can be assigned without waiting until the student has learned to make models, derive equations, and use an equation solver. This provides motivation to learn about bond graph dynamic system modeling and numerical simulation techniques.

The fact that the simulation programs are effective for nonlinear as well as linear models, and for large models as well as small ones, may not be fully appreciated by a beginning student. However, in the course of time the significance of this fact should become apparent.

References [2–4] give the names of some of the more well-known commercial bond graph processors, some of which are used in conjunction with simulation programs to solve the differential equations. A web search will reveal that there are a number of other bond graph processor programs.

Another category of program is based on a stored library of predetermined bond graph submodels, but in use, replaces bond graph submodels with icons. See Reference [5], for example. Such programs are useful for studying large engineering models but they are less useful for learning about bond graph modeling.

REFERENCES

[1] J. W. Forrester, Urban Dynamics, Cambridge, MA: MIT Press, 1969.

[2] CAMP-G, Cadsim Engineering, P.O. Box 4083, Davis, CA 95616.

[3] 20-SIM, Controllab Products B.V., Drienerlolaan 5, EL-CE, 7255 NB Enschede, Netherlands.

[4] SYMBOLS 2000, Hightech Consultants, STEP IIT Ktragpur-721302, WB, India.

[5] The LMS Imagine. Lab AMESim Introduction, http://www.imsintl.com/imagine-amesim-intro.

PROBLEMS

1-1. Suppose you were a heating engineer and you wished to consider a house as a dynamic system. Without a heater, the average temperature in the house would clearly vary over a 24-h period. What might you consider for inputs, outputs, and state variables for a simple dynamic model? How would you expand your model so that it would predict temperatures in several rooms of the house? How does the installation of a thermostatically controlled heater change your model?

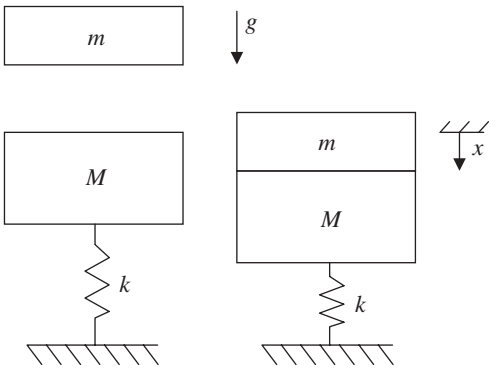
1-2. For a particular car operated on a level road at steady conditions there is a relation between throttle position and speed. Sketch the general shape you would expect for this curve. If recordings were made of instantaneous speed and throttle position while the car was driven normally for several miles on ordinary roads, do you think that the instantaneous values of speed and throttle position would fall on the steady-state curve? What inputs, outputs, and state variables might prove useful in trying to find a dynamic model useful in predicting dynamic speed variations?

1-3. A car is driven over a curb twice—once very slowly and once quite rapidly. What would you need to know about the car in the second case that you did not need to know in the first case if you were required to find the tire force that resulted from going over the curb?

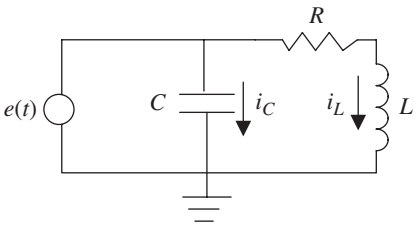
1-4. In the steady state a good weather vane points into the wind, but when the wind shifts, the vane cannot always be trusted to be pointing into the wind. Identify inputs, outputs, and the parameters of the weather vane system that affect its response to the wind. Sketch your idea of how the position of the vane would change in time if the wind suddenly shifted 10°.

1-5. The height of water in a reservoir fluctuates over time. If you had to construct a dynamic system model to help water resource planners predict variations in the height, what input quantities would you consider? How many state variables do you think you would need for your model?

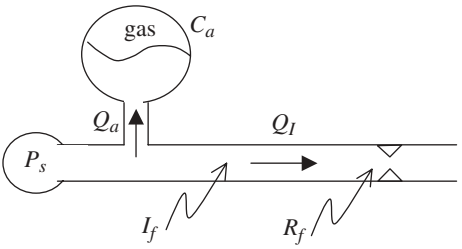
1-6. A mass, M, and spring, k, are at rest in a gravity field, about to be struck by a mass, m, falling from a height, h. The mass, m, sticks to M, and the two move downward. The variable, x, keeps track of the displacement after impact. Sketch the general motion of x for some period of time after impact. How many equations do you think are needed to describe this system mathematically? If the system ever came to rest, what would be the deflection of the spring?



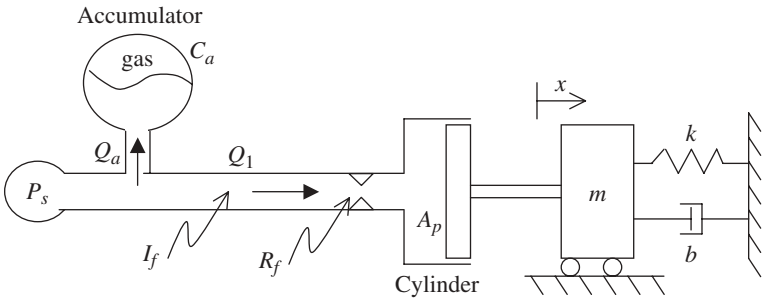
1-7. The system is an electric circuit consisting of an input voltage, $e(t)$, and a capacitor, resistor, and inductor, C, R, L. As will be seen in later chapters, if a voltage is applied to a capacitor, current flows easily at first and then slows as the capacitor becomes charged. Inductors behave just the opposite, in that they reluctantly pass current when a voltage is first applied, and then the current passes easily as time passes. If the input voltage is suddenly raised from zero to some constant value, sketch the current in the capacitor, i_C, and the inductor, i_L, as a function of time. What is the steady-state current in the capacitor and inductor?



1-8. A hydraulic system consists of a supply pressure, P_s, and a long fluid-filled tube. Branching from the tube is an accumulator, C_a, consisting of a compliant gas separated from the fluid by a diaphragm. The long tube has inertia and resistance, I_f and R_f. It may be hard to believe at this point, but this hydraulic system exhibits identical behavior to that of the electric circuit of Problem 1-7. Armed with this information, sketch the volume flow rate of the fluid into the accumulator, Q_a, and in the tube, Q_I, as a function of time.



1-9. Shown here is the hydraulic system in Problem 1-8 connected to a hydraulic cylinder of piston area A_p. The piston is connected to a mass, m, attached to the ground through the spring and damper, k and b. This is a "system" consisting of interacting hydraulic and mechanical components. How many variables do you think are needed to fully describe the motion of the system? Sketch how you think the volume flow rate in the accumulator and in the tube will respond to a sudden elevation of the supply pressure. Sketch the motion, x, of the mass.



2

MULTIPORT SYSTEMS AND BOND GRAPHS

In this chapter the first steps are taken toward the development of system-modeling techniques for engineering systems involving power interactions. First, major subsystems are identified, and the means by which the subsystems are interconnected are studied. The fact that interacting physical systems must transmit power then is used to unify the description of interconnected subsystems. A uniform classification of the variables associated with power and energy is established, and bond graphs showing the interconnection of subsystems are introduced. Finally, the notions of inputs, outputs, and pure signal flows are discussed.

2.1 ENGINEERING MULTIPORTS

In Figure 2.1 a representative collection of subsystems or components of engineering systems is shown. Although the subsystems sketched are quite elementary, they will serve to introduce the concept of an *engineering multiport*. The term "engineering" is used to imply that the devices are physical subsystems used to build up systems such as automobiles, television sets, machine tools, or electric power plants that are designed to accomplish some specific objectives. The term "multiport" refers to a point of view taken in the description of the subsystems.

Inspection of Figure 2.1 reveals that a number of variables have been labeled on the subsystems. These variables are torques, angular speeds, forces, velocities, voltages, currents, pressures, and volume flow rates. The variables occur in pairs associated with points at which the subsystems could be connected with other subsystems to form a system. It would be possible, for example, to couple the shaft of the electric motor (a) to one end of the drive shaft (c) and the hydraulic motor shaft (b) to the other end of the drive shaft. After the coupling, the motor torque and speed would be identical to the torque and speed of one end of the

Copyright © 2012 John Wiley & Sons, Inc.

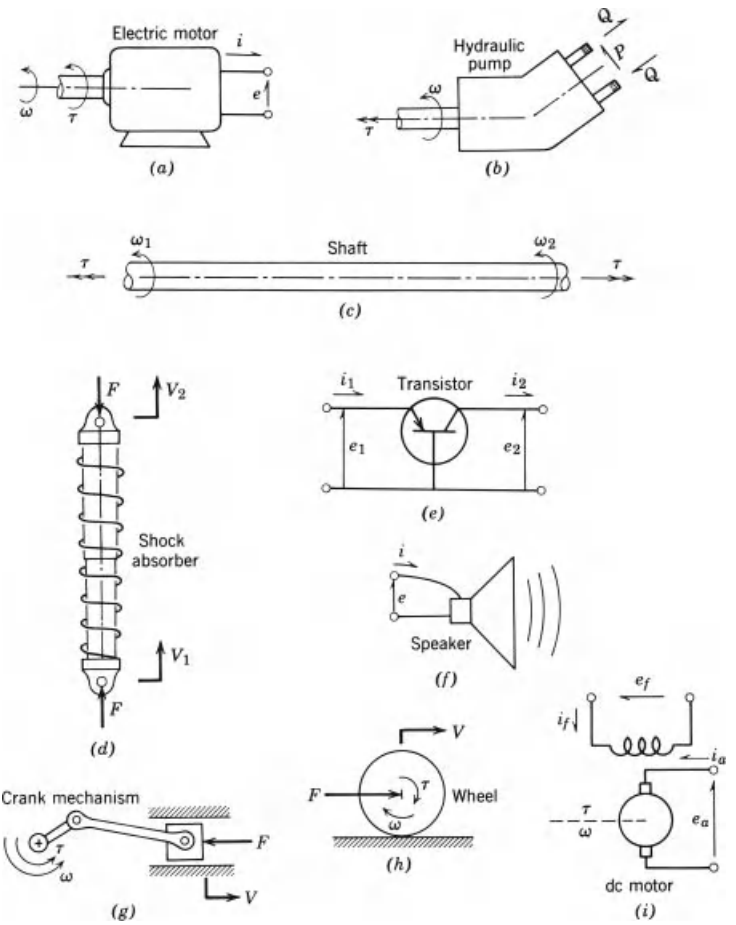


FIGURE 2.1. A collection of engineering multiports. (a) Electric motor: torque τ, angular speed ω, voltage e, current i; (b) hydraulic pump: torque τ, angular speed ω, pressure P, volume flow rate Q; (c) drive shaft: torque τ, angular speeds ω_1 and ω_2; (d) spring shock absorber unit: force F, velocities V_1 and V_2; (e) transistor: voltages e_1 and e_2, currents i_1 and i_2; (f) loudspeaker: voltage e, current i; (g) crank and slider mechanism: torque τ, angular speed ω, force F, velocity V; (h) wheel: force F, velocity V, torque τ, angular speed ω; (i) separately excited direct current (dc) motor: torque τ, angular speed ω, voltages e_a and e_f, currents i_a and i_f.

drive shaft. Similarly, the torque and speed of the other end of the drive shaft would be identical to the torque and speed of the hydraulic pump. (If the drive shaft were not rigid, then the two angular speeds, ω_1 and ω_2, on the drive shaft ends would not necessarily be equal at all times.) Similarly, one could connect the two terminals of the transistor (e) associated with e_2 and i_2 to the terminals of the loudspeaker (f). After the connection, the voltage and current associated

with one terminal pair of the transistor would be identical to the voltage and current associated with the loudspeaker terminals. Generally, when two subsystems or components are joined together physically, two complementary variables are simultaneously constrained to be equal for the two subsystems.

Places at which subsystems can be interconnected are places at which power can flow between the subsystems. Such places are called *ports*, and physical subsystems with one or more ports are called *multiports*. A system with a single port is called a *1-port*, a system with two ports is called a *2-port*, and so on. The multiports in Figures 2.1$a-h$ are shown as 2-ports. Figure 2.1f is a 1-port as long as it is considered only as an electrical element and not as an element coupling electrical and acoustic subsystems. Figure 2.1i is shown as a 3-port.

The variables listed for the multiports in Figure 2.1 and the variables that are forced to be identical when two multiports are connected are called *power variables*, because the product of the two variables considered as functions of time is the instantaneous power flowing between the two multiports. For example, if the electrical motor (Figure 2.1a) were coupled to the hydraulic pump (Figure 2.1b), the power flowing from the motor to the pump would be given by the product of the angular speed and the torque. Since power could flow in either direction over time, a sign convention for the power variables will be established. Similarly, power can be expressed as the product of a force and a velocity for a multiport in which mechanical translation is involved, as the product of voltage and current for an electrical port, and as the product of pressure and volume flow rate for a port at which hydraulic power is interchanged.

Since power interactions are always present when two multiports are connected, it is useful to classify the various power variables in a universal scheme and to describe all types of multiports in a common language. In this book all power variables are called either *effort* or *flow*. Table 2.1 shows effort and flow variables for several types of power interchange.

As Table 2.1 indicates, in general discussions the symbols $e(t)$ and $f(t)$ are used to denote effort and flow quantities as functions of time. For specific applications, more traditional notation suggestive of the physical variable involved may be used. A curse of system analysis that becomes evident as soon as problems involving several energy domains are studied is that it is hard to establish notation that does not conflict with conventional usage. In Table 2.1, for example, a force is an effort quantity, $e(t)$, even though the common use of the letter F to stand for a force might be confused with the $f(t)$, which stands for a flow quantity. These notational difficulties are bothersome but not fundamental, and cannot be avoided except by using entirely new notation. For example, the letter Q has been used for charge in electric circuits, volume flow rate in hydraulics, and heat energy in heat transfer. In this book, both the generalized notation e and f and the physical notation in Figure 2.1 and Table 2.1 are used. The context in which the symbols are used will resolve any possible ambiguities in meaning. The power, $\mathsf{P}(t)$, flowing into or out of a port can be expressed as the product of an effort and a flow variable, and thus in general notation is given by the

TABLE 2.1. Some Effort and Flow Quantities

Domain	Effort, $e(t)$	Flow, $f(t)$
Mechanical translation	Force component, $F(t)$	Velocity component, $V(t)$
Mechanical rotation	Torque component, $\tau(t)$	Angular velocity component, $\omega(t)$
Hydraulic	Pressure, $P(t)$	Volume flow rate, $Q(t)$
Electric	Voltage, $e(t)$	Current, $i(t)$

following expression:

$$\mathsf{P}(t) = e(t)\,f(t). \tag{2.1}$$

In a dynamic system the effort and the flow variables, and hence the power, fluctuate in time. Two other types of variables turn out to be important in describing dynamic systems. These variables, sometimes called *energy variables* for reasons that will become clearer later, are called the *momentum* $p(t)$ and the *displacement* $q(t)$ in generalized notation.

The momentum is defined as the time integral of an effort. That is,

$$p(t) \equiv \int^t e(t)\,dt = p_0 + \int_{t_0}^t e(t)\,dt, \tag{2.2}$$

in which either the indefinite time integral can be used or one may define p_0 to be the initial momentum at time t_0 and use the definite integral from t_0 to t. In the same way, a displacement variable is the time integral of a flow variable:

$$q(t) \equiv \int^t f(t)\,dt = q_0 + \int_{t_0}^t f(t)\,dt. \tag{2.3}$$

Again, the second integral expression in Eq. (2.3) indicates that at time t_0 the displacement is q_0.

Other ways of writing the definitions in Eqs. (2.2) and (2.3) follow by considering the differential rather than the integral forms:

$$\frac{dp(t)}{dt} = e(t), \quad dp = e\,dt; \tag{2.2a}$$

$$\frac{dq(t)}{dt} = f(t), \quad dq = f\,dt. \tag{2.2b}$$

The energy, $\mathbf{E}(t)$, which has passed into or out of a port is the time integral of the power, $\mathsf{P}(t)$. Thus,

$$\mathbf{E}(\mathrm{t}) \equiv \int^t \mathsf{P}(t)\,dt = \int^t e(t)\,f(t)\,dt. \tag{2.4}$$

The reason p and q are sometimes called energy variables is that in Eq. (2.4), one may be able to write $e\,dt$ as dp or $f\,dt$ as dq by using Eq. (2.2a) or (2.2b). Alternative expressions for $\mathbf{E}$ then follow:

$$\mathbf{E}(t) = \int^t e(t)\,dq(t) = \int^t f(t)\,dp(t). \tag{2.5}$$

In the next chapter, cases will be encountered in which an effort is a function of a displacement or a flow is a function of a momentum. Then the energy can be expressed not only as a function of time but also as a function of one of the energy variables; thus,

$$\mathbf{E}(q) = \int^q e(q)\,dq \tag{2.5a}$$

or

$$\mathbf{E}(p) = \int^p f(p)\,dp. \tag{2.5b}$$

This provides the motivation for calling p and q energy variables in distinction to the power variables e and f.

In Figure 2.2 a mnemonic device, fancifully called the "tetrahedron of state," is shown. The four variable types, e, f, p, and q, are associated with the four vertices of a tetrahedron. Along two of the edges of the tetrahedron are indicated the relationships between e and p and between f and q. In Chapter 3 the same figure will be augmented to display the variables related by certain basic multiport elements.

It is an interesting fact that the only types of variables that will be needed to model physical systems are represented by the power and energy variables, e, f, p, and q. To make this statement more plausible, let us study the variables in several energy domains in more detail.

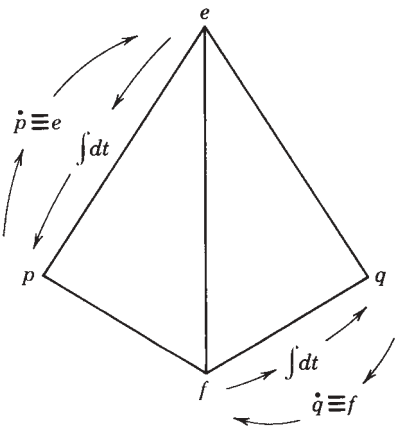


FIGURE 2.2. The tetrahedron of state.

TABLE 2.2. Power and Energy Variables for Mechanical Translational Systems

Generalized Variables	Mechanical Translation	SI Units
Effort, e	Force, F	newtons (N)
Flow, f	Velocity, V	meters per second (m/s)
Momentum, p	Momentum, P	N-s
Displacement, q	Displacement, X	m
Power, P	$F(t)V(t)$	watts (N-m/s = W)
Energy, **E**	$\int^x F\,dx$, $\int^p V\,dP$	joules (N-m = J)

Table 2.2 shows power and energy variables for mechanical translational ports. Since the power variables force and velocity are considered primitive, the units of the remaining variables follow from a choice in units for the power variables. Units are the shoals on which many a system analysis has foundered, and it is worthwhile to consider the great advantage that the metric or International System of Units (SI) has for system dynamic studies as compared with the many other unit systems that have been used in the past. In the SI system, power is always measured in newton-meters per second (N-m/s) or the equivalent watts (W), no matter what type of physical system is being studied. Similarly, energy will always be measured in newton-meters or the equivalent joules (J) for an electrical, mechanical, hydraulic, or any other type of physical system. Thus, if the e, f, p, and q variables are given SI units, no bothersome unit conversions will be necessary to properly account for power and energy interactions.

Anyone who has attempted to describe a complex system using traditional units such as pounds, slugs, feet, volts, pounds per square inch, gallons per hour, and the like will appreciate how difficult it is to ensure that the proper unit conversions have been incorporated. In fact, computer programs for processing bond graphs into the equivalent differential equations for subsequent analysis and simulation are incapable of incorporating conversion factors and thus essentially assume that the SI system will be used. In this text, we will make the same assumption. After an analysis or simulation has been completed, it is a relatively simple matter to convert some results to a traditional unit system if desired. The power demand of an electric car in kilowatts (kW), for example, could readily be converted to horsepower if this would be better understood by consumers, but we believe that it is a mistake to create a mathematical model using units that require conversion factors internally.

Table 2.3 gives power and energy variables for ports involving mechanical rotation. The shafts of motors, pumps, gears, and many other useful devices represent such ports.

The entries in Table 2.4 for hydraulic power again are related to the variables used in solid mechanics, but some unusual quantities are defined. The momentum quantity is defined according to Eq. (2.2) as the integral of the effort, or in this case, the pressure. Not only is the pressure momentum a quantity not often

TABLE 2.3. Power and Energy Variables for Mechanical Rotational Ports

Generalized Variables	Mechanical Rotation	SI Units
Effort, e	Torque, τ	newton-meters (N-m)
Flow, f	Angular velocity, ω	radians per second (rad/s)[a]
Momentum, p	Angular momentum, p_τ	N-m-s
Displacement, q	Angle, θ	rad[a]
Power, P	$\tau(t)\omega(t)$	N-m/s = W
Energy, **E**	$\int^{\theta} \tau\, d\theta$, $\int^{p_\tau} \omega\, dp_\tau$	N-m = J

[a]Radians and other angular measures are dimensionless, but there are scale factors between, say, radians, revolutions, and degrees which can cause errors not discoverable by dimensional analysis. The formulas used in this book all are based on the radian as the unit of angular measure.

TABLE 2.4. Power and Energy Variables for Hydraulic Ports

Generalized Variables	Hydraulic Variables	SI Units
Effort, e	Pressure, P	newtons per square meter Pa = (N/m^2)[a]
Flow, f	Volume flow rate, Q	cubic meters per second (m^3/s)
Momentum, p	Pressure momentum, p_p	N-s/m^2
Displacement, q	Volume, V	m^3
Power, P	$P(t)Q(t)$	N-m/s = W
Energy, **E**	$\int^{v} P\, dV$, $\int^{p_p} Q\, dp_p$	N-m = J

[a]In subsequent tables, when pressure is involved, the units will be given as N/m^2 rather than the equivalent pascals (Pa) for clarity.

encountered in conventional fluid mechanics, but it is also a quantity without an obvious symbol. The symbol p_p is meant to indicate a momentum quantity that is the integral of $P(t)$, just as in Table 2.3 p_τ was a momentum quantity defined as the time integral of $\tau(t)$. Fortunately, the lack of a commonly accepted symbol for certain variables is not a serious handicap. When some facility in system modeling has been developed, the generalized variables, e, f, p, and q, can be used for variables in all the energy domains, if desired.

Finally, Table 2.5 gives power and energy variables for electrical ports. The only new quantity that needs to be defined is the unit of electrical charge, the coulomb. It is common to use *volts* and *amperes* for the units of voltage and current rather than their equivalents in terms of coulombs and SI units. Most of the variables in Table 2.5 should be familiar, with the possible exception of the momentum or flux linkage variable λ. The usefulness of this variable will become evident when inductors are studied in Chapter 3.

TABLE 2.5. Power and Energy Variables for Electrical Ports

Generalized Variable	Electrical Variable	Units
Effort, e	Voltage, e	volt (V) = newton-meter per coulomb (N-m/C)
Flow, f	Current, i	ampere (A) = coulomb per second (C/s)
Momentum, p	Flux linkage variable, λ	V-s
Displacement, q	Charge, q	C = A − s
Power, P	$e(t)i(t)$	V-A = W = N-m/s
Energy, **E**	$\int^q e\,dq$, $\int^\lambda i\,d\lambda$	V-A-s = W-s = N-m = J

The tables of variables presented give at least preliminary evidence that the variables associated with a variety of physical systems can be fitted into the scheme of Figure 2.2. The usefulness of this viewpoint will become increasingly evident as systems are modeled in detail. In the next sections the ways in which subsystem interconnections can be indicated graphically using the e, f, p, q classification will be shown.

2.2 PORTS, BONDS, AND POWER

The devices sketched in Figure 2.1 can all be treated as multiport elements with ports that can be connected to other multiports to form systems. Further, when two multiports are connected, power can flow through the connected ports and the power can be expressed as the product of an effort and a flow quantity, as given in Tables 2.2–2.5. We now develop a universal way to represent multiports and systems of interconnected multiports based on the variable classifications in the tables.

Consider the separately excited dc motor shown in Figure 2.3. Physically, such motors have three obvious ports. The two electrical ports are represented by armature and field terminal pairs, and the shaft is a rotary mechanical port as sketched in Figure 2.3*a*. Figure 2.3*b* is a conventional schematic diagram in which the mechanical shaft is represented by a dashed line, the field coils are represented by a symbol similar to the circuit symbol for an inductance, and the armature is represented by a highly schematic sketch of a commutator and brushes. Note that the schematic diagram does not indicate what the detailed internal model of this subsystem or component will be. To write down equations describing the motor, an analyst must decide how detailed a model is necessary.

Figure 2.3*c* represents a further step in simplifying the representation of this engineering multiport. The name *dc motor* is used to stand for the device, and the ports are simply indicated by single lines emanating from the word representing the device. In a system in which several subsystems are connected, these port

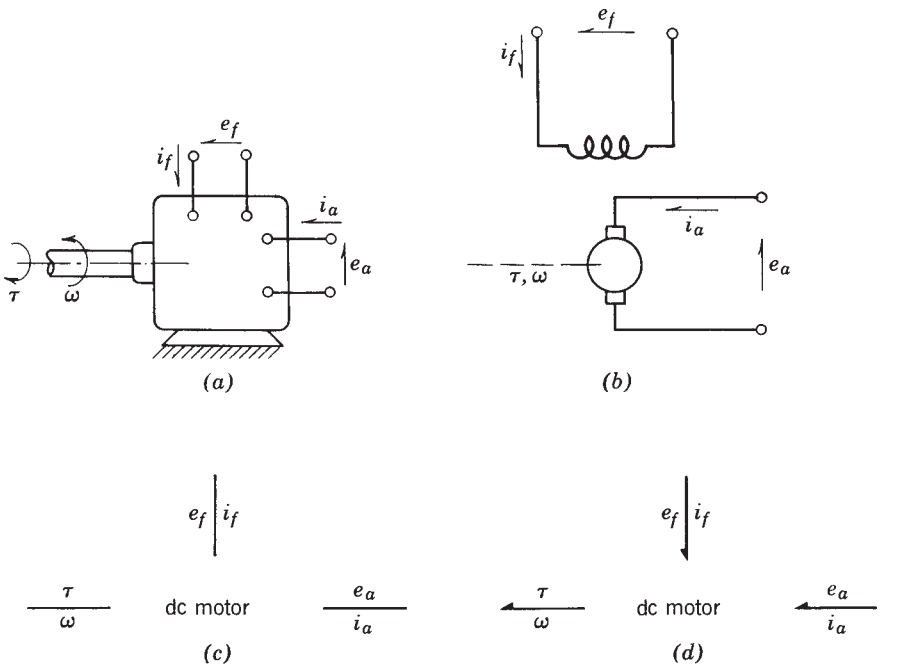


FIGURE 2.3. Separately excited dc motor: (*a*) sketch of motor; (*b*) conventional schematic diagram; (*c*) multiport representation; (*d*) multiport representation with sign convention for power.

lines will be called *bonds*. As a convenience, the effort and flow variables may be written next to the lines representing the ports. The notation can reflect the typical variables used in the particular energy domains involved. Whenever the port lines are either horizontal or vertical, it is useful to use the following convention:

- Efforts are placed either *above* or to the *left* of the port lines.
- Flows are placed either *below* or to the *right* of port lines.
- When diagonal lines are used, some judgment is required for placement of the effort and flow variables.

Note that Figures 2.3*a*, *b*, and *c* all contain the same information, namely, that the dc motor is a 3-port with power variables τ, ω, e_f, i_f, e_a, and i_a. In Figure 2.3*d*, a sign convention has been added: The *half-arrow* on a port line indicates the direction of power flow at any instant of time when the product of the effort and flow variables happens to be positive.

For example, if ω is positive in the direction shown in Figure 2.3*a* and if τ is interpreted to be the torque on the motor shaft resulting from a connection to some other multiport and is positive in the direction shown in Figure 2.3*a*,

then when τ and ω are both positive (or, for that matter, both negative), the product $\tau\omega$ is positive and represents power flowing *from* the motor to some other multiport coupled to the motor shaft. Thus, the half-arrow in Figure 2.3*d* points *away* from the dc motor. Similarly, when e_f, i_f, e_a, and i_a are positive, power flows to the motor from whatever other multiports are connected to the field and armature terminals. Hence, the half-arrows associated with the field and armature ports point *toward* the motor.

Anytime one desires to be specific about the characteristics of a multiport—for instance, in equation form or in the form of tabulated data—then a sign convention is necessary. The establishment of sign conventions is fairly straightforward for electric circuits or for the circuit-like parts of representations of multiports such as those of Figures 2.3*a* and *b*.

In mechanics, however, anyone who has struggled with the definition of forces and moments on interconnected rigid bodies using "free-body diagrams" knows that the establishment of sign conventions in mechanical systems is not trivial. The problem is that the action and reaction forces show up as oppositely directed in most representations. Thus, in Figure 2.3*a*, one must decide whether τ represents the torque *on* the motor shaft or *from* the motor onto some other multiport. On diagrams such as Figure 2.3*b*, the mechanical signs are often not indicated at all, and it is up to the analyst to insert plus or minus signs in the equations without much help from the schematic diagram of the system.

When two multiports are coupled together so that the effort and flow variables become identical, the two multiports are said to have a common *bond*, in analogy to the bonds between component parts of molecules. Figure 2.4 shows part of a system consisting of three multiports bonded together. The motor and pump have a common angular speed, ω, and torque at the coupling, τ. The battery and the motor have a common voltage and current defined at the terminals at which the battery leads connect to the motor armature. To represent this type of subsystem interconnection in the manner of Figure 2.3*c* or *d* is very straightforward; the

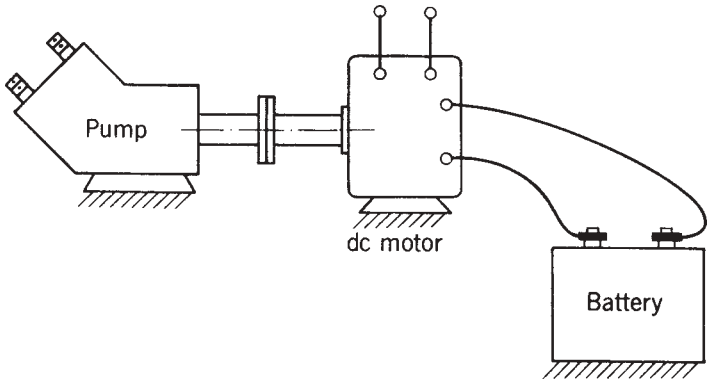


FIGURE 2.4. Partially assembled system.

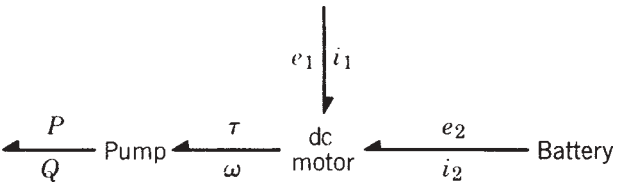


FIGURE 2.5. Word bond graph for system of Figure 2.4.

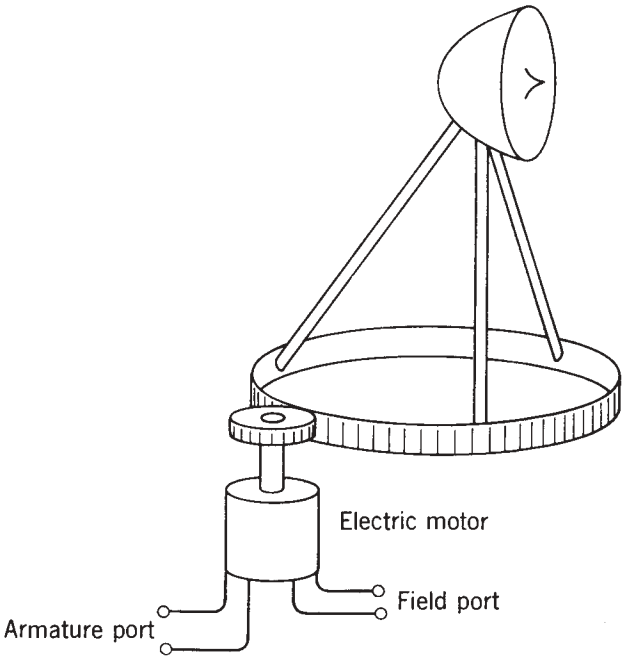
joined ports are represented by a single line or bond between the multiports. This has been done in Figure 2.5. The line between the pump and motor in Figure 2.5 implies that a port of the motor and a port of the pump have been connected, and hence a single torque and a single angular speed pertain to both the pump and the motor. The half-arrow on the bond means that the torque and the speed are defined in such a way that when their product, $\tau\omega$, is positive, power is flowing from the motor to the pump. Thus, lines associated with isolated multiports indicate ports or potential bonds. For interconnected multiports, a single line represents the conjunction of two ports, that is, a bond.

2.3 BOND GRAPHS

The mechanism for studying dynamic systems to be used subsequently in this book is the *bond graph*. A bond graph simply consists of subsystems linked together by lines representing *power bonds*, as in Figure 2.5. When major subsystems are represented by words, as in Figure 2.5, then the graph is called a *word bond graph*. Such a bond graph establishes multiport subsystems, the way in which the subsystems are bonded together, the effort and flow variables at the ports of the subsystems, and sign conventions for power interchanges.

Since the word bond graph serves to make some initial decisions about the representation of dynamic systems, it is worthwhile to consider some example systems even before the details of dynamic systems models have been presented. In Figure 2.6 part of a positioning system for a radar antenna is shown. The word bond graph indicates the major subsystems to be considered, and the bonds with the effort and flow variables indicated introduce some variables that will be useful in characterizing the subsystems at a later stage in the analysis. You should be able to associate all the efforts and flows on the bond graph with physical quantities associated with the physical system being modeled. Try it.

In Figure 2.7 another example system is shown. Again, it is instructive to try to understand the effort and flow quantities associated with the bonds in the word bond graph. For example, what are the three efforts and three flows associated with the 3-port —Diff—? Can you see that —Wheel— is a 2-port that relates a torque and an angular speed to a force and a velocity? Do not be surprised if the construction of a word bond graph for a dynamic system seems less than



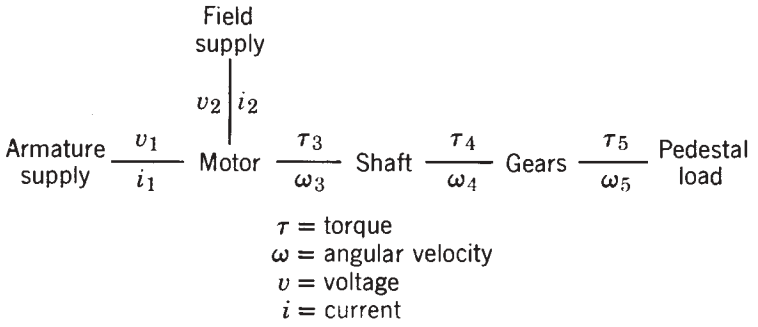


FIGURE 2.6. Schematic diagram and word bond graph for radar antenna pedestal drive system.

obvious at this stage. As you progress in the details of system modeling, it will become easier to recognize ports and bonds and, hence, multiport subsystems.

In Figure 2.7 you will notice that the influences of throttle position, clutch linkage position, and gear selector position are indicated using a bond with a *full* arrowhead. This notation, which is discussed in more detail in the next section, indicates that an influence on the system from its environment occurs at essentially zero power flow. In the present example, the driver of the car is part

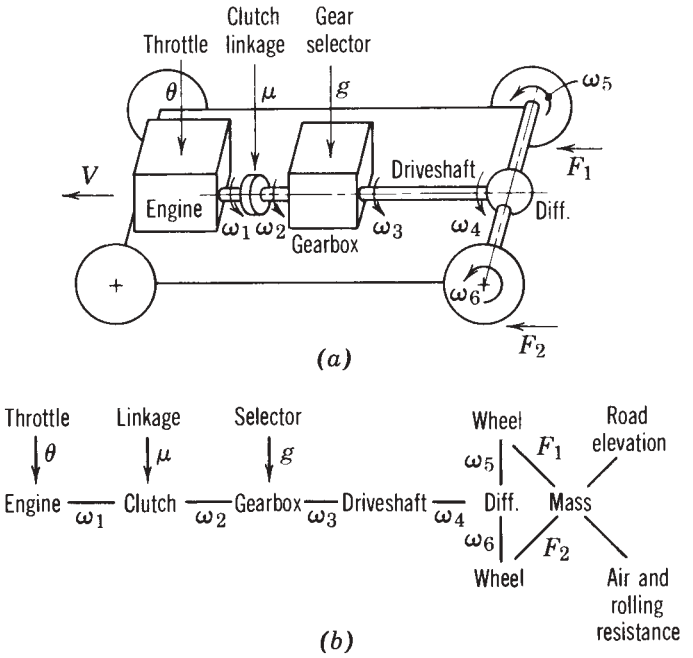


FIGURE 2.7. Automotive drive train example: (*a*) schematic diagram; (*b*) word bond graph.

of the environment of the car and can control the car using the accelerator pedal, clutch pedal, and gear shift using low power, as compared with the power present in the drive train. A bond with a full arrow is an *active bond*, and it indicates a signal flow at very low power. In the present case, we assume that the controls of the car can be moved by the driver at will, and our dynamic model need not concern itself with the forces required to move the controls. A word bond graph is useful for sorting true power interactions (which involve action and reaction) from the one-way influences of active bonds.

Bond graphs will subsequently be used to model subsystems in detail internally. For this purpose, a set of basic multiport elements denoted not by words, but by letters and numbers, will be developed in the next chapter. Ultimately, detailed bond graphs must be substituted for the multiports designated by words in a word bond graph. From a sufficiently detailed bond graph, state equations may be derived using standard techniques or computer simulations of the system can be made. Several computer programs will accept a wide variety of bond graphs directly and produce either state equations for subsequent analysis or system response predictions through computer simulation by solving the state equations numerically. In addition, some types of analyses can be performed on a bond graph without either writing the state equations or using a computer.

2.4 INPUTS, OUTPUTS, AND SIGNALS

Multiport subsystem characteristics typically are determined by a combination of experimental and theoretical methods. It might be fairly easy to compute the moment of inertia of a rotor, for example, merely by knowing the density of the material of which the rotor was made and having a drawing of the part, but to predict the port characteristics of a fan in great detail by theoretical means might be much more difficult than by measuring the characteristics using a prototype. In performing experiments on a subsystem, the notions of *input* and *output* or, equivalently, *excitation* and *response*, arise. The same concepts will carry over when "mathematical" models of subsystems are assembled into a system model.

In performing experiments on a multiport, one must make a decision about what is to be done at the ports. At each port, both an effort and a flow variable exist, and one can control either one but not both of these variables simultaneously. As an example, consider the problem of determining the steady-state characteristics of a dc motor such as the one shown in Figures 2.3–2.5.

Figure 2.8*a* shows a sketch of equipment that could be used in experimenting on the motor. The dynamometer is supposed to be capable of setting the speed of the motor regardless of the torque delivered by the motor. This speed, ω, is then an *input variable* to the motor. The torque being delivered by the motor is then measured by means of a torque gage. The torque, τ, is thus an *output variable* of the motor. Note that it is not possible to adjust the dynamometer so that both torque and speed have arbitrary values. The nature of the experiment is to discover what the motor torque is at a given speed.

Similarly, if voltages are supplied to the two electrical ports, that is, if voltages are input variables, then the motor responds with measurable currents that are output variables of the motor. Figure 2.8*b* is an attempt to use lines and arrows to show which quantities are inputs to the motor and which are outputs. Figure 2.8*b* is a simple example of a *block diagram*, in which lines with arrows indicate the direction of flow of *signals*. For multiports each port or bond has both an effort and a flow, and when these two types of variables are represented as paired signals, it is possible for only one of these signals to be an input and the other will be an output.

To know which of the effort and flow signals at a port is the input of the multiport, only one piece of information must be supplied to Figure 2.3*c*, 2.3*d*, or 2.5. This is because if either the effort or flow variable is an input, the other must be an output. In bond graphs the way in which inputs and outputs are specified is by means of the *causal stroke*. The causal stroke is a short, perpendicular line made at one end of a bond or port line. It indicates the direction in which the effort signal is directed. (By implication, the end of a bond that does not have a causal stroke is the end toward which the flow's signal arrow points.) In Figure 2.8*c* causal strokes have been added to the multiport representation of Figure 2.3*d*. By comparing Figures 2.8*a–c*, all of which contain the same information regarding input and output variables, the meaning of causal strokes may be appreciated.

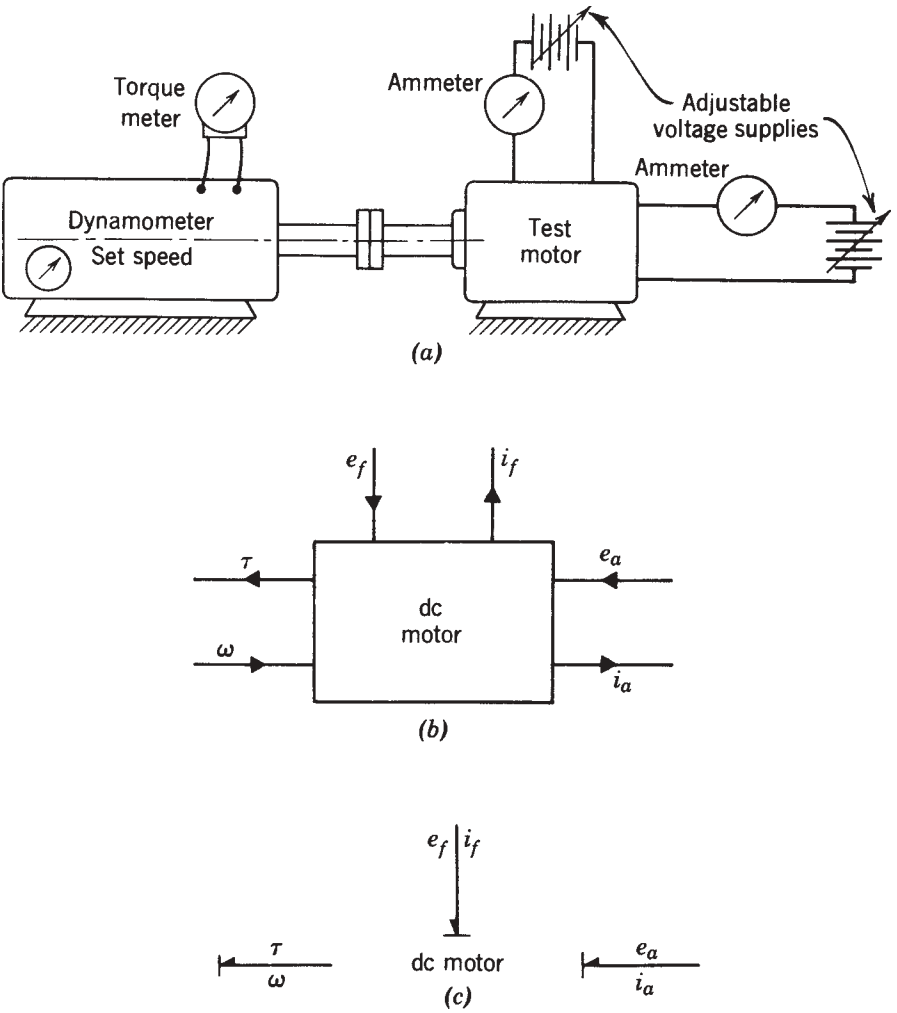


FIGURE 2.8. Experimental testing of a dc motor: (*a*) sketch of the test apparatus; (*b*) block diagram showing input and output signal flow; (*c*) causal strokes added to multiport representation.

The meaning of causal strokes is summarized in Figure 2.9, in which both bond graphs and block diagrams are shown. Note that the half-arrow sign convention for power flow and the causal stroke are completely independent. Thus, using A and B to stand for subsystems as in Figure 2.9, all the following combinations of sign convention and causal strokes are possible: $A \vdash\!\!\!\rightarrow B$, $A \vdash\!\!\!- B$, $A \dashv\!\!\!\rightarrow B$, $A \dashv\!\!\!- B$. The study of input–output *causality*, which is a uniquely useful feature of bond graphs, will be dealt with at length in succeeding chapters.

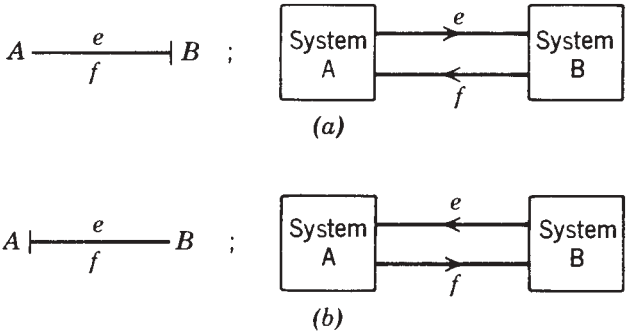


FIGURE 2.9. The meaning of causal strokes: (*a*) effort is output of *A*, input to *B*; flow is output of *B*, input to *A*; (*b*) effort is output of *B*, input to *A*; flow is output of *A*, input to *B*.

Finally, we come to the question of pure signal flow, or the transfer of information with negligible power flow, which we already encountered in the example of Figure 2.7. Multiports in principle all transmit finite power when interconnected. This is correlated with the fact that both an effort and a flow variable exist when multiports are coupled. Thus, systems are interconnected by the matching of a *pair* of signals representing the power variables.

In many cases, however, systems are so designed that only one of the power variables is important, that is, so that a single signal is transmitted between two subsystems. For example, an electronic amplifier may be designed so that the voltage from a circuit influences the amplifier, but the current drawn by the amplifier has virtually no effect on the circuit. Essentially, the amplifier reacts to a voltage but extracts negligible power in doing so, as compared with the rest of the power levels in the circuit. No information can really be transmitted at exactly zero power, but, practically speaking, information can be transmitted at power levels that are negligible, as compared with other system power levels. Every instrument is designed to extract information about some system variable without seriously disturbing the system to which the instrument is attached. An ideal ammeter indicates current but introduces no voltage drop, an ideal voltmeter reads a voltage while passing no current, an ideal pressure gage reads pressure with no flow, an ideal tachometer reads angular speed with no added torque, and the like. When an instrument reads an effort or flow variable, but with negligible power, there is a signal connection between subsystems without the back effect associated with finite power interaction.

The block diagrams of control engineering or the signal flow graphs that were developed first for electrical systems ideally show signal coupling. As Figure 2.8*b* shows, when multiports are considered, power interactions require a pair of bilaterally oriented signals. The bond graph, in which each bond implies the existence of both an effort and a flow signal, is a more efficient way of describing multiports than are block diagrams or signal flow graphs. Yet when the system is

dominated by signal interactions due to the presence of instruments, isolating amplifiers, and the like, then either an effort or a flow signal may be suppressed at many interconnection points. In such a case, a bond degenerates to a single signal and may be shown as an *active bond*. The notation for an active bond is identical to that for a signal in a block diagram; for example, $A \xrightarrow{e} B$ indicates that the effort, e, is determined by subsystem A and is an input to subsystem B. Normally, this situation would be indicated by $A \underset{f}{\overset{e}{\dashv}} B$, in which the flow, f, is determined by B and is an input to A. When e is shown as a signal (by means of the full arrow on the bond) or, in other words, an activated bond, the implication is that the flow, f, has a negligible effect on A.

When automatic control systems are added to physical systems, the control systems usually receive signals by means of nearly ideal instruments and affect the systems through nearly ideal amplifiers. The use of active bonds for such cases simplifies the analysis of the systems. Notice that in using bond graphs, one always assumes that multiports are coupled with both forward and backward effects unless a specific modeling decision has been made that a back effect is negligible and an active bond is specified. As we will see in Chapter 6, active bonds and block diagrams are often useful in studying control systems but for systems that involve real power interactions, the bond graph provides an ideal means for assuring the action–reaction effects on power bonds are properly represented.

In the following chapter we begin the detailed modeling of subsystems by considering a basic idealized set of multiports that can be assembled to model the pertinent physical effects in a subsystem. At this detailed level, physical parameters must be estimated and the rules of causality among ideal multiports must be discovered and obeyed in assembling the subsystem model from elemental multiports. As this process goes on, the notation and concepts briefly introduced in this chapter will become more familiar and useful.

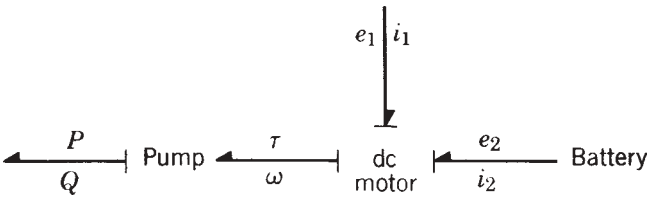
PROBLEMS

2-1. Construct four tetrahedra of state similar to that shown in Figure 2.2 for the following four physical domains: mechanical translation, mechanical rotation, hydraulic systems, electrical systems. Replace e, f, p, and q with their physical counterpart variables, and list the dimensions of each variable.

2-2. For each multiport in Figure 2.1 construct a word bond graph similar to that shown in Figure 2.3. Construct several systems by bonding several multiports together.

2-3. Suppose a pump was tested by running it at various speeds and measuring the volume flow rate and torque for various pressures at the pump outlet. Draw a schematic diagram, block diagram, and bond graph for the pump test analogous to those shown in Figure 2.8 for an electric motor test.

2-4. If the system of Figure 2.5 had the causality show how the signals flow by using a block diagram of the type used in Figure 2.8*b* for each multiport. See also Figure 2.9.



2-5. Apply causal strokes in an arbitrary manner to each bond in the bond graph of Figure 2.6. Construct an equivalent block diagram for this system using one block for each multiport as in Figure 2.8*b*. Indicate the signal flow directions that correspond to your causal marks as in Figure 2.9.

2-6. Repeat Problem 2-5 for the system of Figure 2.7. (Note that active bonds act just like one-way signal flows in a block diagram, but that normal bonds each result in two signal flows for e and f.)

2-7. Consider the system of Problem 2-4. Identify the system input variables that come from the environment of the system, given the causal stroke pattern indicated. What variables are indicated as system outputs (and hence inputs to the environment)?

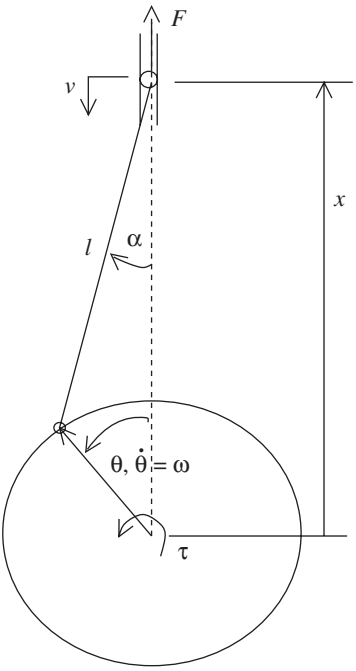
2-8. How long would a 100-W light bulb have to burn to use up the same energy that would be required to raise a 10-kg mass 30 m up in the earth's gravity field?

2-9. Represent an electric drill as a multiport. Consider the switch position influence as occurring on an active bond. Apply causal strokes to your bond graph, assuming that the drill is plugged into a 100-V outlet and that the torque is determined by the material being drilled. Show a block diagram for the drill corresponding to your choice of causality at the ports.

2-10. If a positive-displacement hydraulic pump is 100% efficient (so that the mechanical power is always instantaneously equal to the hydraulic power) and if a torque of 5 N-m produces a pressure of 7.0 MPa, what is the relationship between volume flow and angular speed? (7.0 MPa $= 7.0 \times 10^6$ N/m^2.)

2-11. The slider-crank mechanism is the fundamental kinematic device in virtually all internal combustion engines. This device relates the rotational motion of the crankshaft to the reciprocating motion of the piston. In its most idealized representation, the slider-crank is massless, frictionless, and constructed from rigid components. Under these assumptions, the device is power conserving, in that $\tau\omega = Fv$, where τ is the torque

on the crankshaft, F is the force on the end of the connecting rod, ω is the angular velocity of the crank, and v is the velocity of the rod end. If we can derive how v and ω are related, then we automatically know how F and τ are related.



(a) Slider-crank device

$$\frac{\tau}{\omega} \qquad \text{Slider-crank device} \qquad \frac{F}{v}$$

(b) Word bond graph

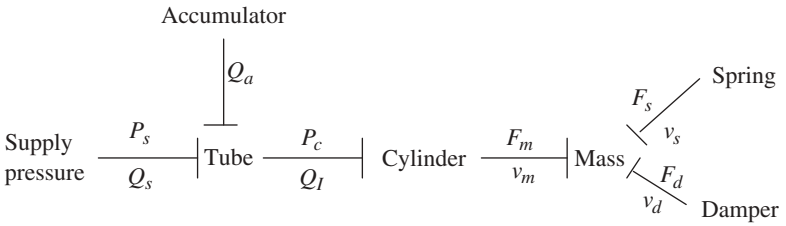
As a word bond graph, the slider-crank will be represented as indicated in the figure. We will soon learn that this device is a modulated transformer. Derive the relationship between v and ω. Here is some help:

$$x = R\cos\theta + l\cos\alpha$$
$$l\sin\alpha = R\sin\theta$$

Solve the second equation for sin α and then use $\cos\alpha = \sqrt{1-\sin^2\alpha}$. Substitute into the first equation. Then differentiate the result to relate $\dot{x} = -v$ to $\dot{\theta} = \omega$. If you complete these steps, you will have derived $v = m(\theta)\omega$, where $m(\theta)$ is a function of the crank angle. Since this device is power conserving, we immediately know that $\tau = m(\theta)F$.

Now try to derive the relationship between τ and F by using force and moment equilibrium conditions. You will find this far more difficult than deriving the velocity–angular velocity relationship.

2-12. The hydraulic system from Problem 1-9 has the word bond graph shown in the following figure. Causality has been assigned to the bonds. Identify the inputs and outputs for each element.



3

BASIC BOND GRAPH ELEMENTS

In Chapter 2 real devices were considered as subsystems from the point of view of power exchanges and external port variables. In this chapter a basic set of multiport elements is defined that can be used to model subsystems in detail. These elements function as components of subsystem and system models and are, in many cases, idealized mathematical models of real components, such as resistors, capacitors, masses, springs, and pipes. In other cases, however, the basic elements are used to model *physical effects* in a device and cannot be put into a one-to-one correspondence with physical components of the device. For example, one might create a model of an electrical or fluid transmission line using a finite collection of resistance, capacitance, and inertia elements, even though in the real device the effects being modeled are distributed along the transmission line and not concentrated into lumps as in the model.

Using bond graphs and the classification of power and energy variables presented in the previous chapter, it turns out that only a few basic types of multiport elements are required in order to represent models in a wide variety of energy domains. The bond graph notation often allows one to visualize aspects of the system more easily than would be possible with just the state equations or with some other graphical notation designed for a single energy domain or for signal flow rather than power flow. The search for a bond graph model of a complex system frequently increases one's understanding of the physical system.

3.1 BASIC 1-PORT ELEMENTS

A 1-port element is addressed through a single power port, and at the port a single pair of effort and flow variables exists. In principle, a 1-port can be a very complex subsystem. For example, an ordinary electrical wall outlet can sometimes be represented as 1-port element in a system analysis. The port actually connects

Copyright © 2012 John Wiley & Sons, Inc.

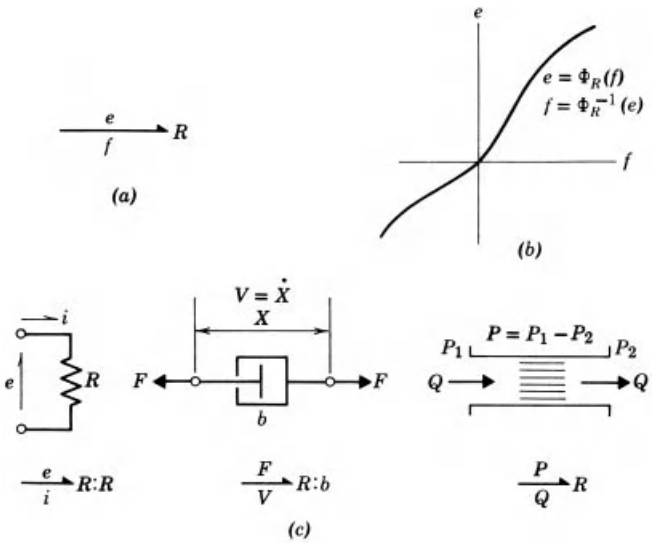


FIGURE 3.1. The 1-port resistor: (*a*) bond graph symbol; (*b*) defining relation; (*c*) representations in several physical domains.

to a vast network of power generation and distribution equipment, yet from the point of view of a system model, a relatively simple characterization of what is behind the wall outlet as a 1-port may suffice.

Here we first deal with the most primitive 1-ports. We consider, in order, 1-port elements that dissipate power, store energy, and supply power.

The *1-port resistor* is an element in which *the effort and flow variables at the single port are related by a static function*. Figure 3.1 shows the bond graph symbol for the resistor, a typical graph of the constitutive relation between e and f, and sketches of resistors in several energy domains. The electrical resistor is an R-element if it can be characterized by a volt–current constitutive relationship such as

$$e = Ri.$$

Since e is an effort variable and i is a flow variable, this constitutive relationship exactly fits our definition of a linear 1-port resistor.

The mechanical dashpot is a 1-port resistor for the same reason as in the electrical resistor. If an ideal dashpot is characterized by a force–velocity relationship such as

$$F = bV,$$

where b is the dashpot constant, then it is represented as a 1-port R element. Since F is an effort and is V a flow, the constitutive relationship also fits our definition of a 1-port resistor.

The hydraulic example is a 1-port resistor because it is characterized by a pressure–volume flow rate relationship and pressure and volume flow rate are bond graph effort and flow variables. In most hydraulic cases the effort–flow relationship is more complicated than the first two examples. This is dealt with in great detail in Chapters 4 and 12; here we just specify that for turbulent flow through a restriction, the pressure–flow relationship is

$$P_1 - P_2 = \frac{1}{2A^2}\rho Q|Q|$$

or

$$Q = A\sqrt{\frac{2}{\rho}|P_1 - P_2|}\,\mathrm{sgn}(P_1 - P_2),$$

where A is the flow area. These relationships are effort–flow relationships and again fit our definition of a one-port resistor.

Usually, resistors dissipate energy; that is, power flows into the resistor but never comes out of it. From the point of view of a system connected to a resistor, over time, energy seems to disappear into a resistor. This must be true for simple electrical resistors, mechanical dampers or dashpots, porous plugs in fluid lines, and other analogous elements. Noting from Figure 3.1*a* that power flows *into* the port when the product of e and f is positive according to the sign convention shown, we may deduce that power is always dissipated if the defining constitutive relation between e and f lies only in the first and third quadrants of the e–f plane as shown in Figure 3.1*b*, for then the product ef is positive when both e and f are positive or when both are negative. Because the resistance function cannot lie in the second or fourth quadrant, the law must pass through the origin.

When the relation between e and f for a 1-port resistor plots as a curved line as in Figure 3.1*b*, then the resistor is a *nonlinear element*. If the relation is a straight line, then it is a *linear element*. In the special case of a linear element, a coefficient, the *resistance*, or its inverse, the *conductance*, may be defined. These parameters are actually just the slopes of the straight line laws when plotted as e versus f or f versus e.

When a resistive element is assumed to be linear, it is conventional to indicate this on the bond graph by appending a colon (:) next to the $-R$ and noting the physical symbol for the resistance parameter. This is done in Figure 3.1 for the electrical resistance and the mechanical dashpot. For the hydraulic resistor, no parameter is indicated since this is a nonlinear element and no single resistance parameter can be identified.

Table 3.1 shows the resistor constitutive laws in general form and in specific form for several energy domains. Note that for power-dissipating linear resistors, with the sign convention shown in Figure 3.1 and Table 3.1, the general resistance and conductance parameters, R and G, respectively, are positive, as are the corresponding parameters in the specific energy domains.

For simplicity, we establish the following arbitrary but useful rule: For passive (power dissipating) resistors, establish the power sign convention by means of

TABLE 3.1. The 1-Port Resistor, $\overset{e}{\underset{f}{\rightharpoonup}} R$

	General Relation	Linear Relation	SI Units for Linear Resistance Parameter
Generalized variables	$e = \Phi_R(f)$	$e = Rf$	$R = e/f$
	$f = \Phi_R^{-1}(e)$	$f = Ge = e/R$	
Mechanical translation	$F = \Phi_R(V)$	$F = bV$	$b = \text{N-s/m}$
	$V = \Phi_R^{-1}(F)$		
Mechanical rotation	$\tau = \Phi_R(\omega)$	$\tau = c\omega$	$c = \text{N-m-s}$
	$\omega = \Phi_R^{-1}(\tau)$		
Hydraulic systems	$P = \Phi_R(Q)$	$P = RQ$	$R = \text{N-s/m}^5$
	$Q = \Phi_R^{-1}(P)$		
Electrical systems	$e = \Phi_R(i)$	$e = Ri$	$R = \text{V/A} = \Omega\,(\text{ohm})$
	$i = \Phi_R^{-1}(e)$	$i = Ge$	

a half-arrow pointing *toward the resistor*. Then linear resistance parameters will be positive, and nonlinear relations will fall in the first and third quadrants of the e–f plane.

Since linear models are of great usefulness in certain fields (vibrations and electric circuits, for example), the linear versions of resistance relations in various energy domains are shown in Table 3.1 with the same notation employed in Chapter 2. The units of the linear resistance parameter are simply the units of effort divided by the units of flow. The units displayed in Table 3.1 are worth studying, since many of them may not be familiar. The only resistance unit dignified with its own name is the electrical ohm.

Next consider a 1-port device in which a static constitutive relation exists between an *effort* and a *displacement*. Such a device stores and gives up energy without loss. In bond graph terminology, an element that relates e to q is called a *1-port capacitor* or *compliance*. In physical terms, a capacitor is an idealization of such devices as springs, torsion bars, electrical capacitors, gravity tanks, and hydraulic accumulators. The bond graph symbol, the defining constitutive relation, and some physical examples are shown in Figure 3.2.

As with the 1-port resistor, there are idealized linear compliance elements as well as nonlinear ones. In Figure 3.2*b*, a general nonlinear constitutive e, q relationship is shown. If the element can be assumed linear, then the e versus q curve will be a straight line and a compliance parameter can be defined such that $e = q/C$. Note that it is customary to define the linear compliance relationship using the inverse of the slope of the e versus q curve. The reason for this will become clear when more physical elements are presented in the next chapter. For the linear case, it is customary to indicate the compliance parameter on the bond graph as shown in Figure 3.2.

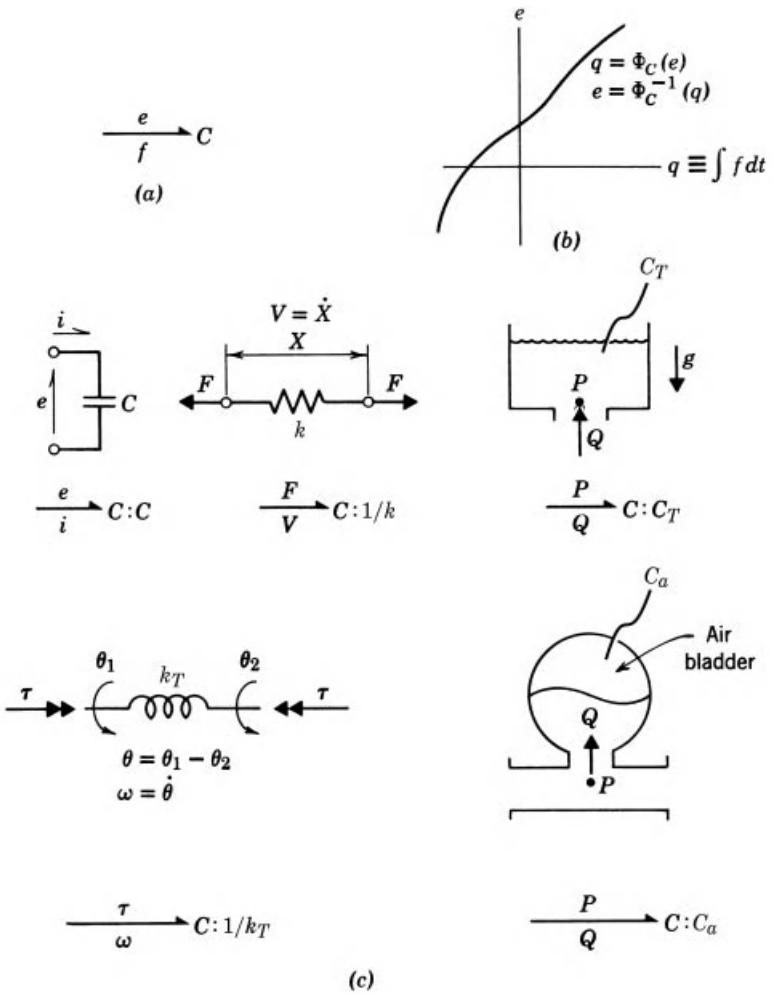


FIGURE 3.2. The 1-port capacitor. (*a*) Bond graph symbol; (*b*) defining relation; (*c*) representation in several physical domains.

The electrical capacitor, of capacitance C farads, is a compliance element because its idealized behavior is

$$e = \frac{q}{C},$$

where $q = \int i\,dt$ is the charge on the capacitor. This fits perfectly with our definition of a linear 1-port capacitor. The spring of stiffness, k, is a 1-port capacitor because it is characterized by

$$F = kx,$$

TABLE 3.2. The 1-Port Capacitor, $\underset{f=\dot{q}}{\overset{e}{\rightharpoonup}} C$

	General Relation	Linear Relation	SI Units for Linear Capacitance Parameter
Generalized	$q = \Phi_C(e)$	$q = Ce$	$C = q/e$
	$e = \Phi_C^{-1}(q)$	$e = q/C$	$1/C = e/q$
Mechanical	$X = \Phi_C(F)$	$X = CF$	$C = \text{m/N}$
Translation	$F = \Phi_C^{-1}(X)$	$F = kX$	$k = \text{N/m}$
Mechanical	$\theta = \Phi_C(\tau)$	$\theta = C\tau$	$C = \text{rad/N-m}$
Rotation	$\tau = \Phi_C^{-1}(\theta)$	$\tau = k\theta$	$k = \text{N-m/rad}$
Hydraulic	$V = \Phi_C(P)$	$V = CP$	$C = \text{m}^5/\text{N}$
Systems	$P = \Phi_C^{-1}(V)$	$P = V/C$	
Electrical	$q = \Phi_C(e)$	$q = Ce$	$C = \text{A-s/V}$
Systems	$e = \Phi_C^{-1}(q)$	$e = q/C$	$= \text{farad (F)}$

where $x = \int V\,dt$ is the relative displacement across the spring. This definition fits the general definition of a 1-port capacitor and the specific definition of a linear 1-port C-element as indicated in Table 3.2. In this case, as indicated on the bond graph, the compliance parameter is $C = 1/k$.

A water storage tank is discussed in the next chapter, and it is ideally a linear compliance element. The torsional spring of stiffness, k_τ, N-m/rad, is a linear 1-port compliance for the identical reason as in the linear spring. The compliance parameter for the torsional spring is $1/k_\tau$, as indicated on the bond graph of Figure 3.2.

An air bladder is also discussed in the next chapter. There are circumstances in which it may be modeled as a linear C-element, at least for small pressure excursions. However, in general, compressing air is a nonlinear process. If the process was isentropic, then the behavior of the air bladder could be characterized by

$$P = \frac{P_0 V_0^\gamma}{V^\gamma},$$

where $V = \int Q\,dt$, P_0, V_0 are initial pressure and volume in the bladder, and γ is the ratio of specific heats for air. This behavior is nonlinear, but still fits our general definition of a 1-port capacitor as indicated in Table 3.2. Thus, for the case here, the air bladder is a compliance element, but it is not a linear one, and no compliance parameter can be identified nor indicated on the bond graph.

Note that when a sign convention similar to that used for the resistor, namely, $\rightharpoonup C$, is used for the C-element, then ef represents power flowing *to* the capacitor and

$$\mathbf{E}(t) = \int_0^t e(t) f(t)\,dt + \mathbf{E}_0 \tag{3.1}$$

represents the energy stored in the capacitor at any time t. The energy stored initially at $t = 0$ (if any) is called $\mathbf{E}_0$.

Since from Eq. (2.2b) the displacement q is defined so that $f\,dt \equiv dq$, and the constitutive relation of a C-element implies that e is a function of q, $e = e(q)$, then Eq. (3.1) can be rewritten as

$$\mathbf{E}(q) = \int_{q0}^{q} e(q)\,dq + \mathbf{E}_0 \tag{3.2}$$

where $\mathbf{E}_0$ is the energy stored when $q = q_0$. Usually, it is convenient to define the energy stored to be zero when the effort is zero. Then, if q_0 is that value of q at which $e = 0$, and $\mathbf{E}_0 = 0$, Eq. (3.2) may be written as

$$\mathbf{E}(q) = \int_{q_0}^{q} e(q)\,dq. \tag{3.2a}$$

The operation indicated in Eq. (3.2a) may be interpreted graphically as shown in Figure 3.3. As q varies, the area under the curve of e versus q varies, and this area is equal to the stored energy, $\mathbf{E}$. The *conservation of energy* for —C is almost obvious. If q goes from q_0 to $\hat{q}$ as in Figure 3.3*a*, then energy is stored; if q then ever returns to q_0, the shaded area disappears and all the stored energy disappears. The power flow into the port, which resulted in the storage of energy, reverses and power flows out of the port. During the process, no energy is lost; in other words, energy is conserved.

Table 3.2 summarizes the relationships characterizing capacitors. The units for linear capacitance parameters are given, and again it may be noted that only the electrical unit is given a name, the farad. For linear mechanical systems, it is common to use the *spring constant*, k, rather than the *compliance*, $C \equiv 1/k$, which is analogous to the electrical capacitance, C, and the parameter C in generalized variables. In mixed electrical–mechanical systems, one must simply be careful to note whether a numerical parameter corresponds to C or the inverse of C in a bond graph. Once again, the reader is urged to study the units shown, since some units will probably be unfamiliar.

A second energy-storing 1-port arises if the momentum p is related by a static constitutive law to the flow f. Such an element is called an *inertia* in bond graph terminology. The bond graph symbol for an inertia, the constitutive relation, and several physical examples are shown in Figure 3.4. The inertia is used to model inductance effects in electrical systems, and mass or inertia effects in mechanical or fluid systems.

The 1-port inertia is characterized by an f, p relationship as indicated in Figure 3.4*b*. If the relationship is linear, then it will plot as a straight line and the constitutive relationship will have the form $f = p/I$, where I is the inertia parameter and $p = \int e\,dt$. Note that, as with the linear compliance element, it is customary to define the inertance parameter as the inverse of the slope of the linear relationship of f versus p. Shown in Figure 3.4 is an electrical inductor

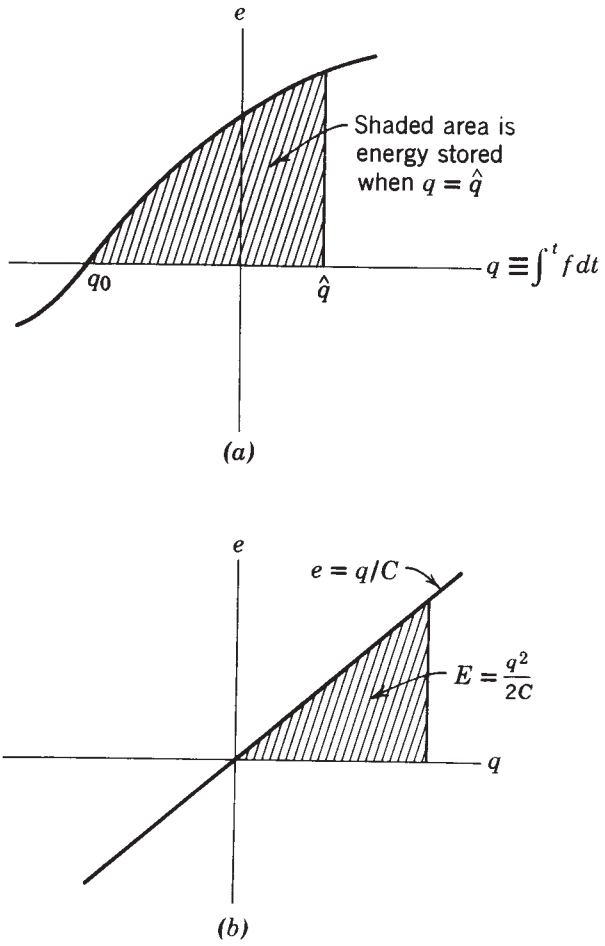


FIGURE 3.3. Area interpretation of stored energy for a 1-port capacitor: (*a*) nonlinear case; (*b*) linear case.

with inductance L, a mass m, a section of fluid-filled pipe with fluid inertia, I_f, and a rotating disk with moment of inertia, J. These are all examples of linear 1-port inertia elements. The inductor is ideally represented by the constitutive relationship $i = \lambda/L$, where $\lambda = \int e\,dt$; the mass is represented by $V = p/m$, where $p = \int F dt$; and the rotating disk has the ideal behavior, $\omega = p_\tau/J$, where $p_\tau = \int \tau\,dt$. All these elements fit exactly our definition of a 1-port inertia as shown in Table 3.3. The fluid inertia is also a linear 1-port inertia, which is covered thoroughly in Chapter 4.

Using the sign convention $\rightharpoonup I$, the power flowing into the inertia is given by the expression in Eq. (3.1). In the present case, Eq. (2.2a) allows us to write

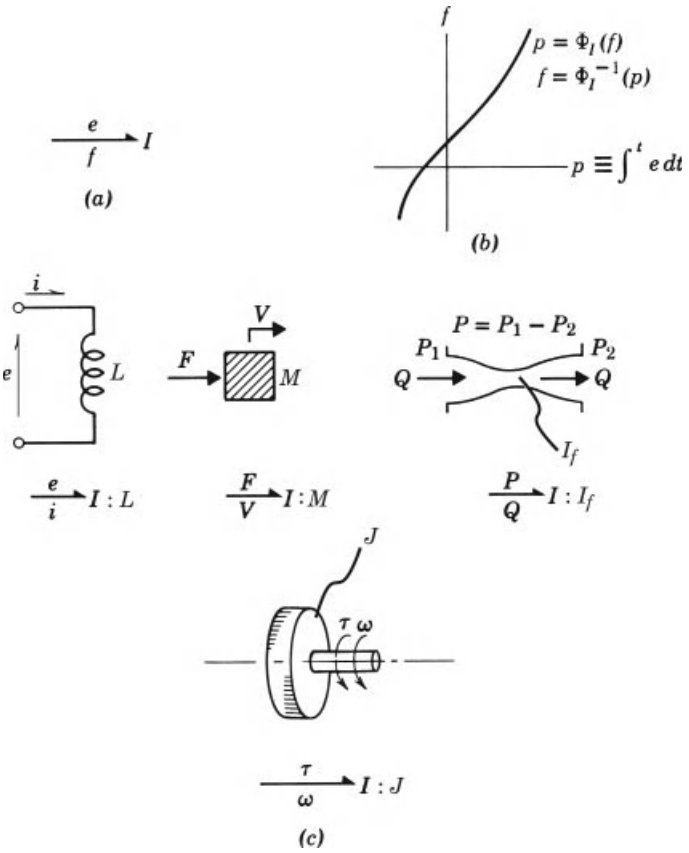


FIGURE 3.4. The 1-port inertia: (*a*) bond graph symbol; (*b*) defining relation; (*c*) representation in several physical domains.

$e\,dt \equiv dp$, and if $f = f(p)$, then Eq. (3.1) can be written thus:

$$\mathbf{E}(p) = \int_{p_0}^{p} f(p)\,dp + \mathbf{E}_0. \tag{3.3}$$

If the energy is defined to vanish when f vanishes and if p_0 corresponds to that point in the plot of f versus p at which $f = 0$, then

$$\mathbf{E}(p) = \int_{p_0}^{p} f(p)\,dp. \tag{3.3a}$$

The similarities between Eqs. (3.2) and (3.3) should be noted. Often the energy associated with a capacitor is called *potential energy*, whereas the energy associated with an inertia is called *kinetic energy*. These names are applied primarily to

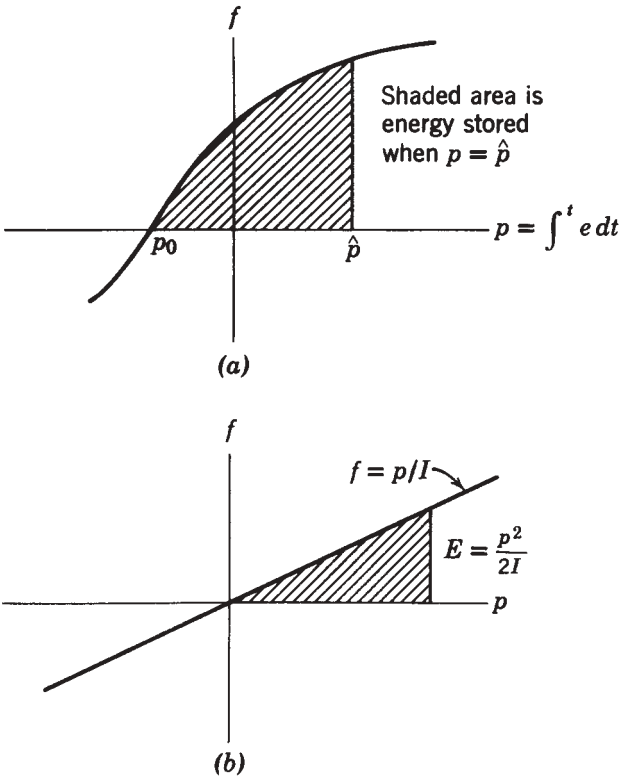


FIGURE 3.5. Area interpretation of stored energy for 1-port inertia. (*a*) Nonlinear case; (*b*) linear case.

mechanical systems. In electrical systems, the corresponding two forms of stored energy are sometimes called *electric* and *magnetic* energy.

As in the case of the capacitor, if the constitutive relation of the inertia is plotted, then there is an area interpretation of the stored energy. This interpretation is shown in Figure 3.5, and, again, you should be able to demonstrate that any energy stored in an I-element can be recovered without loss.

Table 3.3 shows the constitutive relations for inertias and gives units for inertance parameters for the linear case. Since most engineering work is accomplished successfully using Newton's law rather than the postulates of relativity, the relation between velocity and momentum is linear and the mass or moment of inertia is the inertia parameter. Although it is common to think of mass as a ratio of force to acceleration from the equation

$$F = ma, \quad a \equiv \dot{V}, \tag{3.4}$$

TABLE 3.3. The 1-Port Inertia, $\overset{e=\dot{p}}{\underset{f}{\rightharpoonup}} I$

	General Relation	Linear Relation	SI Units for Linear Inertance Parameter
Generalized	$p = \Phi_I(f)$	$p = If$	$I = p/f$
variables	$f = \Phi_1^{-1}(p)$	$f = p/I$	$1/I = f/p$
Mechanical	$p = \Phi_I(V)$	$p = mV$	$m = \text{N-s}^2/\text{m} = \text{kg}$
translation	$V = \Phi_1^{-1}(p)$	$V = p/m$	
Mechanical	$p_\tau = \Phi_I(\omega)$	$p_\tau = J\omega$	$J = \text{N-m-s}^2 = \text{kg-m}^2$
rotation	$\omega = \Phi_1^{-1}(p_\tau)$	$t\omega = p_\tau/J$	
Hydraulic	$p_p = \Phi_I(Q)$	$p_p = IQ$	$I = \text{N-s}^2/\text{m}^5$
systems	$Q = \Phi_1^{-1}(p_p)$	$Q = p_p/I$	
Electrical	$\lambda = \Phi_I(i)$	$\lambda = Li$	$L = \text{V-s/A}$
systems	$i = \Phi_1^{-1}(\lambda)$	$i = \lambda/L$	$= \text{henrys (H)}$

the table gives the fundamental definition of a mass according to

$$p \equiv mV \tag{3.5}$$

with

$$\dot{p} \equiv F. \tag{3.6}$$

Clearly, when Eq. (3.5) is differentiated with respect to time, and Eq. (3.6) is used, then Eq. (3.4) can be derived. If, on the other hand, Eq. (3.5) is replaced with a nonlinear relation,

$$p = \Phi_I(V) = \frac{mV}{(1 - V^2/c^2)^{1/2}}, \tag{3.7}$$

where m is the rest mass and c is the velocity of light, then Eqs. (3.5) and (3.6) hold for the special theory of relativity. See Reference [1, p. 19], for example. Thus, there is some justification for the general constitutive relations given for mechanical systems, even though engineering is overwhelmingly concerned with the linear case. For electrical systems, however, the relation between the flux linkage variable (the time integral of the voltage) and the current in an inductor is nonlinear in typical cases. The use of the linear parameter L is then the result of a modeling decision. It is more satisfactory to generalize $\lambda = Li$ to $\lambda = \Phi_I(i)$ with $\dot{\lambda} = e$ than to try to generalize $e = L\, di/dt$ to the nonlinear case.

As an aid in remembering the three 1-port relationships, the tetrahedron of state introduced in Figure 2.2 may be used. See Figure 3.6. We now know something about five of the six edges of the tetrahedron. The sixth edge, which stretches

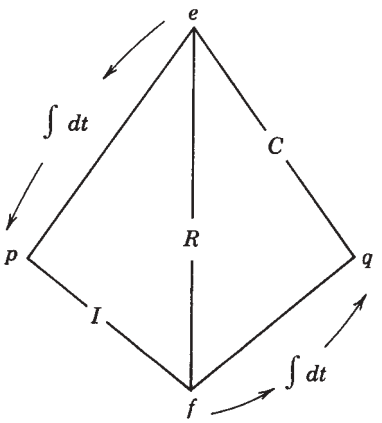


FIGURE 3.6. The three 1-ports placed on the tetrahedron of state according to the variables to which they relate.

between the vertices representing p and q, is hidden from view in Figure 3.6. This is just as well, since no basic element will relate p and q.*

Finally, two useful and rather simple 1-ports must be defined—the *effort source* and the *flow source*: the 1-port sources are idealized versions of voltage supplies, pressure sources, vibration shakers, constant-flow systems, and the like. In each case, an effort or flow is either maintained reasonably constant, independent of the power supplied or absorbed by the source, or constrained to be some particular function of time. As an example of a constant-effort source, consider the gravity force on a mass. Near the surface of the earth, this force is essentially independent of the velocity of the mass. As an example of a time-varying source, the electrical wall outlet will serve. The wall outlet enforces a sinusoidal voltage across the power cord wires of most small appliances. Over a reasonable range of currents, the voltage is independent of fluctuations in the current. Of course, the voltage is actually affected by large currents, and a fuse will blow to protect the circuits if very large currents build up, but this simply means that the real outlet is not modeled exactly by an ideal source of effort. Real sources can be modeled as combinations of effort and flow sources combined with other bond graph elements.

Table 3.4 presents bond graph symbols and the constitutive relations for sources. In this table, physical names are given for the respective energy domains. There are sources of velocity, S_V, sources of force, S_F, sources of pressure, S_P, and so forth. In general, it is best to use the generalized name for the effort

*One can, in fact, define an element corresponding to the hidden edge, the "memristor." While interesting and occasionally useful, memristors can be represented in terms of other elements to be introduced later, so the memristor will not be considered to be a basic element. See G. F. Oster and D. M. Auslander, "The memristor: a new bond graph element," *Trans. ASME, J. Dynamic Systems, Measurement, and Control*, **94**, Ser. G, no. 3, pp. 249–252 (Sept. 1972).

TABLE 3.4. The 1-Port Source Elements

	Bond Graph Symbol	Defining Relation
Generalized variables	$S_e \rightharpoonup$	$e(t)$ given, $f(t)$ arbitrary
	$S_f \rightharpoonup$	$f(t)$ given, $e(t)$ arbitrary
Mechanical translation	$S_F \rightharpoonup$	$F(t)$ given, $V(t)$ arbitrary
	$S_V \rightharpoonup$	$V(t)$ given, $F(t)$ arbitrary
Mechanical rotation	$S_\tau \rightharpoonup$	$\tau(t)$ given, $\omega(t)$ arbitrary
	$S_\omega \rightharpoonup$	$\omega(t)$ given, $\tau(t)$ arbitrary
Hydraulic systems	$S_P \rightharpoonup$	$P(t)$ given, $Q(t)$ arbitrary
	$S_Q \rightharpoonup$	$Q(t)$ given, $P(t)$ arbitrary
Electrical systems	$S_e \rightharpoonup$	$e(t)$ given, $i(t)$ arbitrary
	$S_i \rightharpoonup$	$i(t)$ given, $e(t)$ arbitrary

or flow source, S_e or S_f, and denote next to the generalized name the specific energy domain being represented. This is done by example in most of the following chapters. It should be further noted that the symbols *SE* and *SF* for effort and flow source are often used in computer programs.

Typically, source elements are thought of as supplying power to a system. This accounts for the sign-convention half-arrow shown, which implies that when $e(t)$ $f(t)$ is positive, power flows from the source to whatever system is connected to the source. Later we will encounter modeling situations in which it is convenient to have the half-arrow pointing in the opposite direction to the directions shown in Table 3.4.

Since a source maintains one of the power variables constant or a specified function of time no matter how large the other variable may be, a source can supply an indefinitely large amount of power. This is, of course, not a realistic assumption, and real devices are not really ideal sources even though they may be modeled approximately by sources. As an example, consider the problem of predicting the current flowing from a 12-V automotive battery into a variable resistor connected to the battery. Figure 3.7 shows a circuit diagram, a bond graph, and a plot of voltage versus current. This is a static system operating at points at which the source characteristic intersects the resistor characteristics. For small currents (or high values of the resistance R), the battery is almost a constant-voltage source. When the resistance is lowered toward zero, the predicted current approaches infinity. Actually, when the current gets large, the internal resistance in the battery reduces the voltage below the nominal 12 V. In fact, if the resistance approaches zero, as it will when a shorting bar is put across the battery terminals, the battery current will approach a finite value, labeled "short-circuit current" in Figure 3.7*c*. When more basic multiports have been defined, it will be possible to model the battery with an ideal source and a resistor in such a way that the actual characteristics in Figure 3.7 will be reproduced by the model. For now, we simply note that ideal sources are useful in modeling real devices but should

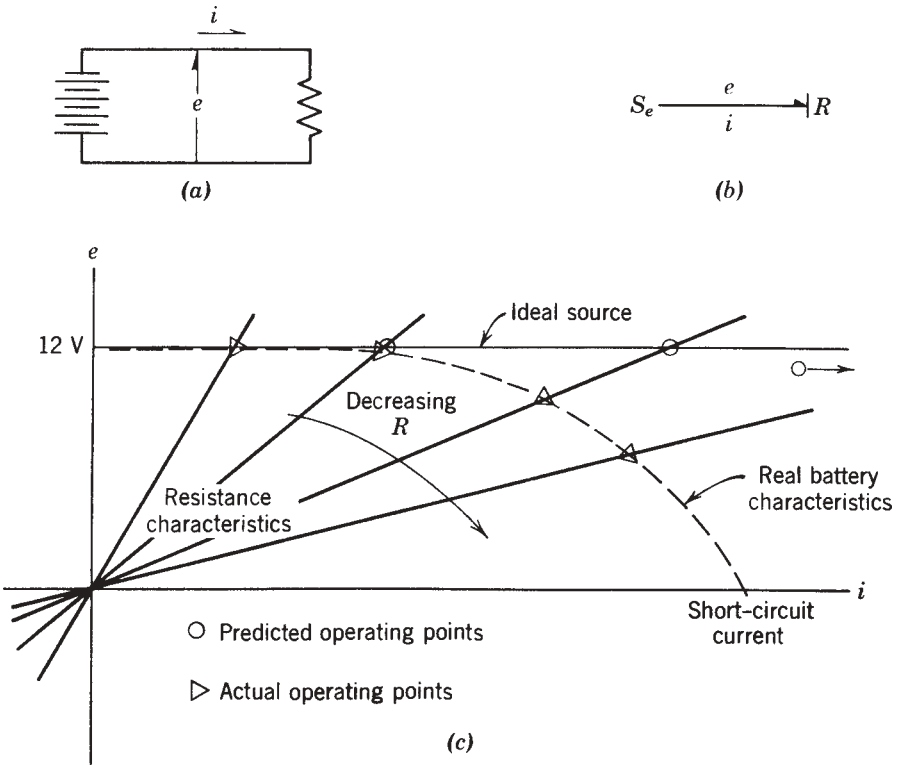


FIGURE 3.7. Study of a battery connected to a variable resistance: (*a*) electric circuit diagram; (*b*) bond graph; (*c*) plot of source, real battery, and resistance characteristics.

not be expected to be realistic models in all power ranges unless supplemented by other multiports.

A universe made up only of 1-ports would be very simple, since bond graphs more complicated than that in Figure 3.7*b* would be impossible. This leads one to anticipate that 1-ports are not the whole story. A logical next step is to consider 2-ports.

3.2 BASIC 2-PORT ELEMENTS

One might expect that it would be necessary to define more basic types of 2-ports than 1-ports, but, in fact, only two basic types of 2-ports need to be defined here. There are, of course, an unlimited number of 2-port subsystems, but we need to discuss here only those that cannot be modeled using the basic 1-ports of the previous section and other elements to be defined later. (Chapter 7 discusses the more advanced topic of 2-port and multiport versions of the 1-port R-, C-, and I-elements discussed above.)

The 2-ports to be discussed here are ideal in the specific sense that *power is conserved*. If any 2-port, —*TP*—, has the sign convention

$$\underset{f_1}{\overset{e_1}{\rightharpoonup}} TP \underset{f_2}{\overset{e_2}{\rightharpoonup}},$$

then power conservation means that at every instant of time

$$e_1(t) f_1(t) = e_2(t) f_2(t), \tag{3.8}$$

where the bonds are numbered so that all variables can be simply described as a type, e, f, p, or q with a bond number subscript.

The power sign convention implied in Eq. (3.8) and shown in the bond graph just above is a *through power* sign convention in the sense that power is thought of as flowing through the 2-port. Equation (3.8) states that whatever power is flowing into one side of the 2-port is simultaneously flowing out of the other side.

One way in which Eq. (3.8) can be satisfied is found in the 2-port known as a *transformer* and given in the bond graph symbol ⇀ *TF* ⇀. The constitutive laws of the ideal 2-port transformer are

$$e_1 = m e_2, \quad m f_1 = f_2, \tag{3.9}$$

in which the parameter m is called the *transformer modulus** and the subscripts 1 and 2 correspond to the two ports, as shown in Figure 3.8*a*. Multiplying the two equations in Eq. (3.9) yields Eq. (3.8). Note that Eqs. (3.9) and (3.8) both imply the use of the through sign convention shown in the figure.

Also shown in Figure 3.8 are a number of devices that in idealized form are modeled by transformers. In no case is the physical device exactly a transformer. For example, the lever in Figure 3.8*b* would only be a ⇀ *TF* ⇀ if it were massless, rigid, and frictionless. Similar restrictions can be made on the validity of the transformer as a model for the other physical devices. Actual models of the devices can be made using the ideal transformer and other multiports to account for nonideal effects if these effects are important to the system under study.

The lever is an ideal transformer because kinematics dictates that $(b/a)V_1 = V_2$ and moment equilibrium requires $F_1 = (b/a)F_2$. This is exactly the definition for the ideal transformer from Eq. (3.9) with the ratio (b/a) playing the role of the modulus m. If the velocity relationship had been derived first, then there would be no need to derive the force relationship, as it comes automatically due to the power conserving nature of the transformer.

Similarly, the gear set is an ideal transformer because kinematics dictates that $(r_1/r_2)\omega_1 = \omega_2$ and moment equilibrium requires $\tau_1 = (r_1/r_2)\tau_2$. This also fits exactly the definition from Eq. (3.9). If the angular velocity relationship had been

*There is an ambiguity here since one could just as well write Eq. (3.9) as, $m e_1 = e_2, f_1 = m f_2$, in which case the modulus would be defined as the inverse of the modulus in Eq. (3.9).

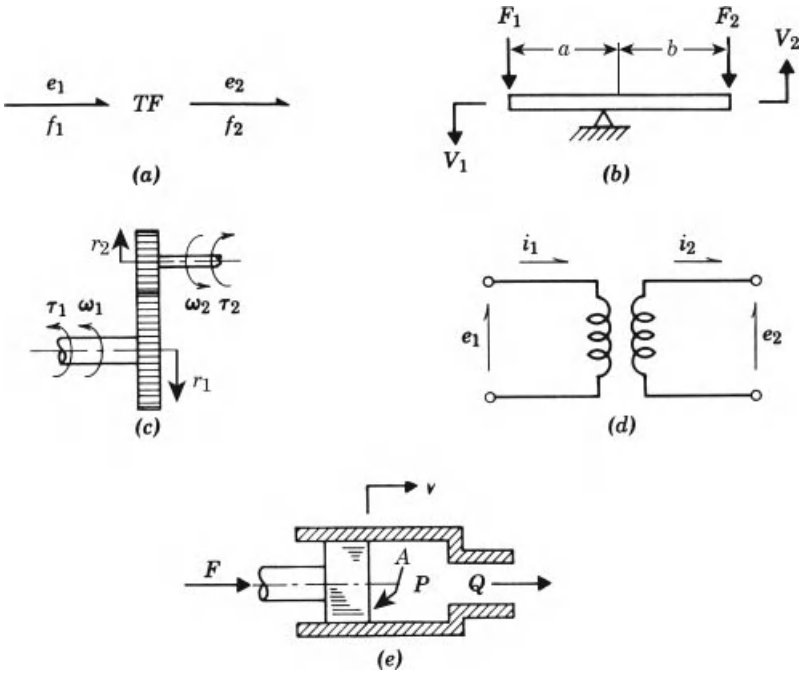


FIGURE 3.8. Transformers: (*a*) bond graph; (*b*) ideal rigid lever; (*c*) gear pair; (*d*) electrical transformer; (*e*) hydraulic ram.

derived first, then there would be no need to derive the torque relationship, as it also comes automatically due to the power conserving nature of the transformer. The electrical transformer from Figure 3.8*d* is a transformer in the bond graph sense because the voltage is stepped up or down as the current is stepped down or up according to the turns ratio of the windings of the transformer. In these first three cases, the transformer relates similar power variables on each side of the device. An extremely important use of the transformer is to cross from one energy domain to another. The hydraulic ram is a first example of this.

The hydraulic ram is shown in Figure 3.8*e*. Hydraulic power is *transduced* into mechanical power. In such a case, the *transformer* can also be called a *transducer*. The constitutive laws of the ideal version of this device are

$$F = AP, \quad AV = Q, \tag{3.10}$$

in which the area of the piston, A, functions as the transformer modulus, m, as it appears in Eq. (3.9). The two equations of (3.10) can be derived separately from physical considerations, or if one is derived, the other follows because of power conservation, $FV = PQ$. The fact that only a single modulus exists serves as a useful check on constitutive equations such as in Eq. (3.10).

Another way in which the power balance of Eq. (3.8) may be satisfied is embodied in the *gyrator*, which is symbolized thus: ⇀ GY ↽. If the bonds are numbered 1 and 2, the constitutive laws of the gyrator are

$$e_1 = rf_2, \quad rf_1 = e_2, \tag{3.11}$$

in which r is the *gyrator modulus* and the through sign convention of Figure 3.9*a* is implied. The modulus is called r because Eq. (3.11) reminds one of the 1-port linear resistance law, as shown in Table 3.1. In Eq. (3.11), however, the effort and flow at two *different* ports are statically related. Thus, a 1-port resistor dissipates power, whereas the 2-port gyrator conserves power, as can be seen by multiplying the two equations in (3.11) together.

Figure 3.9 shows some physical devices that are at least approximately gyrators. The electric circuit symbol is used to represent gyrators in electric network diagrams. Electric gyrators can be made using the Hall effect, and the gyrator is needed to model effects at microwave frequencies even though in that case it cannot be identified as a separate physical device.

Anyone who has played with a toy gyroscope has observed a gyrator. If the rotor of Figure 3.9*c* spins very rapidly, a gentle push in the direction of F_1 will yield a proportional velocity V_2. Similarly, a force F_2 will result in a velocity V_1. The counterintuitive behavior of the gyroscope is predicted by Eq. (3.11). For example, if the gravity force is in the direction of F_2, then the device precesses in a horizontal path. If a gyroscope rotates slowly, or if large disturbances are applied, the gyroscope must be modeled as a multidimensional rigid body. The bond graph, then, is much more complex than a simple ⇀ GY ↽, but even then

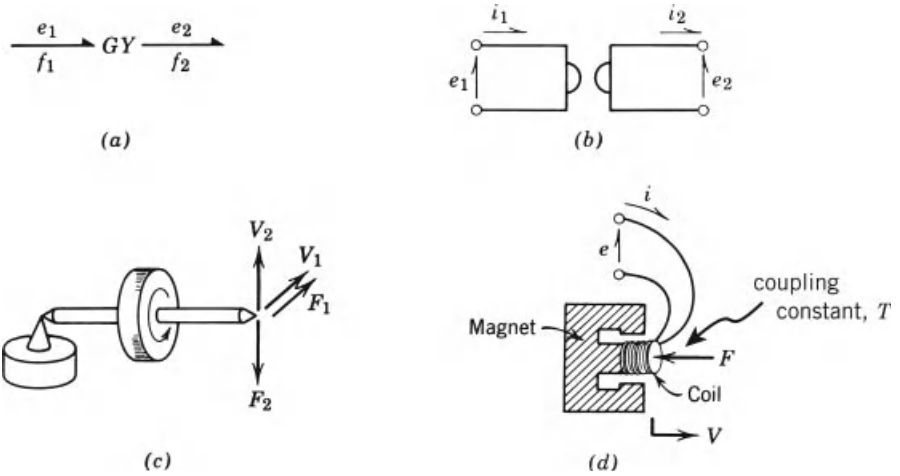


FIGURE 3.9. Gyrators: (*a*) bond graph; (*b*) symbol for electrical gyrator; (*c*) mechanical gyrator; (*d*) voice coil transducer.

the model does contain gyrators. For a restricted range of spin speeds and forces, the gyroscope is approximately a gyrator and gives the gyrator its name.

Figure 3.9*d* shows a useful transducer, which is a gyrator if certain nonideal effects may be neglected. It is the *voice coil* used in electrodynamic loudspeakers, vibration shakers, seismic mass accelerometers, and many other devices. The voice coil and other similar electromechanical conversion devices are covered thoroughly in Chapter 4. Here we simply specify its behavior as an example of a gyrator transducer.

The constitutive laws of the device are

$$e = TV, \quad Ti = F, \tag{3.12}$$

in which T plays the role of r in Eq. (3.11) and is the coupling constant for the device. The units of T might be thought of as volts per feet per second from one equation and pounds per ampere from the other equation. Thus, one might be tempted to measure the two versions of T that have such different units in two separate experiments. Actually, there is only a single T for a gyrator *when power is measured in consistent units.* The different units merely arise if power is measured in volt-amperes (watts) on one side of the transducer and in foot-pounds per second on the other.

Thus, one may measure T using one of the equations (3.12), invoke power conservation using the conversion factor between W and ft-lb/s to deduce the T in the other equation. The two values of T are numerically different *only* because of the choice of units. The recognition that the device is power conserving and representable as a gyrator helps one avoid the mistake of using two approximate numerical values for the two Ts in Eq. (3.12). Unless the two values of T are exactly correct, the model will inappropriately create or dissipate power.

An important advantage of the SI system (International System of Units) is that the numerical values of T in the two parts of Eq. (3.12) are identical because power in all physical domains is measured in watts. We recommend that the SI system always be used in order to avoid any unit inconsistencies, especially when multiple energy domains are being modeled. Practically, the SI system avoids the troublesome necessity of making unit conversions when power in different domains is involved.

The gyrator always seems to be a more mysterious element than the transformer. Before the significance of the gyrator was recognized, it was common to make equivalent electrical network diagrams for electromechanical or electrohydraulic systems using only transformers. This is not possible, in general, but in many special cases one may apply the analogy between electrical and mechanical or hydraulic variables so that a gyrator is treated as a transformer. In bond graph terms, the voice coil is a transformer if, for example, we call current an effort and voltage a flow. This switching of the identification of effort–flow variables is entirely unnecessary if one only recognizes that gyrators are really necessary in order to model systems containing devices such as those shown in Figures 3.9*b* and *c*, in general.

In fact, a gyrator is a more fundamental element than a transformer. Two gyrators cascaded are equivalent to a transformer:

$$\underset{f_1}{\overset{e_1}{\rightharpoonup}} GY_1 \underset{f_2}{\overset{e_2}{\rightharpoonup}} GY_2 \underset{f_3}{\overset{e_3}{\rightharpoonup}} = \underset{f_1}{\overset{e_1}{\rightharpoonup}} TF_3 \underset{f_3}{\overset{e_3}{\rightharpoonup}}$$

$$e_1 = r_1 f_2, \quad r_2 f_2 = e_3 \;\rightarrow\; e_1 = (r_1/r_2)e_3,$$

$$r_1 f_1 = e_2, \quad e_2 = r_2 f_3 \;\rightarrow\; (r_1/r_2) f_1 = f_3.$$

In contrast, cascaded transformers are equivalent only to another transformer

$$\underset{f_1}{\overset{e_1}{\rightharpoonup}} TF_1 \underset{f_2}{\overset{e_2}{\rightharpoonup}} TF_2 \underset{f_3}{\overset{e_3}{\rightharpoonup}} = \underset{f_1}{\overset{e_1}{\rightharpoonup}} TF_3 \underset{f_3}{\overset{e_3}{\rightharpoonup}}$$

$$e_1 = m_1 e_2, \quad e_2 = m_2 e_3 \;\rightharpoonup\; e_1 = m_1 m_2 e_3,$$

$$m_1 f_1 = f_2, \quad m_2 f_2 = f_3 \;\rightharpoonup\; m_1 m_2 f_1 = f_3.$$

Thus, one could, in principle, consider every transformer as a cascade combination of two gyrators and dispense with $\rightharpoonup TF \rightharpoonup$ as a basic 2-port. It is more convenient, however, to retain $\rightharpoonup TF \rightharpoonup$ as a basic bond graph element.

It is also important to realize that the gyrator essentially interchanges the roles of effort and flow. This may be seen by replacing r in Eq. (3.11) with unity. Then the effort at one port of the $\rightharpoonup GY \rightharpoonup$ is just the flow at the other, and vice versa. Thus, the combination $\rightharpoonup GY \rightharpoonup I$ is equivalent to $\rightharpoonup C$. To see this, recall that $\rightharpoonup I$ relates f to the integral of e, or p. After the gyrator is added and the roles of e and f are interchanged, the combination relates e and the integral of f or q at the external port. The element relating e to q is $\rightharpoonup C$. Similarly, $\rightharpoonup GY \rightharpoonup C$ is equivalent to $\rightharpoonup I$. Thus, one could in principle dispense with either $\rightharpoonup C$ or $\rightharpoonup I$ as a basic 1-port as long as $\rightharpoonup GY \rightharpoonup$ was available. Again, it is more convenient and natural to retain both $\rightharpoonup C$ and $\rightharpoonup I$ as basic 1-ports.

As an example of a deduction about a system based purely on its bond graph representation, consider the equivalence $I_1 \rightharpoonup GY \rightharpoonup I_2 = I_1 \rightharpoonup C$, in which $\rightharpoonup GY \rightharpoonup I_2$ has been replaced by $\rightharpoonup C$. An $\rightharpoonup I$ bonded to a $\rightharpoonup C$ is an oscillator. In physical terms, it could be a mass–spring or inductor–capacitor system. Thus, we see that $I_1 \rightharpoonup GY \rightharpoonup I_2$ (or, for that matter, $C_1 \rightharpoonup GY \rightharpoonup C_2$) will act just like an inertia–capacitor system.

Finally, there is a generalization of the transformers and gyrators discussed above, based on the curious fact that in both Eqs. (3.9) and (3.11) the power conservation between the two ports is maintained even when the moduli m and r are not constant. This gives rise to the *modulated transformer* and the *modulated gyrator*, denoted in bond graph symbolism by

$$\underset{f_1}{\overset{e_1}{\rightharpoonup}} \overset{\downarrow m}{MTF} \underset{f_2}{\overset{e_2}{\rightharpoonup}} \quad \text{and} \quad \underset{f_1}{\overset{e_1}{\rightharpoonup}} \overset{\downarrow r}{MGY} \underset{f_2}{\overset{e_2}{\rightharpoonup}}.$$

Note that m and r are shown as *signals* on an *activated bond*. This means that no power is associated with the changes in m and r, and $e_1 f_1$ is always exactly equal to $e_2 f_2$, as in the case of the constant modulus, $\rightharpoonup TF \rightharpoonup$ and $\rightharpoonup GY \rightharpoonup$.

Many physical devices may be modeled by the modulated 2-ports. For example, the electrical autotransformer contains a mechanical wiper, which, when moved, alters the turns ratio between the primary and secondary coils and thus changes the transformer ratio. This alteration takes no power (if we can assume mechanical friction is negligible), and for any wiper position the device essentially conserves electrical power.

Both gyrators in Figure 3.9*c* and *d* have moduli that can be changed without changing the fact that power is conserved at the two ports. For the gyroscope, a motor can be made to change the spin speed of the rotor, which changes r. Similarly, for the voice coil, if an electromagnet is substituted for a permanent magnet, then the transduction coefficient T in Eqs. (3.12) may be varied. At every instant, the power is conserved at the two ports, but the characteristics of the device change.

In mechanics, the *MTF* is particularly important and may be used to represent geometric transformations or kinematic linkages. As a simple example, consider the rotating arm shown in Figure 3.10. The arm is in equilibrium under the action of the torque τ and force F, and it provides a relation between θ and y or $\dot{\theta} \equiv \omega$ and $\dot{y} \equiv V_y$. Writing the displacement relation first,

$$y = l \sin\theta, \tag{3.13}$$

we can differentiate this to yield a constitutive relation for velocities,

$$\dot{y} = (l \cos\theta)\dot{\theta}$$

or

$$V_y = (l \cos\theta)\omega. \tag{3.14}$$

The equilibrium relation between τ and F is

$$(l \cos\theta)F = \tau. \tag{3.15}$$

The occurrence of $l \cos\theta$ in both Eqs. (3.14) and (3.15) might appear to be coincidental until we remember that the device must conserve power. Equations (3.14) and (3.15) are embodied in the bond graph of Figure 3.10*b*. The *MTF* is called a *displacement*-modulated transformer because the modulus m is a function of the displacement variable, θ. Such transformers allow one to create a bond graph for the extremely complex dynamic systems associated with three-dimensional rigid-body motion.

Although some useful 2-port components have been defined, still only very simple bond graphs can be assembled from 2-ports and 1-ports. Only chains of 2-ports with 1-ports at the ends can be made. To create the complex models necessary in engineering system dynamics, it turns out that 3-ports are required. But surprisingly, only two basic 3-ports are required to model a very rich variety of systems.

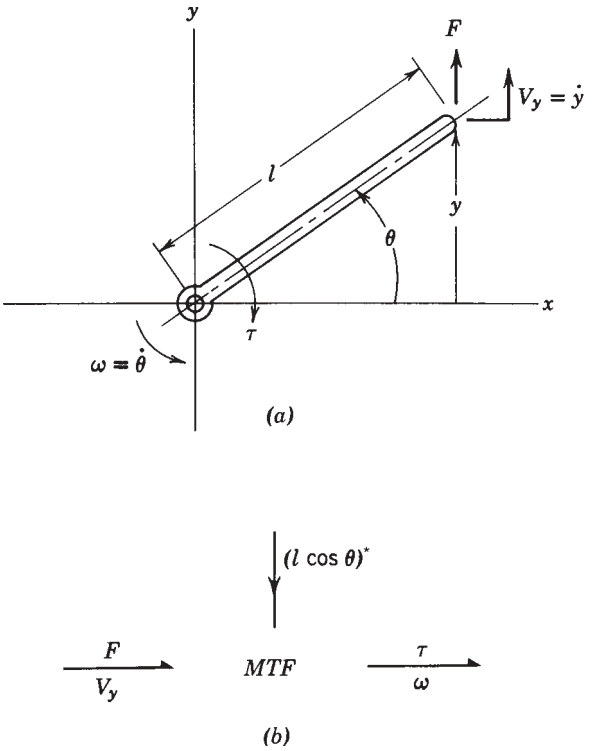


FIGURE 3.10. A displacement-modulated transformer: (*a*) sketch of rigid, massless rotating arm; (*b*) bond graph.*

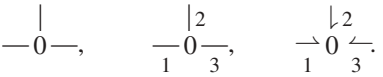
3.3 THE 3-PORT JUNCTION ELEMENTS

Imagine that you are surrounded with electrical components such as resistors, capacitors, inductors, motors, and so forth, mechanical components such as springs, shock absorbers, flywheels, and the like, and hydraulic components such as pipes, tees, accumulators, and so on. Further imagine that you are to connect these components any way you desire. You might imagine that there are hundreds of ways these diverse components could be connected. In fact, any possible connection of components can be realized using only two different connection patterns, and this brings us to the 3-port junction elements that allow elements in all energy domains to be assembled into overall system models.

*If one wishes to follow the convention of Figure 3.8 and Eq. (3.9), one might wish to indicate the modulus as the inverse of $l \cos \theta$. Equations (3.14) and (3.15) show unambiguously how the modulus $l \cos \theta$ is to be used.

We now introduce these two connection patterns as 3-port components themselves, which, like the 2-ports of the previous section, are power conserving. These 3-ports are called *junctions*, since they serve to interconnect other multiports into subsystem or system models. These 3-ports represent one of the most fundamental ideas behind the bond graph formalism. The idea is to represent in multiport form the two types of connections, which, in electrical terms, are called the *series* and *parallel* connections. As we shall see, such connections really occur in all types of systems, even though traditional treatments may not recognize the existence of the junctions as multiports.

First, consider the *flow junction*, *0-junction*, or *common effort junction*. The symbol for this junction is a zero with three bonds emanating from it (as will become evident, it is easy to extend the definition to a 4-, 5-, or more-port version of this 3-port element):



This element is ideal in that power is neither dissipated nor stored. Using the *inward power sign convention shown* in the last version of the junction, this implies

$$e_1 f_1 + e_2 f_2 + e_3 f_3 = 0. \tag{3.16}$$

The 0-junction is defined such that all efforts are the same, thus,

$$e_1(t) = e_2(t) = e_3(t). \tag{3.17}$$

Combining (3.16) and (3.17) yields

$$f_1(t) + f_2(t) + f_3(t) = 0. \tag{3.18}$$

In words, the efforts on all bonds of a 0-junction are always identical, and the algebraic sum of the flows always vanishes. In other words, if power is flowing in on two ports of the three, then it must be flowing out of the third port.

The use of the 0-junction is suggested by Figure 3.11*a*. The most obvious examples of 0-junctions are the electrical conductors connected as shown to provide three terminal pairs and the pipe tee junction, which is an idealized version of the hardware store variety. The mechanical example may seem obscure, and it is contrived. Mechanical 0-junctions are just as necessary as electrical or hydraulic ones, but they do not appear so readily in gadget form. The two carts in the mechanical example of Figure 3.11*a* are supposed to be rigid and massless. Note that $V_3 = -V_1 - V_2$, which conforms with Eq. (3.18). If F is the force across the gap, X_3, then F is the port effort for V_1, V_2, and V_3, in accordance with Eq. (3.17). Such a force would, in fact, exist if F were due to a massless spring or damper connected between the two carts.

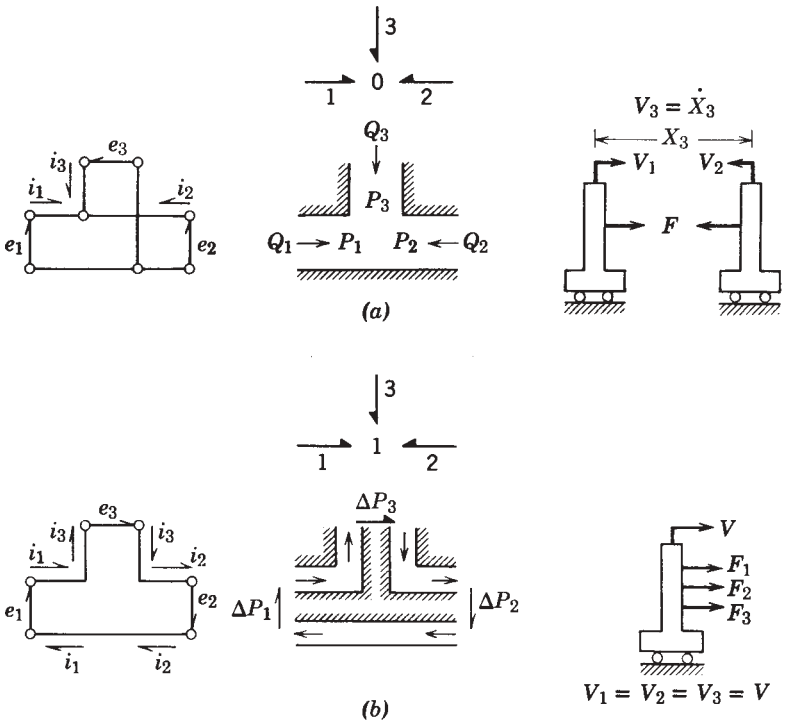


FIGURE 3.11. Basic 3-ports in various physical domains: (*a*) 0-junction; (*b*) 1-junction.

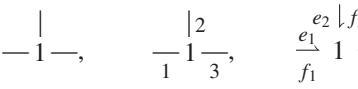
It is important to note that power conservation Eq. (3.16) and flow variable summation Eq. (3.18) change if not all the power half-arrows point inward as in the example above. Suppose that bond 3 had a half-arrow pointing *outward* instead of *inward* as is the case for bonds 1 and 2. Then a minus sign would appear in the power conservation equation and the flow summation, thus:

$$e_1 f_1 + e_2 f_2 - e_3 f_3 = 0. \tag{3.16a}$$

$$f_1(t) + f_2(t) - f_3(t) = 0. \tag{3.18a}$$

This is why we stress that for a 0-junction, the flows add *algebraically* to zero. Depending on the orientation of the sign-convention half-arrows on the bonds, minus signs or plus signs will appear in the equations of the junction. One simple way to decide on the proper signs in equations such as Eq. (3.18) or Eq. (3.18a) is to use a *plus sign* for all bonds with *inward*-pointing half-arrows and a *minus sign* for bonds with *outward*-pointing half-arrows. Note however, that the half-arrow directions have no effect on the strict equality of the efforts on all bonds of a 0-junction.

Before considering more examples in which the 0-junction is used, consider the *dual* of the 0-junction, that is, a multiport element in which the roles of effort and flow are interchanged. Such an element is an *effort junction*, a 1-*junction*, or a *common flow junction*. The symbol for this multiport is a 1 with three bonds:



This element is again power conserving according to Eq. (3.16) *if all half-arrows point inward*; however, this time the element is defined such that every bond has the identical flow, thus,

$$f_1(t) = f_2(t) = f_3(t), \tag{3.19}$$

which, when combined with the power conserving idealization, requires

$$e_1(t) + e_2(t) + e_3(t) = 0. \tag{3.20}$$

As with the 0-junction, the constitutive equations for the 1-junction, Eqs.(3.19 and 3.20), combine to ensure power conservation in the form of Eq. (3.16) as long as all bonds have an inward sign convention.

If a 1-junction should have an *inward* sign convention on bonds 2 and 3 but an *outward* sign convention on bond 3, the power conservation would be expressed as in Eq. (3.16a) and the effort summation of Eq. (3.20) would change to

$$e_1(t) + e_2(t) - e_3(t) = 0. \tag{3.20a}$$

The equality of the 1-junction flows, Eq. (3.19), is not affected by the power sign conventions on the bonds.

The 1-junction has a single flow, and the *algebraic* sum of the effort variables on the bonds vanishes. Figure 3.11*b* shows some instances in which a 1-junction can be used to model physical situations. Both the electrical conductors and the hydraulic passages are arranged so that if 1-port components were attached to the ports, one could describe the resulting connection as a series connection. A single current or volume flow would circulate, and the voltages and pressures at the ports would sum algebraically to zero. In the mechanical example, the three forces are all associated with a common velocity, and the forces must sum to zero, since the cart is assumed to be massless.

An understanding of the meaning of the 0- and 1-junctions is important for anyone learning bond graph techniques, and it may be helpful to give some physical interpretation for these multiports in several physical domains.

Electrical circuits:		represents Kirchhoff's current law for a node where three conductors join
		represents Kirchhoff's voltage law written along a loop in which a current flows and experiences three voltage drops
Mechanical systems:		represents geometric compatibility for a situation involving a single force and three velocities that algebraically sum to zero
		represents dynamic equilibrium of forces associated with a single velocity—when an inertia element is involved, the junction enforces Newton's law for the mass element
Hydraulic systems:		represents the conservation of volume flow rate at a point where three pipes join
		represents the requirement that the sum of pressure drops around a circuit involving a single flow must sum algebraically to zero

As might be expected, the existence of 0- and 1-junctions within complex systems is not always obvious, but in succeeding chapters formal techniques for modeling systems using these basic elements are presented.

To make clear the utility of the junctions, four elementary example systems are displayed in Figure 3.12. Note that only two bond graphs are involved. The series and parallel aspects of the junctions are more obvious in the electrical than in the mechanical cases. The reader should study these examples to understand how the sign conventions are transferred from the physical sketches to the bond graph.

Note that in Figure 3.12 the 1-ports have the sign-convention half-arrows as they were presented in the tables at the beginning of this chapter, but the junction signs are not all inward pointing. As we have noted, when the sign-convention arrows are changed from the inward-pointing convention used to introduce the 3-ports, Eqs. (3.18) and (3.20) must be modified with a minus sign for each port with an outward-pointing sign. Equations (3.17) and (3.19) remain invariant, however, under changes in a sign convention. Remember, *a 0-junction has only a single effort and a 1-junction has only a single flow, independent of the sign convention*. As an example, consider

$$\xrightarrow[f_1]{e_1}\ e_2|f_2\ 1\ \xrightarrow[f_3]{e_3},$$

the equations of which are

$$f_1 = f_2 = f_3, \quad e_1 - e_2 - e_3 = 0. \tag{3.21}$$

The reader should verify that the systems and bond graphs are consistent by writing equations such as (3.21).

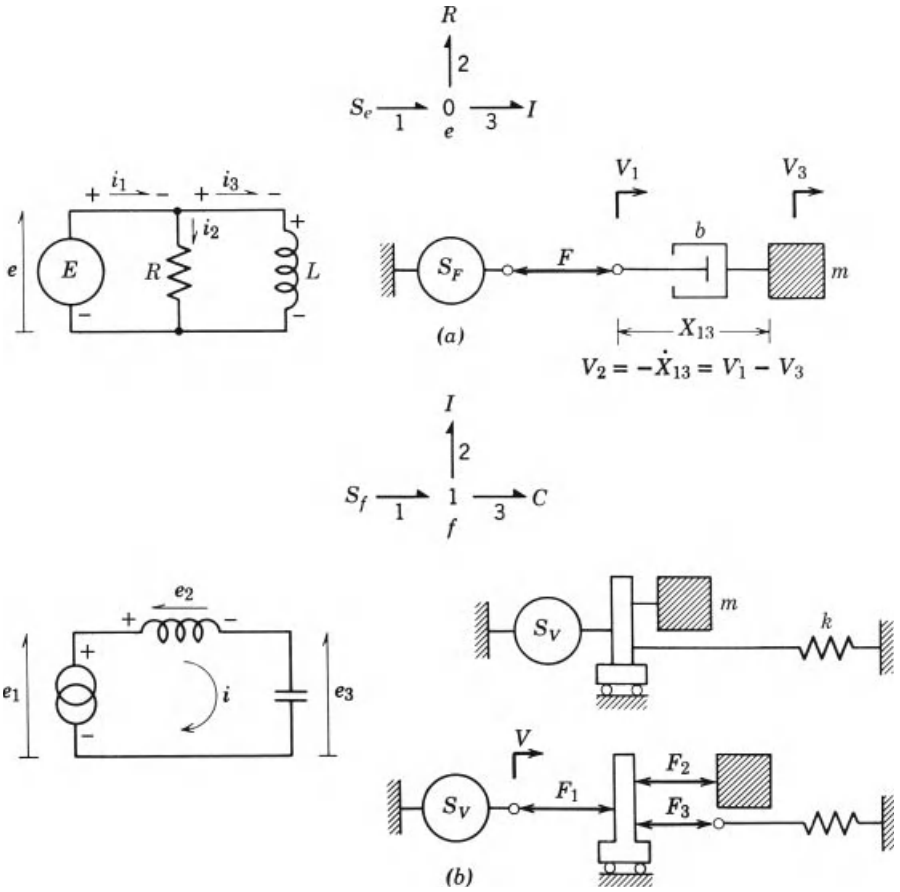
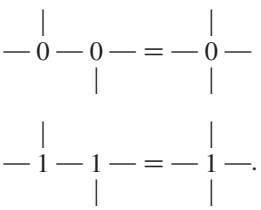


FIGURE 3.12. Example systems involving basic 3-ports. (*a*) Systems using 0-junctions; (*b*) systems using 1-junctions.

The slight generalization from 3-port junctions to 4- or n-port junctions is worth emphasizing. In bond graph symbolism, two similar 3-ports may be combined into a 4-port thus:



An n-port 0- or 1-junction has a common effort or flow on all bonds, and the algebraic sum of the complementary power variables on the bonds vanishes.

Occasionally, 2-port junctions arise, and, in some cases, these are precisely equivalent to a single bond. The following bond graph identities are always valid:

$$\rightharpoonup 0 \rightharpoonup = \rightharpoonup, \qquad \rightharpoonup 1 \rightharpoonup = \rightharpoonup .$$

However, with some sign patterns, the 2-port 0- and 1-junctions serve to reverse the sign definition of an effort or flow. For example,

$$\underset{f_1}{\overset{e_1}{\rightharpoonup}} 0 \underset{f_2}{\overset{e_2}{\leftharpoondown}} \quad \text{implies} \quad e_1 = e_2,\ f_1 = -f_2,$$

and

$$\underset{f_1}{\overset{e_1}{\rightharpoonup}} 1 \underset{f_2}{\overset{e_2}{\leftharpoondown}} \quad \text{implies} \quad f_1 = f_2,\ e_1 = -e_2.$$

Such 2-ports are sometimes necessary when two multiports are to be joined by a bond, but the two multiports have been defined with signs that are not compatible with a single bond. For example, in connecting a spring, —C, to a mass, —I, one could define a common velocity but use a 2-port 1-junction to express the fact that the spring force is the negative of the force on the mass. The resulting bond graph would then be $C \leftharpoondown 1 \rightharpoonup I$, in which the two passive 1-ports have the appropriate inward sign convention. In this case, both bonds have the same flow but one effort is the negative of the other.

The constitutive relations for 0- and 1-junctions are summarized in Table 3.5.

3.4 CAUSALITY CONSIDERATIONS FOR THE BASIC ELEMENTS

The concept of causality was discussed in general terms in Section 2.4, and we may now consider some more specific uses of the idea with respect to the basic multiports. Some of the causal properties developed here will be applied in later chapters. For now, we simply note that some of the basic multiports are heavily constrained with respect to possible causalities, some are relatively indifferent to causality, and some exhibit their constitutive laws in quite different forms for different causalities.

TABLE 3.5. Summary of Basic 3-Ports

Flow junction, or 0-junction	$\underset{f_1}{\overset{e_1}{\rightharpoonup}} 0 \underset{f_2}{\overset{e_2}{\leftharpoondown}} \underset{f_3}{\overset{e_3}{\leftharpoondown}}$	$e_1 = e_2 = e_3,$ $f_1 + f_2 + f_3 = 0$
Effort junction, or 1-junction	$\underset{f_1}{\overset{e_1}{\rightharpoonup}} 1 \underset{f_2}{\overset{e_2}{\leftharpoondown}} \underset{f_3}{\overset{e_3}{\leftharpoondown}}$	$f_1 = f_2 = f_3,$ $e_1 + e_2 + e_3 = 0$

3.4.1 Causality for Basic 1-Ports

The effort and flow sources are the most easily discussed from a causal point of view, since, by definition, a source impresses either an effort or flow time history upon whatever system is connected to it. Thus, if we use the symbols S_e— and S_f— for the abstract effort and flow sources, the only permissible causalities for these elements are

$$S_e \dashv \quad \text{and} \quad S_f \vdash.$$

in which the causal stroke indicates the direction that the effort signal is oriented. (Remember that a sign-convention half-arrow could be placed on either end of the bond without affecting the causality.) The causal forms for effort and flow sources are summarized in the first two rows of Table 3.6.

In contrast to the sources, the 1-port resistor is normally indifferent to the causality imposed upon it. The two possibilities may be represented in equation form as follows:

$$e = \Phi_R(f), \quad f = \Phi_R^{-1}(e),$$

where we use the convention that the variable on the left of the equality sign represents the output of the resistor (the dependent variable), and that appearing in the function of the right side is the input (independent) variable for the element. This convention is used commonly, but not universally, in writing equations and corresponds to the notation used in computer programming.

The correspondences between the causally interpreted equations and the causal strokes on the bond of the R— element are shown in the third row of Table 3.6. As long as both the functions Φ_R and Φ_R^{-1} exist and are known, either causal version of them could be used in a system model. It is possible, however, that the static relation between e and f shown in Figure 3.1 is multiple valued in one direction or the other; that is, either Φ_R or Φ_R^{-1} might be multiple valued. In

TABLE 3.6. Causal Forms for Basic 1-Ports

Element	Acausal Form	Causal Form	Causal Relation
Effort source	$S_e \rightharpoonup$	$S_e \dashv$	$e(t) = E(t)$
Flow source	$S_f \rightharpoonup$	$S_f \vdash$	$f(t) = F(t)$
Resistor	$R \leftharpoonup$	$R \vdash$	$e = \Phi_R(f)$
		$R \vdash$	$f = \Phi_R^{-1}(e)$
Capacitor	$C \leftharpoonup$	$C \dashv$	$e = \Phi_C^{-1}\left(\int^t f\, dt\right)$
		$C \vdash$	$f = \dfrac{d}{dt}\Phi_C(e)$
Inertia	$I \leftharpoonup$	$I \vdash$	$f = \Phi_I^{-1}\left(\int^t e\, dt\right)$
		$I \dashv$	$e = \dfrac{d}{dt}\Phi_I(f)$

such a case, the single-valued causality would be clearly preferable. In the linear case, with a finite slope of the e–f characteristic, the 1-port resistor is indifferent to the causality imposed upon it, although the resistance law would be written in two forms:

$$e = Rf \quad \text{or} \quad f = (1/R)e = Ge.$$

The constitutive laws of the C— and I— elements are expressed as static relations between e and $q = \int^t f dt$ and f and $p = \int^t e dt$, respectively. In expressing causal relations between e's and f's, we will find that the choice of causality has an important effect. Taking the capacitor, we may rewrite the relations from Table 3.2 as follows:

$$e = \Phi_C^{-1}\left(\int^t f\,dt\right), f = \frac{d}{dt}\Phi_C(e), \tag{3.22}$$

in which causality is implied by the form of the equation. Note that when f is the input to the C—, e is given by a static function of the time integral of f, but when e is the input, f is the time derivative of a static function of e. The correspondences between these causal equations and the causal stroke notation for the capacitance are shown in the fourth row of Table 3.6. The implications of the two types of causality, which are called *integral causality* and *derivative causality*, respectively, will be discussed in some detail in later chapters.

Since inertia is the dual* of the capacitor, similar effects occur with the two choices of causality. Rewriting the inertia element relations from Table 3.3, we have

$$f = \Phi_I^{-1}\left(\int^t e\,dt\right), e = \frac{d}{dt}\Phi_I(f). \tag{3.23}$$

In this case, integral causality exists when e is the input to the inertia, and derivative causality exists when f is the input. These observations are summarized in the fifth row of Table 3.6. Equations (3.22) and (3.23) are written in a form suitable for nonlinear C— and I— elements, but the distinction between integral and derivative causality remains for the special case of linear elements with their compliance and inertance parameters.

3.4.2 Causality for Basic 2-Ports

Proceeding now to the basic 2-ports, one might think initially that there would be a total of four possibilities for the assignment of causality of a transformer, namely, any combination of the two possible causalities for each of the two ports. However, there are only two possible causality assignments, as the defining

*Dual elements have identical types of constitutive laws except that the roles of effort and flow are interchanged.

relations (3.9) and (3.11) show. As soon as one of the e's or f's has been assigned as an input to the —TF—, the other e or f is constrained to be an output by Eq. (3.9). Thus, in fact, the only two possible choices for causality for the transformer are $\vdash TF\vdash$ and $\dashv TF\dashv$. The possible causalities are tabulated in the first row of Table 3.7. Again, causal equation equivalents to the causal stroke notation are given for all elements in Table 3.7.

For the gyrator, Eqs. (3.11) show that as soon as the causality for one bond has been determined, the causality for the other is also. Thus, the only permissible causal choices for the —GY— are $\dashv GY\vdash$ and $\vdash GY\dashv$. The choices for the causality for the gyrator are summarized in the second row of Table 3.7.

3.4.3 Causality for Basic 3-Ports

The causal properties of 3-port 0- and 1-junctions are somewhat similar to those of the basic 2-ports. Although each bond of the 3-ports, considered alone, could have either of the two possible causalities assigned, not all combinations of bond causalities are permitted by the constitutive relations of the element. For example, the constitutive relations for the 0-junction given in Table 3.5 indicate that all efforts on all the bonds are equal and the flows must sum to zero. Thus, if on any bond the effort is an input to a 0-junction, then all other efforts are determined, and on all other bonds they must be outputs of the 0-junction. Conversely, if all the flows on all bonds except one are inputs to the 0-junction, the flow on the remaining bond is determined and must be an output of the junction. A typical

TABLE 3.7. Causal Forms for Basic 2-Ports and 3-Ports

Element	Acausal Graph	Causal Graph	Causal Relations
Transformer	$\overset{1}{\rightharpoonup} TF \overset{2}{\rightharpoonup}$	$\overset{1}{\vdash\!\rightharpoonup} TF \overset{2}{\rightharpoonup\!\vdash}$	$e_1 = me_2$ $f_2 = mf_1$
		$\overset{1}{\dashv\!\rightharpoonup} TF \overset{2}{\rightharpoonup\!\dashv}$	$f_1 = f_2/m$ $e_2 = e_1/m$
Gyrator	$\overset{1}{\rightharpoonup} GY \overset{2}{\rightharpoonup}$	$\overset{1}{\vdash\!\rightharpoonup} GY \overset{2}{\rightharpoonup\!\dashv}$	$e_1 = rf_2$ $e_2 = rf_1$
		$\overset{1}{\dashv\!\rightharpoonup} GY \overset{2}{\rightharpoonup\!\vdash}$	$f_1 = e_2/r$ $f_2 = e_1/r$
0-Junction	$\overset{1}{\rightharpoonup} 0 \overset{2}{\leftharpoondown}$ $3\uparrow$	$\overset{1}{\dashv\!\rightharpoonup} 0 \overset{2}{\leftharpoondown\!\dashv}$ $3\perp$	$e_2 = e_1$ $e_3 = e_1$ $f_1 = -(f_2 + f_3)$
1-Junction	$\overset{1}{\rightharpoonup} 1 \overset{2}{\leftharpoondown}$ $3\uparrow$	$\overset{1}{\vdash\!\rightharpoonup} 1 \overset{2}{\leftharpoondown\!\vdash}$ $3\top$	$f_2 = f_1$ $f_3 = f_1$ $e_1 = -(e_2 + e_3)$

permissible causality for a 0-junction is shown in the third row of Table 3.7. Here the causal stroke on the end of bond 1 nearest the 0 indicates that e_1 is an input to the junction and that all other bonds must have causal strokes at the end away from the 0. To interpret the diagram another way, the flows on bonds 2 and 3 are inputs to the 0-junction. These considerations are also expressed by the causal equations shown in Table 3.7. For a 3-port 0-junction, then, there are only three different permissible causalities in which each of the three bonds in succession plays the role assigned to bond 1 in the example shown in the table. For an n-port 0-junction this description of the constraints on causality is still valid, and there are exactly n different permissible causal assignments.

For a 1-junction the same considerations apply as for a 0-junction except that the roles of the efforts and flows are interchanged. Table 3.5 indicates that flows on all the bonds are equal and the efforts sum to zero. Thus, if the flow on any single bond is an input to the 1-junction, the flows on all other bonds are determined and must be considered outputs of the junction. Alternatively, when the efforts on all bonds except one are inputs to the 1-junction, the effort on the remaining bond is determined and must be an output of the junction. A typical permissible causality is shown in the fourth row of Table 3.7. In this example, bond 1 plays the special role of determining the common flow at the junction, and the remaining bonds supply effort inputs that suffice to determine the effort on bond 1. Clearly, there are three permissible causalities for a 3-port 1-junction, and there are n permissible different causal assignments for an n-port 1-junction.

Although the causal considerations have been stated for all the basic multiports defined so far (summarized in Tables 3.6 and 3.7), it can hardly be clear as to what all the implications of causality are. The study of causality is very important, and bond graphs are uniquely suited to this study. However, only when some real system models have been assembled is it clear why causal information is so important. In the next chapter, system models are built up using the basic multiports just discussed. Using the rules of causality, it is then possible to predict many important features of these systems even before the constitutive laws for all the elements have been decided upon. For instance, it will be possible to predict the mathematical order of the system model before any equations are written and before a decision has been made about whether the model should be linear or nonlinear. In addition, causal considerations will prove invaluable in writing a set of state equations or in setting up computational program for system simulation.

3.5 CAUSALITY AND BLOCK DIAGRAMS

Block diagrams indicate how input and output signals flow in a system. The blocks show how the signals are transformed, for example, by integration, differentiation, multiplication by a constant or through some other functional relationship. Block diagrams are inherently causal since they show how an input signal is transformed into an output signal.

When causal strokes are added to a bond graph, one may represent the information in the bond graph by a block diagram. For example, the block-diagram versions of the causal forms for the R, C, and I 1-ports shown in Table 3.6 are given in Figure 3.13. Similarly, block diagrams for 2-ports and 3-ports corresponding to entries in Tables 3.7 are shown in Figures 3.14 and 3.15. It should be possible for you to correlate the signal flow paths in the block diagrams with the equations in the tables and with the bond graph representation. Note that when one rigorously maintains the spatial arrangements with efforts above and to the left of bonds and flows below and to the right, the block diagrams have fixed patterns.

It may also be seen that block diagrams are more complex graphically than bond graphs because a single bond implies two signal flows on a block diagram. A single causal mark indicates the signal flow directions on two (effort and flow) signals. Also, a single sign-convention half-arrow on a bond is replaced on the equivalent block diagram by a plus or minus sign at summation points.

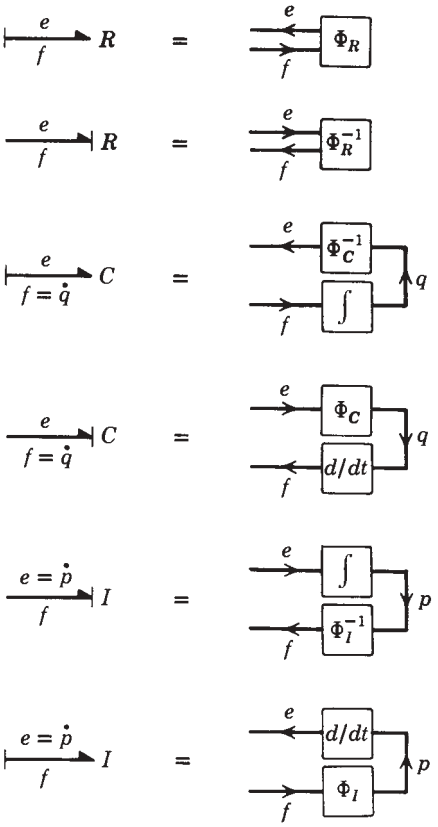


FIGURE 3.13. Block diagrams for 1-ports.

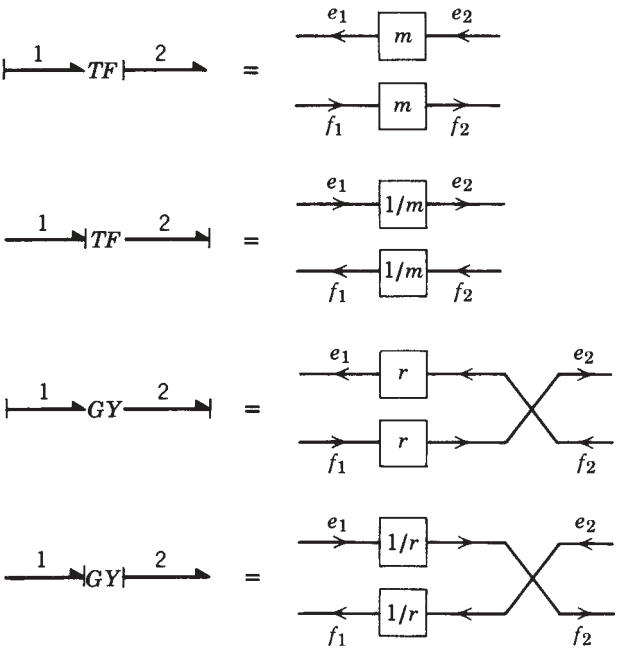


FIGURE 3.14. Block diagrams for 2-ports.

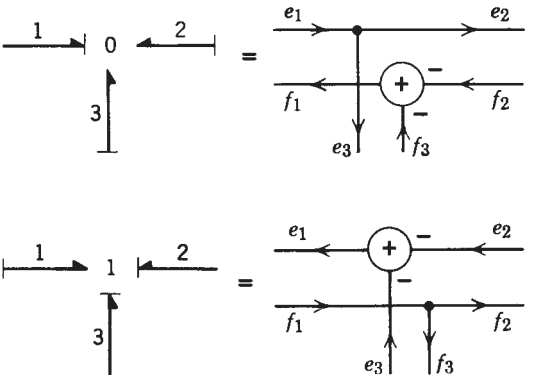


FIGURE 3.15. Block diagrams for 3-ports.

Initially, block diagrams may be easier to understand than bond graphs because they contain redundant information. For systems with some complexity, however, block diagrams rapidly become so complicated that the conciseness of equivalent bond graphs is an advantage. For example, Figure 3.16 shows a block diagram equivalent to a bond graph model of the automotive drive train system of Figure 2.7. Note that the sign-convention half-arrows have yet to be put on

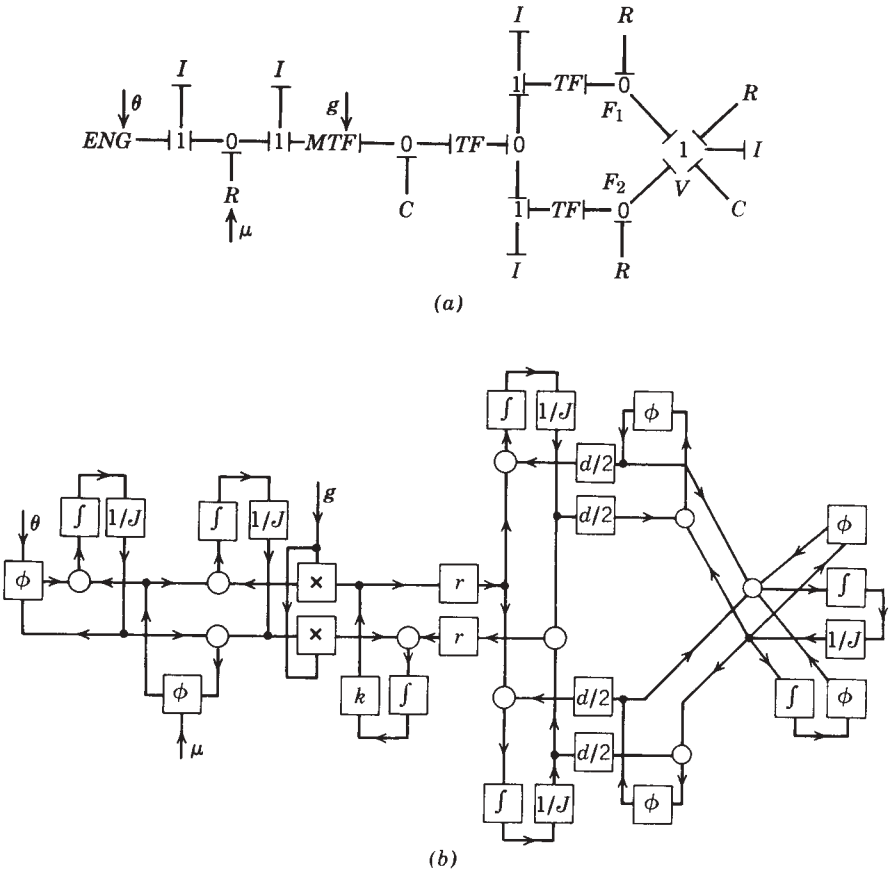


FIGURE 3.16. Interconnected drive train model. (*a*) Bond graph; (*b*) block diagram.

the bond graph, and the corresponding + and − signs do not appear in the block diagram near the circles representing signal summation.

After procedures for constructing bond graph models and adding causal strokes to them have been discussed in the following chapters, one option is to construct a block diagram from the bond graph. There are computer programs that accept block diagram models for the analysis and simulation of systems so there are some reasons for converting a bond graph model of a system to a block diagram. In the following chapters we will focus more on deriving state equations for bond graph models but the ability to show a block diagram for a bond graph is often useful in explaining the assumptions behind a model to those unfamiliar with bond graph techniques. Also, particularly in the case of linearized system models, block diagrams are commonly used in the design of control systems. It is also true that a detailed block diagram for a complex system model may

become so complex as to be virtually useless in understanding the assumptions that have been made.

REFERENCE

[1] S. H. Crandall, D. C. Karnopp, E. F. Kurtz, and D. C. Pridmore-Brown, *Dynamics of Mechanical and Electromechanical Systems*, New York: McGraw-Hill, 1968.

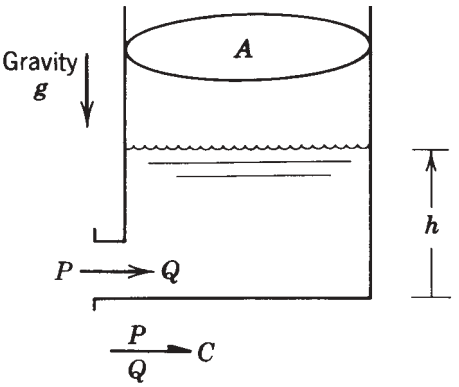
PROBLEMS

3-1. A nonlinear dashpot has as its constitutive relation the "absquare law,"

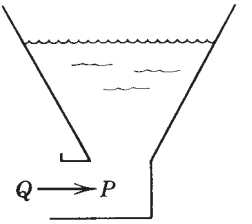
$$F = AV\,|V|\,,$$

where F and V are the force and velocity across the dashpot and A is a constant. Plot this relation in a sketch, and indicate the bond graph sign convention implied if $A > 0$ and which causality the equation implies in the form given. Try to invert the constitutive law to yield the velocity as a function of the force.

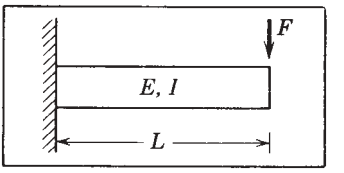
3-2. A fluid of mass density ρ is pumped into an open-topped tank of area A. If P is the pressure at the tank bottom and Q the volume flow rate, the tank is approximately a $-\!\!C$ for slow changes in the volume of fluid stored. Is the $-\!\!C$ a linear element in this case, and, if so, what is the capacitance? *Hint*: It is useful to compute the height of fluid, h, as a function of the total volume of fluid as an intermediate step.



3-3. Reconsider Problem 3-2, but let the tank have sloping walls as shown. What does this do to the constitutive law for the device?



3-4. Consider a uniform cantilever beam of length L, elastic modulus E, and area moment of inertia I. If a force F is applied at the tip of the beam, it will deflect. If the beam is supposed to be massless, decide what type of 1-port it is and compute its constitutive law.

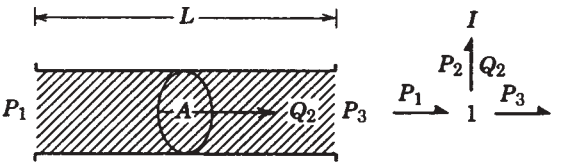


3-5. Linear electrical inductors can be characterized by a law relating the voltage e to the rate of change of current:

$$L\frac{di}{dt} = e.$$

Convert this to a law relating the flow i to the momentum λ, which in this case is the time integral of e and is called the *flux linkage*. Plot a linear flow–momentum constitutive law, and show on the plot where the inductance L appears. Now sketch a nonlinear flow–momentum law. Convert the nonlinear law back to a relation between e and di/dt if this is possible.

3-6. A rigid pipe filled with incompressible fluid of mass density ρ has length L and cross-sectional area A.

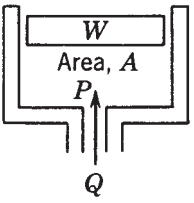


If P_1 and P_2 are the pressures at the ends of the pipe and the volume flow rate is Q_2, convince yourself that the bond graph shown correctly represents the pipe in the absence of friction. Show that the correct constitutive law relating pressure momentum and volume flow is

$$p_{P_2} = \int^t P_2 dt = \left(\rho \frac{L}{A}\right) Q_2$$

by writing Newton's law for the slug of fluid in the pipe. (It should come as a surprise that small-area tubes have a lot of inertia when P, Q variables are used.)

3-7. An accumulator consists of a heavy piston in a cylinder. If the pressure is determined primarily by the weight of the piston, sketch the constitutive law for this 1-port device.



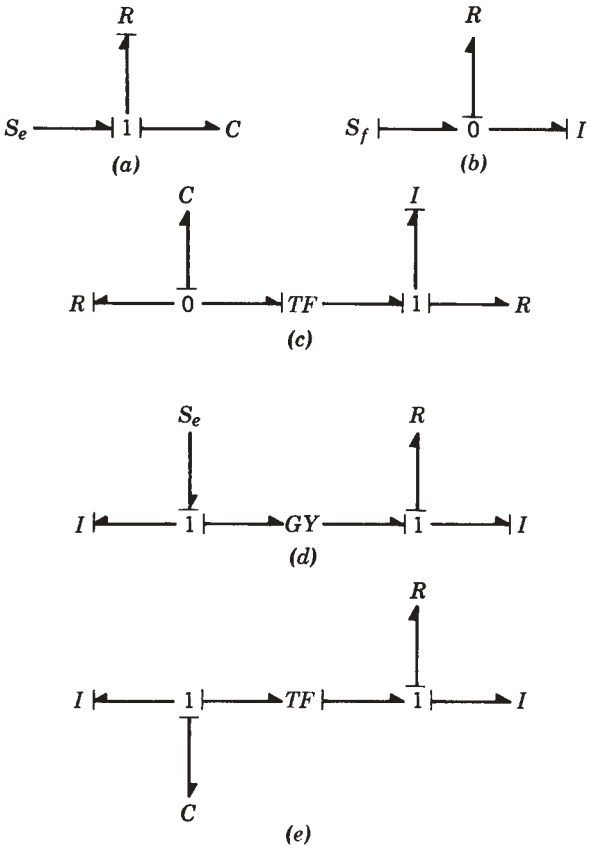
3-8. A flywheel is a uniform disk of radius R and thickness t and is made of a material of mass density ρ. Write the constitutive law for the 1-port representing the flywheel in its flow–momentum form. Evaluate the inertance parameter for a steel disk 1 in. in thickness and 10 in. in diameter.

3-9. Assume any needed dimensions for the hydraulic ram of Figure 3.8*e*, and write the constitutive law for this 2-port.

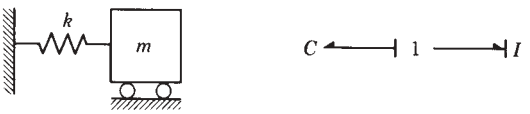
3-10. Repeat Problem 3-9 for the devices of Figure 3.8*b* and *c*.

3-11. In Figure 3.9*c* assume that the rotor has moment of inertia J and spins at a high angular rate Ω. If the rotor is centered on an axle of length L, relate F_1, V_1, F_2, V_2, and thus demonstrate that the device is indeed approximately a —*GY*—.

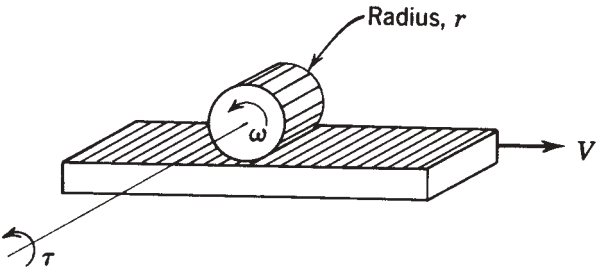
3-12. Draw block diagrams for the following bond graphs, assuming all 1-ports are linear:



3-13. Draw a block diagram for the oscillator using the bond graph shown.



3-14. Consider an ideal rack and pinion with no friction losses:



If the pinion has radius r and torque and speed τ and ω, and if the rack has velocity V and force F, what type of element would represent the device? Write the appropriate constitutive laws.

3-15. An electrodynamic loudspeaker is driven by a voice coil transducer described by Eq. (3.12). Show that if only the mass of the speaker cone is considered, then at the electrical terminals the device will act like a capacitor. The bond graph identity

$$\underset{i}{\overset{e}{\rightharpoonup}} \dot{G}\dot{Y} \overset{T}{\underset{V}{\overset{F}{\rightharpoonup}}} I = \underset{i}{\overset{e}{\rightharpoonup}} C$$

is to be verified using the element constitutive equations directly.

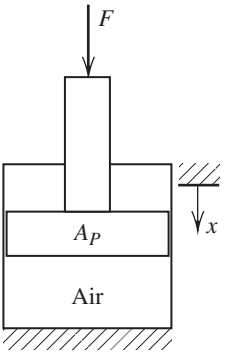
3-16. An air spring is idealized as a piston in a cylinder with no leakage and no heat transfer through the cylinder walls. The process the air undergoes is assumed to be isentropic, such that

$$PV^{\gamma} = P_0 V_0^{\gamma},$$

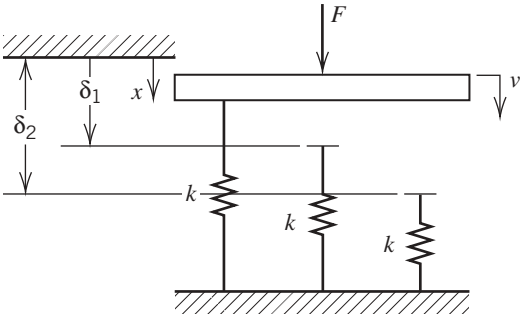
where V is the instantaneous volume,

$$V = V_0 - A_p x,$$

V_0 is the volume when $x = 0$, P_0 is atmospheric pressure, γ is the ratio of specific heats ($\approx$1.4 for air), and P is the absolute cylinder pressure. Derive the nonlinear constitutive relationship for $F = F(x)$.



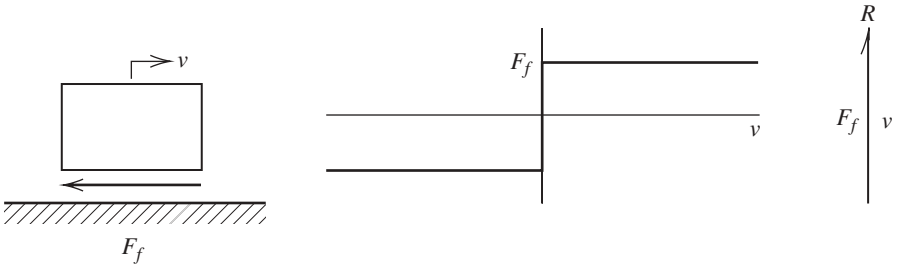
3-17. Three springs are used in a parallel configuration as shown. Sketch the constitutive behavior as seen from the $F-v$ port. Show both compression and tension.



3-18. Friction is always dissipative and is therefore a resistance in a bond graph model. Its constitutive behavior is sketched in an ideal sense.

(a) Discuss the only possible causality for this element.

(b) Discuss the problem of using this device near $v = 0$. Propose a change to the constitutive behavior that will retain its fundamental character but avoid the problem near $v = 0$.



4

SYSTEM MODELS

Using the bond graph arsenal of R, C, I, S_e, S_f, 0, 1, TF, and GY it is possible to create bond graph mathematical models for a surprisingly wide variety of physical systems. However, it is not true that every system you may encounter can be modeled as a simple bond graph with the elements defined so far. For electrical circuits, certain classes of mechanical systems, hydraulic circuits, and some transducer systems, the elements studied so far will be sufficient. Chapters 7 through 12 present advanced techniques for an even broader class of systems. From this chapter through Chapter 6, we purposely limit the systems to be considered in order to present the concepts of modeling, analysis, simulation, and control without introducing undue complexity.

In this chapter we first show how to represent any electrical circuit in bond graph form by using a direct, simple modeling procedure. Then we use a similar approach to problems involving mechanical translation. With slight extensions of the procedure, the next class of problems we treat contains fixed-axis rotation. We next couple translation and rotation and introduce plane motion dynamics. By generalizing our approach slightly, we are able to model hydraulic and acoustic circuits effectively, since they are similar in several respects to systems already treated.

All of the previously mentioned systems involve only a single type of power within a single system. For that reason they are said to be *single-energy-domain* systems. However, many devices involve two or more types of power. These must have *transducer* elements such as motors, pumps, sensors and actuators coupling subsystems in the different energy domains. Just imagine all the mechanical, hydraulic, electrical, and thermal subsystems present in a modern automobile—a device that illustrates the utility of the term *mechatronic* to indicate that a system uses electronic computer control of subsystems in several energy domains.

Copyright © 2012 John Wiley & Sons, Inc.

Some basic transducer system models are introduced as the next topic. The last section of this chapter is devoted to constructing bond graph models of multi-energy-domain systems. We will combine electrical, hydraulic, and mechanical devices to yield realistic overall physical systems that are ready for analysis and simulation.

In succeeding chapters, you will learn how to derive differential equations from a bond graph model, how to analyze certain types of equations, and how complex nonlinear systems can be handled in a computer. In this chapter, however, we want to emphasize that bond graphs are themselves precise mathematical models of systems that can be processed automatically by a computer program and that can, in many cases, yield time histories of interesting variables almost automatically, requiring only the specification of parameters, initial conditions, and forcing functions. A variety of computer programs are now available that will process bond graphs and put the implied equations into a form suitable for computational solution by a simulation program. The use of graphical displays of the results allows one to understand the dynamic responses of the system model and to optimize the system.

After studying the succeeding chapters, you will have an understanding of how automated simulation programs work, as well as the mathematical difficulties that may arise from some modeling decisions. For now, we wish only to emphasize that after a bond graph model has been devised, in most cases it is easy and convenient to perform computer experiments on the model and thus to size components or to change the system configuration if the predicted system behavior is not satisfactory. Beyond this, the dynamic bond graph model will prove useful in the design of control systems to improve the performance of the overall system.

4.1 ELECTRICAL SYSTEMS

We begin by observing that any electrical *circuit* can be modeled by a bond graph containing elements of the set $\{0, 1, R, C, I, S_e, S_f\}$. Notice that the elements TF and GY are not included. That is because these elements are properly used in representing electrical *networks*, a more general class than circuits. First we shall model circuits; then we shall extend the procedure to include network elements.

4.1.1 Electrical Circuits

In Chapter 3, the electrical resistance, inductance, and capacitance were shown to be bond graph —R, —I, and —C elements. It remains to determine how to use the junction elements (0- and 1-junctions) to construct an overall bond graph model of an electrical circuit. Sometimes, for simple circuits, it is easy to recognize that some elements have the same current (flow) and others have the same voltage (effort). For these circuits, bond graph construction can be accomplished by inspection. For example, Figure 4.1*a* shows a simple circuit with positive voltage drops and current directions defined and node voltages labeled.

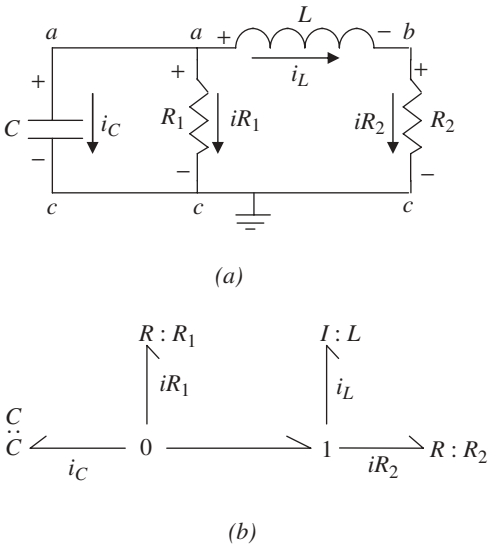


FIGURE 4.1. A simple electric circuit. Example 1.

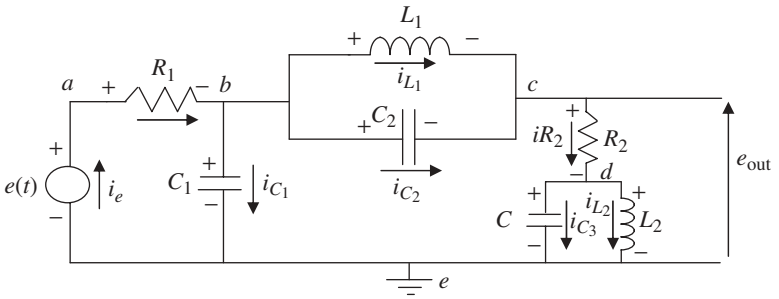
For convenience, some node voltages are repeated to emphasize their association with specific elements. The circuit is grounded at the bottom, meaning that the voltage labeled c is considered to be zero, $v_c = 0$.

To arrive at the bond graph in part b, we argue as follows: The elements C and R_1 have the same voltage ($v_a - v_c = v_a - 0 = v_a$) and thus are attached to the same 0-junction (common effort junction); the elements L and R_2 have the same current ($i_L = i_{R_2}$), so they are attached to the same 1-junction (common flow junction). The bond joining the 0- and 1-junctions enforces the fact that the current through the inductor is the sum of the currents through the capacitor and the resistor, R_1 (actually, with the sign convention shown in Figure 4.1a, $i_L = -i_C - i_{R_1}$). Notice that all the 1-port R, C, and I elements have the power half-arrows defined such that whenever the voltage drop across the element is in the direction defined as positive in Figure 4.1a, and the current is simultaneously in the defined positive direction, the power is flowing into the element. We always define positive power directions for R, C, and I such that this is true by pointing the half-arrows *toward* the elements.

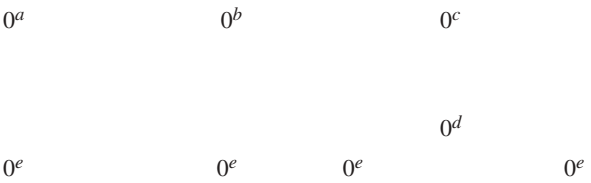
Most of the time, electric circuits are too complex to model by inspection. There may be some parts that are obviously in series or parallel (common current or common voltage), but constructing the overall bond graph model is much easier if a procedure can be followed that ensures success regardless of the complexity of the circuit. Here we present a foolproof circuit construction procedure and develop the procedure along with an example shown in Figure 4.2. This is a voltage-excited circuit, grounded at the bottom, and open circuited at the right side, exposing an output voltage, e_{out}.

Circuit Construction Procedure

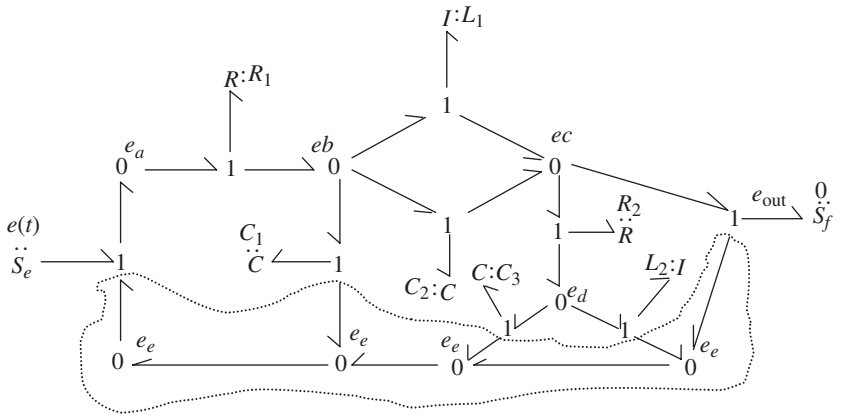
1. **Assign a power convention to the circuit schematic diagram.**
   This step must always be done regardless of the modeling procedure being used. On the circuit, this is done by showing the positive voltage drop and current directions. For the —I, —R, and —C elements, the positive voltage drop (+ to −) is shown in the same direction as the positive current.



(a)

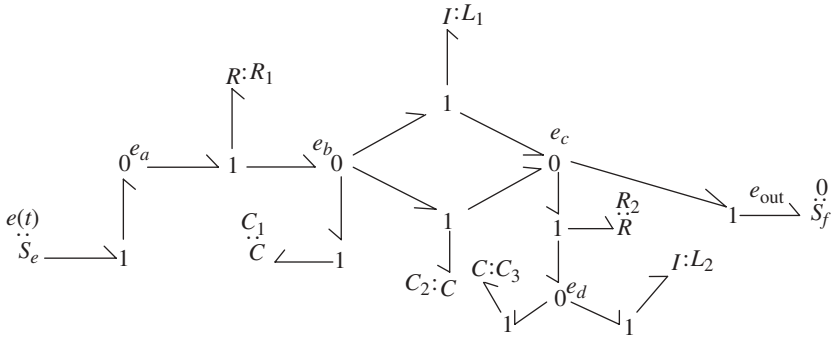


(b)

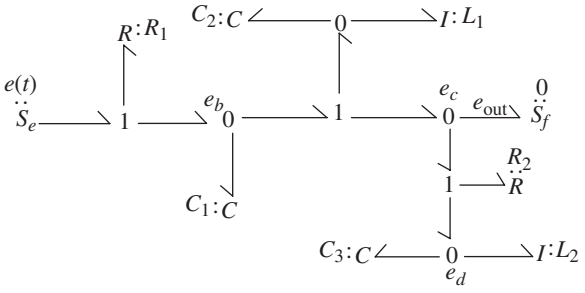


(c)

FIGURE 4.2. A complicated electric circuit. Example 2.



(d)



(e)

FIGURE 4.2. (*Continued*)

This ensures that when the power is positive, power is flowing *toward* the passive 1-ports.

For the source elements (S_f for current source and S_e for voltage source), it is not automatic which directions should be chosen for positive voltage drop and positive current (and eventually which end of the bond should have the half-arrow). If positive current is defined such that the current moves "uphill" against the positive voltage as is done in Figure 4.2*a*, then if the current is positive, power will come from the source into the rest of the circuit. If either the positive voltage direction or current direction is chosen in the opposite direction, then positive power will flow toward the source and negative power will flow away from the source. There is absolutely nothing wrong with this, and, in fact, real time-varying sources may sometimes absorb power from the attached system and sometimes supply power to the attached system.

2. **Label each node voltage on the circuit schematic diagram and use a 0-junction to represent each node voltage as shown in Figure 4.2*b*.**

   Node voltages are all the independent voltages that exist at points in the circuit, including the ground node voltage that later will be set to zero. The

node voltages in Figure 4.2*b* appear above and below or to the left and right of each circuit element. In the figure, the node voltages are labeled using letters. For convenience, the ground voltage, e, is repeated several times. Remember, every bond that touches a particular 0-junction has the voltage associated with that junction.

3. **Establish the positive voltage drops across the circuit elements using 1-junctions.**

   Remember that 1-junctions add efforts (voltages) algebraically according to the power convention represented by the half-arrows. By properly directing the half-arrows on 1-junctions, the voltage drop can be established across each bond graph element. See the discussion about signs on 0- and 1-junctions in Section 3.3 and particularly the example leading to Eq. (3.21). (This step will require practice to accomplish rapidly with confidence but this is a worthwhile endeavor.)

   In every case the sign half-arrows flow *through* the 1-junctions, indicating that the circuit elements react to node voltage *differences*, never node voltage sums. Which way the arrows flow through the 1-junctions depends on the sign of the voltage difference established in step 1.

   Figure 4.2*c* shows this construction. For example, for the —R element representing the resistor, R_1, the effort on the bond attached to the R is $e_a - e_b$, which corresponds to the positive voltage drop defined in the circuit schematic.

   Notice that positive power flows out of the voltage source element and the voltage on the source bond is $e_a - e_e$, as was defined in the schematic. Also, the output voltage, e_{out}, is exposed using a flow source, S_f, of zero current. (The idea of an open circuit is that no current flows even if there is a voltage across the terminals.) The voltage drop across this flow source is $e_{\text{out}} = e_c - e_e$. The reader should check the other elements and ensure that all have their defined positive voltages.

4. **Chose a ground or reference node and remove all bonds that have zero power.**

   Before the bond graph can be used for equation derivation or simulation, the *reference or ground node* must be established. (The reference node is the point at which we would put the black probe of a voltmeter if we were measuring voltages in the circuit. In other words it is the voltage we assume to be zero when we measure other voltages.)

   The reference in Figure 4.2 has been chosen to be e_e, which we now set to zero. Since every bond that touches a 0-junction has the identical voltage, all the bonds inside the dotted line in Figure 4.2*c* have zero voltage and therefore each of those bonds carries no power. We can either append an effort source of zero voltage to one of the 0-junctions representing e_e, or we can simply erase all the bonds that carry no power, with the result shown in Figure 4.2*d*.

5. **Simplify the bond graph by using the bond graph identities defined in Chapter 3**.

   This is not an absolutely necessary step but it is generally useful. By removing any 2-port 0- and 1-junctions that happen to have a through power convention, a much neater picture emerges, as shown in Figure 4.2*e*. (These identities were discussed in Section 3.3.)

   Also, the loop structure from e_b to e_c has been reduced using another type of bond graph identity. Noticing that L_1 and C_2 actually have the same voltage drop, we establish the voltage drop $e_b - e_c$ once using a 1-junction, and then establish a 0-junction having this voltage difference. Finally, the —I and —C elements associated with L_1 and C_2 are attached to this 0-junction, yielding the bond graph of Figure 4.2*e*.

   For equation derivation or automated simulation as described in Chapter 5, it is not necessary to reduce the bond graph to its simplest form, but it does make the final bond graph easier to use and understand.

   Note that the flow source introduced to expose the output voltage, e_{out}, may also be erased since there is no power associated with the source bond. It is left in Figure 4.2*e* as a convenience to remind us that we are interested in that particular output voltage. Later we will see that any effort or flow on any bond can be simply established as an output, and we will not have to construct an artificial means to expose desired outputs.

The trick to creating arbitrary circuit connections exclusively out of series and parallel connections (or 0- and 1-junctions in bond graph terms) lies in first representing *every* node voltage, including what will be the ground or reference voltage, by establishing a set of 0-junctions. Only after representing voltage differences using 1-junctions for each circuit element in step 3, do we set the ground voltage to zero.

This means that different bond graphs will result for a given circuit if we chose different ground nodes. This corresponds to the fact that if different points in a real circuit are used as a reference point for voltage measurements, the measured voltages will have different values. The different bond graphs will correctly predict the voltages with respect to the different choices for the ground node.

As a final example of circuit modeling using the procedure just presented, consider the Wheatstone bridge shown in Figure 4.3*a*. This circuit is typically used with strain gages as the resistive elements R_1 through R_4, and the voltage across the load resistance, R_L, is the output that is indicative of any change in the bridge resistances. We are simply going to model this circuit as an exercise in using the bond graph circuit construction procedure. In Figure 4.3*a*, the positive voltage drop and current directions are shown along with labels for the node voltages. This is step 1 of the procedure.

Step 2 of the procedure is done in Figure 4.3*b*, where the node voltages are represented by 0-junctions. Step 3 is shown in Figure 4.3*c*, where 1-junctions are used to appropriately add voltages such that the proper positive voltage drops are across each element. Also shown in *c* is the dotted line surrounding all the

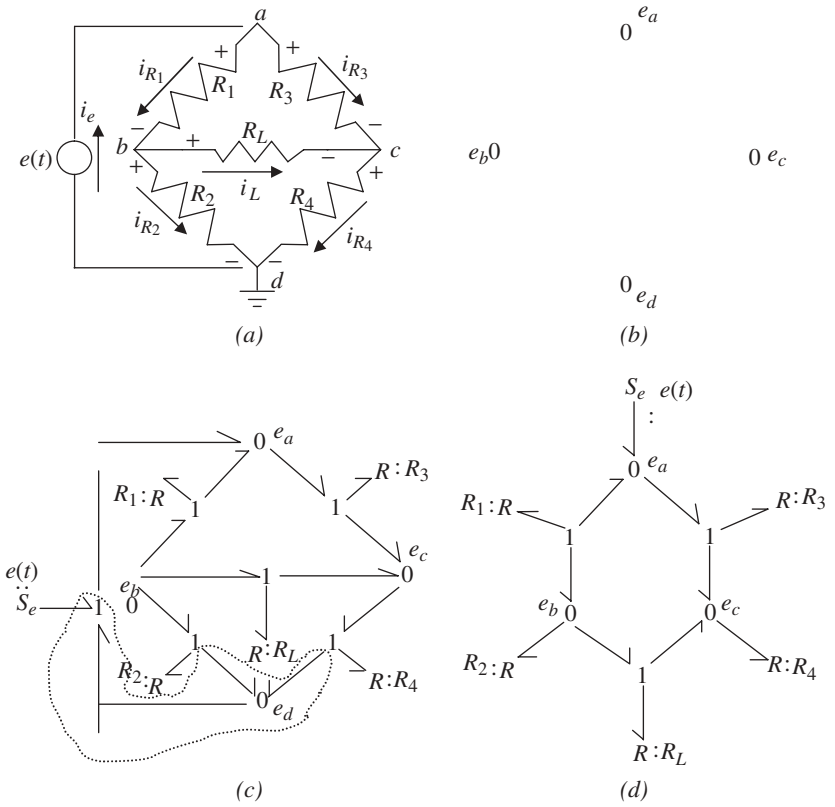


FIGURE 4.3. A Wheatstone bridge circuit. Example 3.

bonds at the ground voltage, $e_d = 0$. Steps 4 and 5 are done in Figure 4.3*d*. The bonds with no power are erased, and the bond graph simplifications have been done, where all 2-port 1-junctions with a "through" power convention have been reduced to single bonds.

The resulting final bond graph reveals a beautiful structural symmetry quite similar to that of the bond diagram for the hydrocarbon molecule called the "benzene ring." It was this similarity to chemical bond diagrams that prompted the "bond graph" name of our modeling approach. The idea is that physical systems consist of components or elements bonded together by power interactions much as atoms are bonded together in chemistry.

4.1.2 Electrical Networks

An electrical network is an extension of electrical circuits to include idealized transformers and gyrators (see Reference [3]). An electrical transformer is a common electromagnetic device used to step voltages up or down while doing the

opposite to the current. Electrical gyrators are exhibited in Hall effect transducers (see Reference [1]), where voltage across a semiconductor material is related to a current through the material perpendicular to the voltage drop direction. The basic rules for bond graph construction remain unchanged.

Figure 4.4*a* shows the electrical symbol for a transformer, where N indicates the turns ratio across the device. Positive voltage drops and current directions are chosen such that positive power is *into* the device on the left side and *out of* the device on the right side.

Since ideal transformers and gyrators are power-conserving elements, it is logical to always define positive power such that it flows in on one side and out on the other. In Figure 4.4*b*, 1-junctions are used to establish the positive voltage drops across the input and output sides of the transformer. Notice that on the input side the voltage is $e_a - e_b$ and on the output side the voltage on the transformer output bond is $e_c - e_d$. These voltages are as defined in the schematic of Figure 4.4*a*. (It may require some thought to see that the sign-convention half-arrows on the two 1-junctions correctly imply the defined voltage differences at the two transformer bonds.)

Although the 2-port transformers and gyrators seem to involve four voltages, one at each terminal, they really relate only the voltage *differences* at the two ports. This means that we may have to choose *two reference voltages* on the two sides of the transformer. It is not necessary to have any direct metallic connection between the windings of a transformer so such an *isolating transformer* can be used to decouple the voltages in one part of a network from another part. If the two ground nodes are selected at the two sides of the transformer as in Figure 4.4*b*, then the transformer appears as in Figure 4.4*c*.

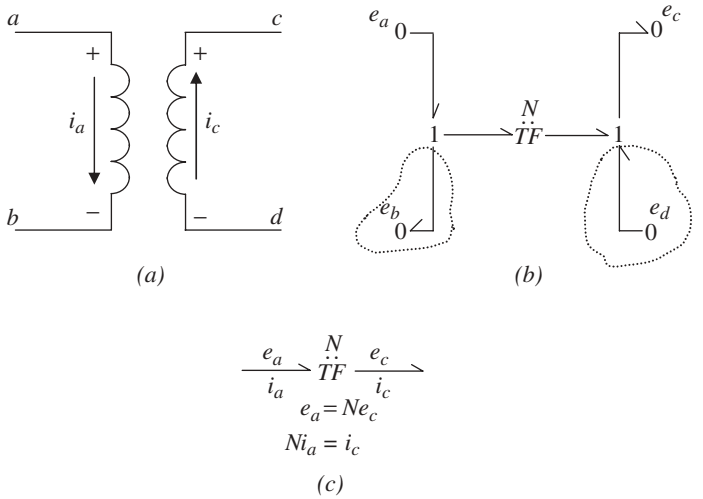
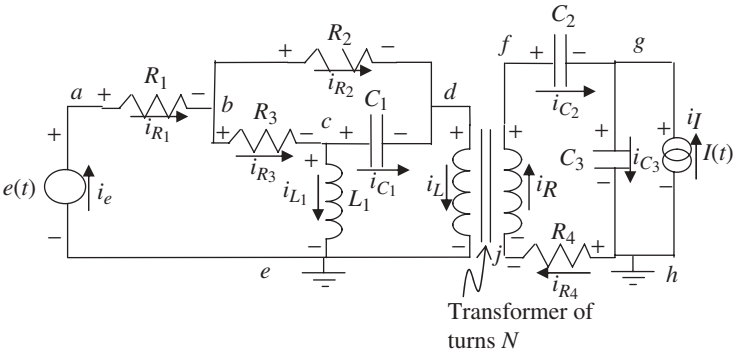


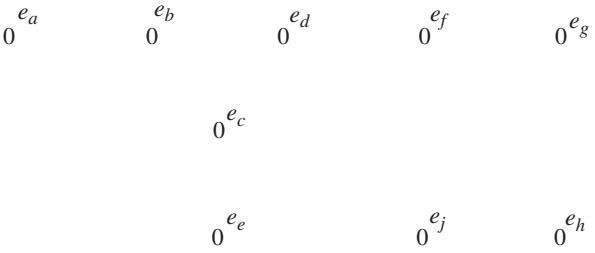
FIGURE 4.4. The ideal electrical transformer and its bond graph representation.

In some networks, each side of a transformer is connected to a common ground. If this is the case, $e_b = e_d = 0$. Then their associated bonds can be erased, some reductions can be performed, and the typical appearance of a transformer emerges as shown in Figure 4.4*c*.

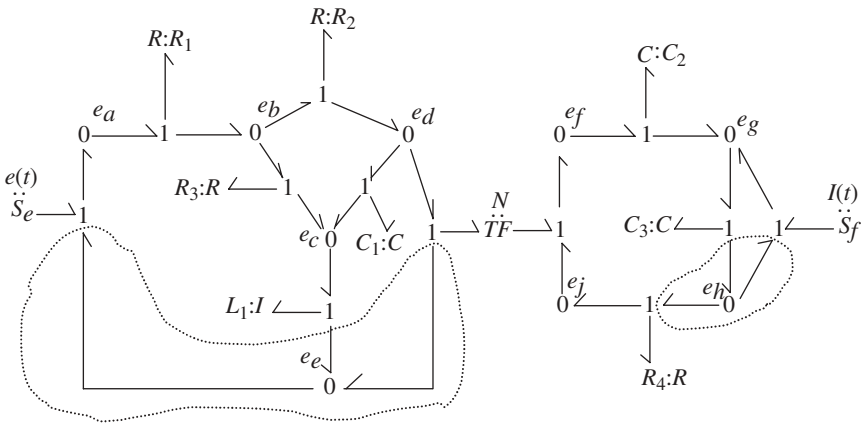
An electrical network with an isolating transformer is shown in Figure 4.5*a*. Positive voltage drops and current directions are shown in this figure. Notice



(a)

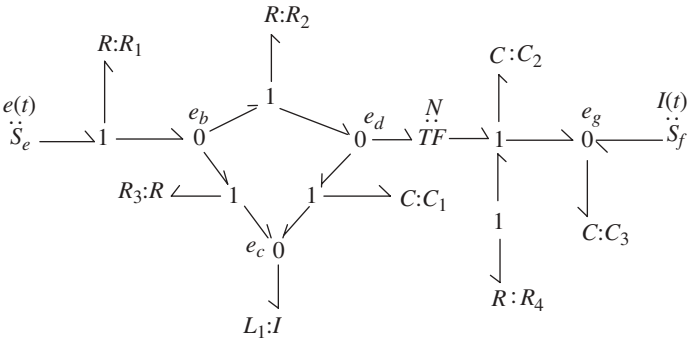


(b)



(c)

FIGURE 4.5. An electrical network with an isolating transformer.



(d)

FIGURE 4.5. (*Continued*)

that positive power flows *in* on the left side of the transformer and *out* on the right side and that *two* ground nodes are specified for the two parts of the network coupled only by the transformer. Following the bond graph construction procedure of the previous section, Figure 4.5*b* uses 0-junctions to expose the node voltages labeled in part *a*, and Figure 4.5*c* uses 1-junctions to establish the proper voltage drops across all the elements. Also in part *c* are dotted lines enclosing all the bonds with zero power. These are erased in Figure 4.5*d*, and bond graph simplifications have been performed to yield the final result.

A final example of the using the construction procedure is the network shown in Figure 4.6*a*. Most of the network is composed of circuit elements that must be looking pretty routine by now, but the new component is the *voltage-modulated current source*, or "controlled current source," labeled $I(t)$ in the figure. The meaning of this schematic symbol is that the voltage, e, from the left side of the circuit modulates, with virtually no power, the current source that is an input to the right side of the circuit.

Active bonds were introduced in Section 2.4 and here an active bond is used to model a device such as a transistor in which a voltage on one terminal has a major effect on the current flowing out of another terminal. Almost no current is associated with the controlling voltage. Of course, this model is valid only for a limited range of voltages and currents, but the point is that, in the model, the voltage signal has zero current associated with it. Thus, the active bond with a full arrow resembles a *signal* in a block diagram without the back effect normal power bonds have.

Positive voltage and current directions are shown in the schematic diagram of Figure 4.6*a*. The labeled node voltages are shown using 0-junctions in Figure 4.6*b*, and 1-junctions are used in Figure 4.6*c* to establish voltage drops across the elements. Of particular interest is the establishment of the modulating voltage, e, and the use of the modulated flow source, S_f, for the current source generating $I(t)$.

Note that the 2-port 0-junction on which the voltage e appears has both an active bond and a normal bond. However, since the active bond implies zero current, the

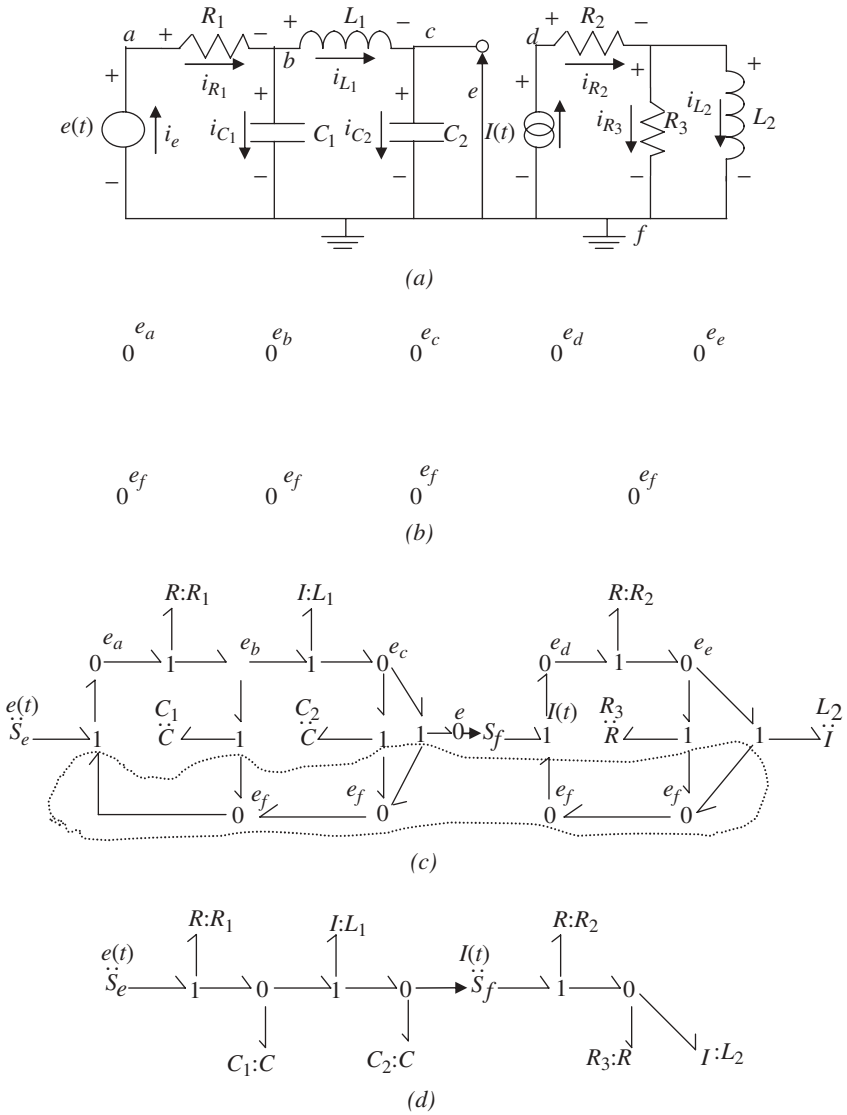


FIGURE 4.6. An electrical network with a controlled source.

normal bond also has zero current. (On any 2-port 0-junction the sum of the currents adds to zero—so if one current is zero, so is the other.) This normal bond is attached to a 1-junction, which also then has zero current on all of its bonds.

Altogether, the bond graph elements imply that the active bond is an effort signal (since it stems from a 0-junction) and that the effort is $e = e_c - e_f$ (because of the signs on the 1-junction bonds) and that the currents on all the local bonds are zero (because of the active bond). The schematic diagram is intended to imply

all of this, but the bond graph makes these modeling assumption more explicit. The rest of the network is processed in Figure 4.6*c* and *d* in the same fashion as in the previous examples.

The zero ground voltage applies to the bonds enclosed by the dotted line, and the zero power bonds have been erased and simplifications performed to yield the final result of Figure 4.6*d*.

The network construction method presented here will allow you to model the most complicated electrical networks once you have acquired facility with the basic steps of node representation, element insertion, power definition, ground definition, and graph simplification. With practice you will begin to observe certain recurrent patterns, and your ability to do more by inspection will increase. Many interesting discoveries await you as you explore systems containing structures like *ladders*, *pi*'s, and *tees*.

Speaking of ladders, let us make use of our ability to model the topology of circuits by bond graph junction structures in order to examine the bond graph form of a ladder network. An example of a resistive ladder circuit is given in Figure 4.7*a*. We are not really concerned with the nature of the 1-ports (source and resistances), but rather with their interconnection patterns (or circuit topology).

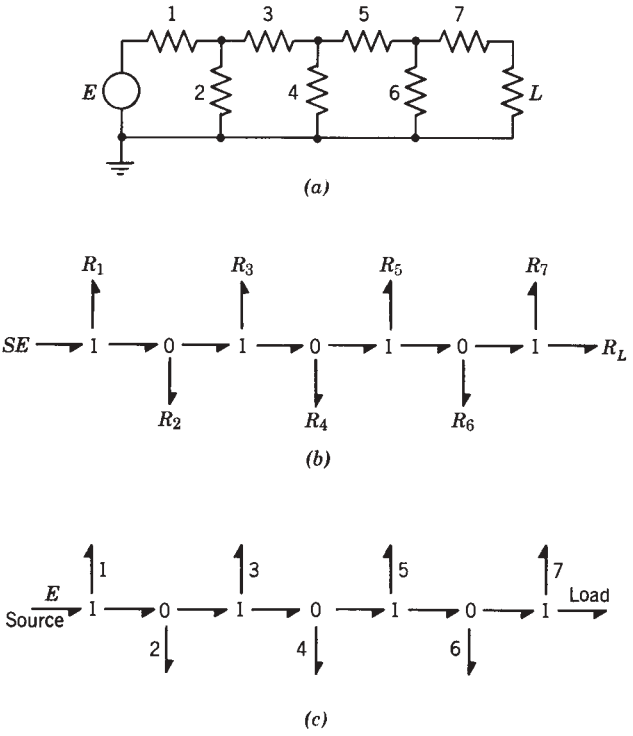


FIGURE 4.7. A ladder network example.

The bond graph of part *b* may be found by inspection, using arguments like "..., R_2 in parallel, R_3 in series, R_4 in parallel, ..." Or the formal construction procedure may be used. To display the system structure without regard to what is (or might be) connected to it, we use a bond graph like that of Figure 4.7*c*, which is a direct representation of the ladder structure. So ladders are nothing more than chains of alternating 3-port 0- and 1-junctions.

Next we shall use such a graph to explore the idea of dual topologies in circuits. A pi network is shown in Figure 4.8*a* in resistive form. The pi gets its name because of the Greek letter pi-like appearance of the schematic. The structure is represented as parallel–series–parallel in part *b* by the bond graph. Resistances 1, 2, and 3 would go on the corresponding bonds to complete the model. In Figure 4.8*c* a tee network is shown, again in resistive form. The tee gets its name from the T-like appearance of the schematic. The bond graph in part *d* is a series–parallel–series type. We observe that part *d* can be obtained from part *b* by switching the roles of 0 and 1, and part *b* can be obtained from part *d* similarly. The formal idea behind the switching is that if the roles of voltage and current are interchanged, a *dual* network results. In terms of bond graph structure, this implies a switching of 0 and 1 elements. By this technique the topological dual of a complex circuit may be obtained from its bond graph in a very simple fashion.

The pi and tee structures are simplified versions of the ladder structure of Figure 4.7. We will see later in the text that electrical transmission lines and long hydraulic lines have models that can be represented by ladder structures that represent a large number of short segments cascaded in the manner of finite element representations. The pi and tee structures become truncations of the ladder structure that allow for reasonable inclusion of some of the transmission line dynamics without having the overall system become too large.

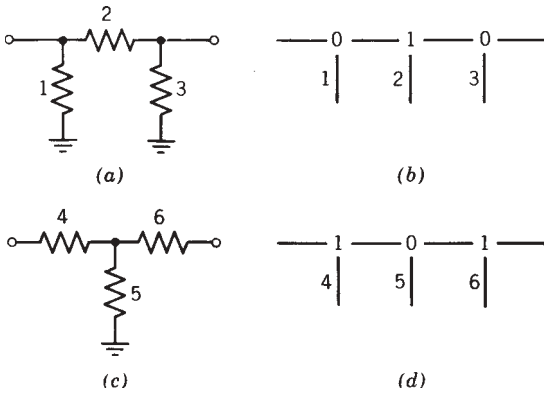


FIGURE 4.8. Truncations of the ladder structure—the pi and tee structures.

4.2 MECHANICAL SYSTEMS

Mechanical systems are those composed of such components as masses, springs, dampers, levers, flywheels, gears, shafts, and so forth. When dynamic systems are put together from these components, we must interconnect rotating and translating inertial elements with axial and rotational springs and dampers, and we must appropriately account for the kinematics of the systems. Bond graphs are well suited for this task. In the next section we address the specifics of mechanical translation; that is, we restrict our attention to mechanical systems moving in a straight line. Then in Section 4.2.2 we focus on mechanical systems that rotate, such as flywheels connected by a torsionally flexible shaft or a gear pair in a transmission. Finally, in Section 4.2.3, we study the dynamics of plane motion where inertial elements have both mass and rotational inertia and bodies can rotate and translate simultaneously.

4.2.1 Mechanics of Translation

In the previous sections dealing with electrical systems, the construction procedure started with identifying the important node voltages and representing them with 0-junctions. Then 1-junctions were used to create appropriate voltage drops across the elements. Finally, the reference voltage had to be identified and eliminated to yield the final model.

For mechanical systems, the effort variable is the force and the flow variable is the absolute velocity (velocity with respect to inertial space). It has been found from years of experience that constructing bond graph models for mechanical systems is easier if a procedure is followed that is dual to the procedure for electrical systems. For mechanical systems, we start by representing system velocities using 1-junctions (common velocity, forces add) and then we create the appropriate relative velocities across the springs and dampers using 0-junctions (common force, velocities add). This is best illustrated with a simple example as shown in Figure 4.9. The system consists of a mass, m, suspended in a gravity field from a spring with stiffness and damper with damping constant b, attached to an inertial frame.

Just as with the electrical systems, a power convention must be established using the schematic diagram. This is a necessary step regardless of the modeling procedure being used. Figure 4.9*a* shows the assumed positive velocity direction for the mass and that the spring and damper relative velocities are assumed to be positive in extension rather than in compression. This is indicated by the symbol $+T$ (for *tension*) in Figure 4.9*a*.

Notice that the inertial reference is labeled as a velocity equal to zero. This is our reference velocity. This 1-junction and all its bonds ultimately will be erased, just as the ground 0-junction was eliminated in electrical systems. (In later examples, it will be seen that as long as we always define *mass velocities*

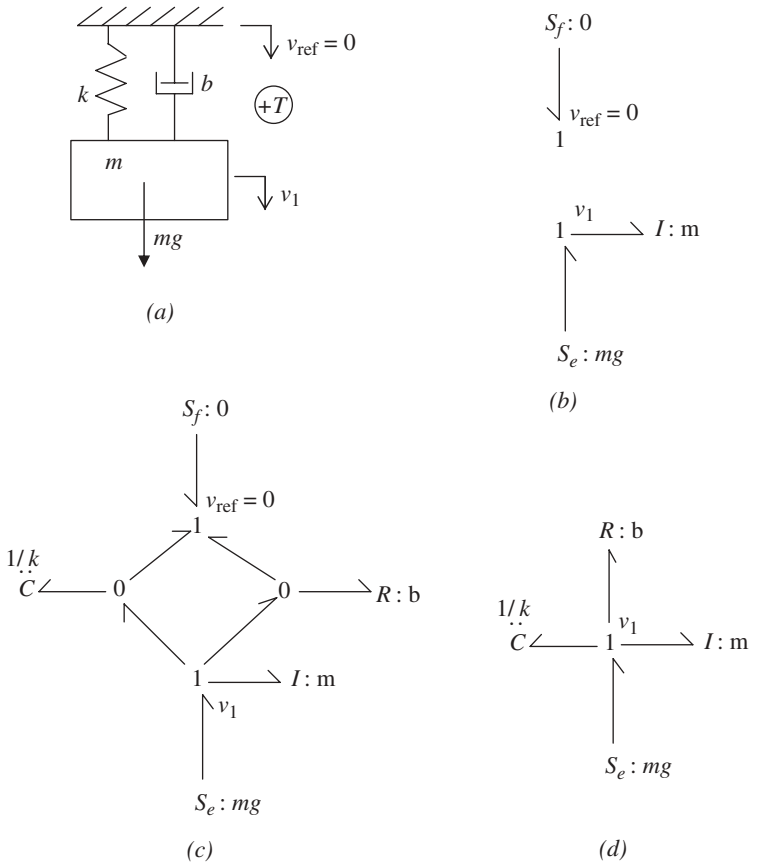


FIGURE 4.9. Mechanical translation. Example 1.

with respect to *inertial space*, we will not need to define an inertial velocity 1-junction to be subsequently be removed.)

In Figure 4.9*b*, a 1-junction is shown such that any bonds that are connected to this junction will have the velocity v_1, and a second 1-junction is shown such that any bonds connected to it will have the reference velocity of zero. The mass is a 1-port inertia that has the absolute velocity, v_1, so the —I element is attached to that 1-junction. The gravity force, mg, is modeled as an effort source in the bond graph (any force that is a known input to a system, whether time varying or constant, is modeled as an effort source in a bond graph). Since this force is moving at the velocity v_1, the effort source is attached to the 1-junction representing this velocity. The power out convention on the effort source comes from the fact that if the velocity is in the positive, downward direction and the gravity force is acting downward, then positive power is coming from the source

into the system. A flow source equal to zero is attached to the reference 1-junction to enforce that the velocity is zero.

The spring and damper in general react to the relative velocity across them. The choice of how to properly add the velocity components at each end of these elements is a topic that causes much student frustration. For this first example, we simply note that for the elements with forces positive in tension, with the positive velocity directions indicated, the relative velocity across both the spring and damper should be $v_1 - v_{\text{ref}}$. That is, if $v_{\text{rel}} = v_1 - v_{\text{ref}}$ is positive, then the spring and the damper are extending. This is accomplished using 0-junctions with the power convention indicated in Figure 4.9*c*.

The reader should verify that the velocities on the C-element and the R-element bonds really do represent the *rate of extension* of these elements. If this is so, when the forces are positive in tension, power will flow toward the C- and R-elements as indicated by their sign half-arrows. One should review the velocity summations in Eqs. (3.18) and (3.18*a*) to check the sign conventions used in Figure. 4.9*c*.

Finally, in Figure 4.9*d*, the bonds with zero power are removed and the final bond graph emerges after some bond graph reductions are used. It makes sense that, since all the elements have the same velocity, they all end up on the same 1-junction. We will be able to do many mechanical system models by inspection after we gain some experience.

To formalize the bond graph construction procedure for mechanical translation, here are the steps to follow:

Construction Procedure for Mechanical Translation

1. **On a schematic diagram of the physical system, use arrows and symbols to indicate the positive direction of absolute velocity components**.

   These elements include all individual mass elements, all prescribed input velocities, and the velocities of any other physical locations that may prove useful in establishing useful relative velocities. State whether force-generating elements, springs and dampers, are positive in tension or compression using symbols such as $+T$ or $+C$.

2. **Use 1-junctions to represent each distinct velocity from step 1**.

   Label the 1-junction with the velocity symbol from the schematic diagram. This will help to remind you which junction is associated with which velocity component. You might use a 1-junction to represent the reference of zero absolute velocity. This will later be eliminated, but it might be helpful for establishing relative velocities.

3. **Attach to each 1-junction any element that relates to the absolute velocity represented by the junction.**

   In general, mass elements are inertias associated with absolute velocities represented by some of the 1-junctions. Remember, positive power is *always* directed into an —I, —R, or —C element.

4. **Use 0-junctions to establish proper relative velocities.**
   Use sign-convention half-arrows so that the connecting elements are positive in compression or tension as was assumed in the schematic diagram. This involves the sign-convention rules discussed in Section 3.3.
5. **Eliminate the bonds with zero velocity and reduce to the final model.**

As an example of the use of the construction procedure, consider the so-called *quarter-car model* of Figure 4.10. This represents one corner of an automobile

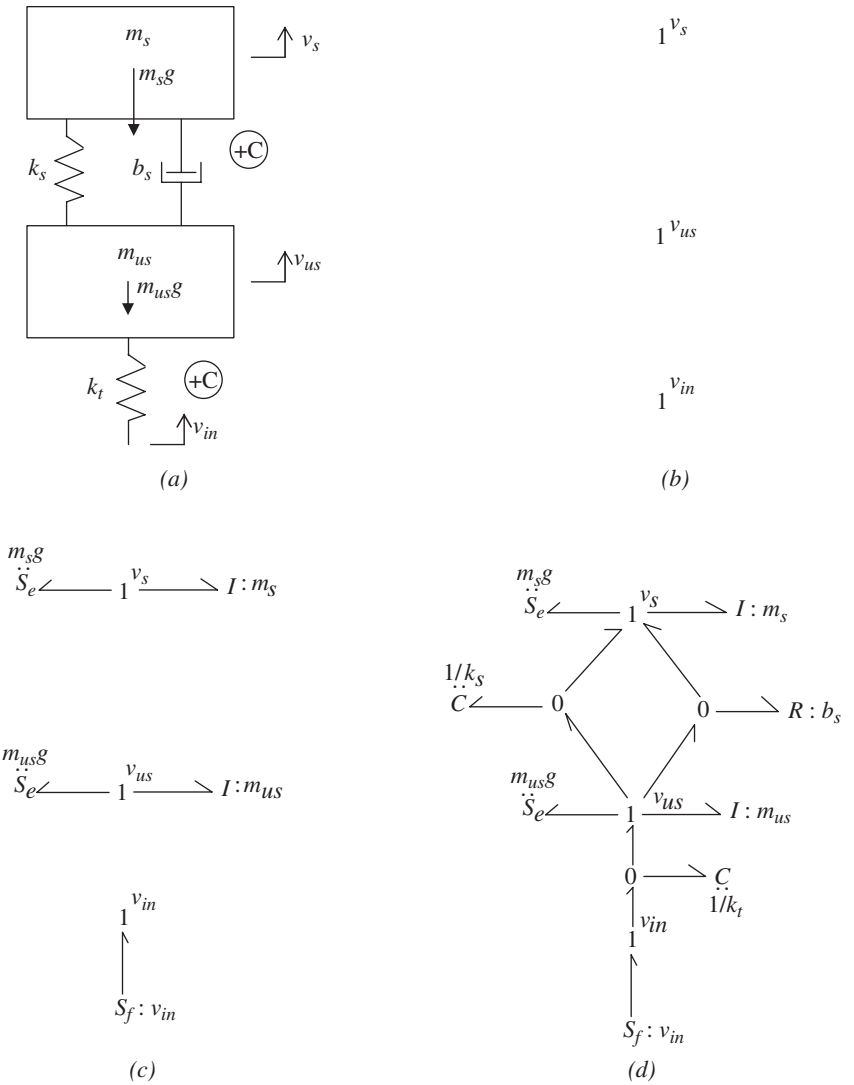


FIGURE 4.10. Quarter-car model. Example 2.

where m_s is the sprung mass (one-quarter of the body mass); m_{us} is the unsprung mass, which includes the tire, wheel, and some part of the brakes and suspension; k_s, b_s are the suspension spring and damper constants; and k_t is the tire stiffness.

Quite some liberty has been taken in this model in representing the suspension and tire as linear elements. The system is constrained to move only vertically, and the velocity input, $v_{\text{in}}(t)$, at the base is representing the roadway unevenness experienced as the vehicle moves forward. The vehicle is also under the influence of gravity acting vertically downward.

Construction rule 1 is used in the schematic diagram of Figure 4.10*a*. The distinct velocities are labeled with arrows indicating positive directions, and, as indicated, the springs and dampers are all assumed positive in compression. In Figure 4.10*b*, step 2 is performed, where 1-junctions are used to represent each distinct velocity, and the junctions are labeled to remind us which velocity they represent. In this case, all the velocities are inertial velocities and we have not included a reference velocity as was done in the first example.

Step 3 is accomplished in Figure 4.10*c*, where the inertial —*I* elements, with inward power convention, are attached to the appropriate 1-junctions and labeled to indicate which mass they represent. Notice that the velocity input is represented by a flow source, S_f, attached to the 1-junction for v_{in}. The power arrow is out from the flow source because if the tire spring is in compression (positive by assumption) and the input velocity is up (positive by assumption), then power is flowing from the source into the system.

The weight of each mass element, $m_s g$ and $m_{\text{us}} g$, is modeled as an effort source attached to the 1-junctions having the velocity of the associated mass. The power arrow is directed away from the 1-juntion and into the source because if the respective mass element is moving upward (positive by assumption) and the gravity force is acting downward (as it always must), then power is flowing from the system and into the source representing the gravity force. (The weight force source in the first example had a different sign convention because there, the absolute velocity was defined to be positive downward whereas here the velocities are positive upward.)

For step 4, shown in Figure 4.10*d*, 0-junctions are used to establish relative velocities across the remaining elements. For the positive velocity directions defined, and the positive in compression force directions defined, the proper relative velocity across the tire spring is $v_{\text{in}} - v_{\text{us}}$ and the proper relative velocity across the suspension spring and damper is $v_{\text{us}} - v_s$. These are the proper relative velocities because, if they are positive, then the respective elements will, in fact, be compressing. Notice that the relative velocities across the suspension spring and suspension damper are independently constructed, as dictated by the construction procedure. This is perfectly correct and can always be done; however, a simpler realization for these relative velocities will be shown below. There is one simplification that could be done for this model. The 2-port 1-junction at the bottom of Figure 4.10*d* could be removed because it has the sign half-arrows passing "through" it, and the flow source could be attached directly to the 0-junction. (See again Section 3.3.)

It should be recognized that we paid attention only to establishing the proper velocity components for this construction procedure, and we paid no attention to the forces. The beauty of using the power conservation properties of bond graph junctions is that we need only to constrain the velocities and the forces will automatically be balanced. For mechanical systems, if we establish proper velocities, the forces are guaranteed to be correct.

To see this, consider the free-body diagram for the quarter-car example, shown in Figure 4.11. The masses are isolated and the forces from the springs, dampers, and gravity are exposed. The arrows for the forces are directed in the assumed positive directions, namely, in compression for this example, and the positive velocities are directed upward as was assumed positive previously. Notice that we must show all the forces as having equal but opposite effects, as required by Newton's laws. If we were to sum the forces on the sprung mass, positive upward, the resultant force, F_s, would be

$$F_s = F_{k_s} + F_{b_s} - m_s g, \tag{4.1}$$

and for the unsprung mass, the resultant force, F_{us}, would be

$$F_{\text{us}} = F_{k_1} - F_{k_s} - F_{b_s} - m_{\text{us}} g. \tag{4.2}$$

On the bond graph of Figure 4.10*d*, the forces on the respective mass elements are the efforts associated with the —*I* elements emanating from the 1-junctions labeled v_s and v_{us}. The 1-junctions, in addition to being common flow or velocity, add efforts, or forces, according to the power convention. Amaze yourself by

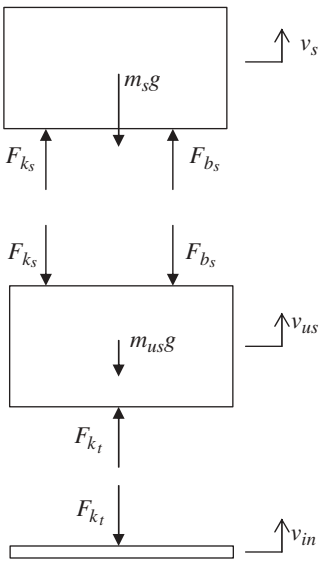


FIGURE 4.11. Free-body diagram of the quarter-car system.

adding the forces on the —*I* element bonds and seeing that the forces add exactly as required by Eqs. (4.1) and (4.2). The forces were constrained without further effort (no pun intended) after enforcing the velocity constraints. And we never had to show equal and opposite reaction forces.

The construction procedure used in the example for the suspension spring and damper calls for the establishment of the relative velocity across each element using 0-junctions to appropriately add velocity components. Such adherence to the construction rules will always produce a correct result. The resulting bond graph structure, referred to as a *loop with a through power convention* or a *reducible loop*, comes up a lot in bond graph modeling. (We saw a similar loop in the electrical system of Figure 4.2.)

This loop is isolated in Figure 4.12*a*, where it has been generalized using efforts and flows rather than forces and velocities. The relative velocities across the spring and damper are identical, and each is equal to $f_1 - f_2$.

In Figure 4.12*b*, the relative velocity is established using a single 0-junction, and then a 1-junction is used to ensure that any bond attached to that 1-junction will have the relative velocity, $f_1 - f_2$. Since both the —*C* and —*R* elements have this relative velocity, they both get attached to the 1-junction as shown in Figure 4.12*b*. The reader can check to see that the forces still add up properly in this alternative representation. The result is a slightly simpler bond graph emphasizing that several elements have the *exactly* the same relative velocity. (It is possible to make a mistake with a loop by defining one side as positive in extension and the other in compression, which doesn't square with the schematic diagram. The pattern in Figure 4.12*b* prevents that type of mistake, since in that version, both elements clearly have the same relative velocity.)

In addition to mechanical systems having one-directional translational motion, it requires only a modest extension of our procedure to deal with translational systems containing levers, pulleys, and other simple motion–force transforming devices. Consider the system shown in Figure 4.13*a*. A velocity input is prescribed in the horizontal direction on the end of a spring with constant k_1. The

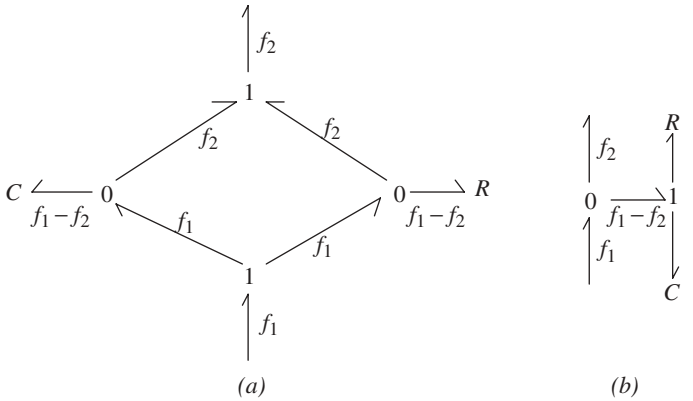
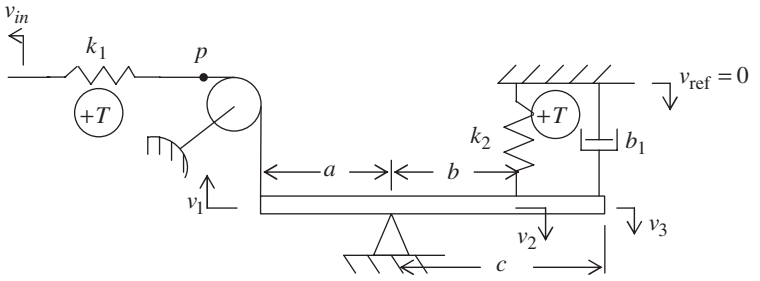
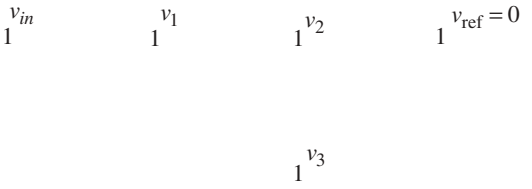


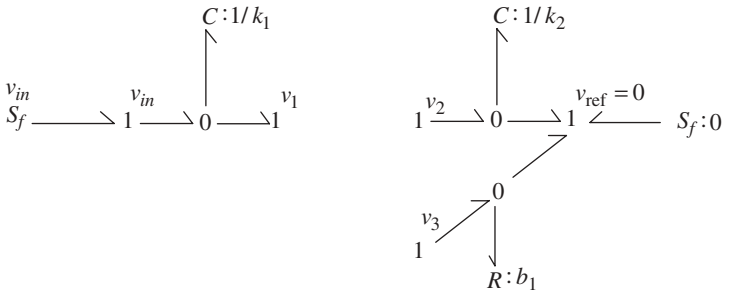
FIGURE 4.12. Loop with through power convention that can be reduced.



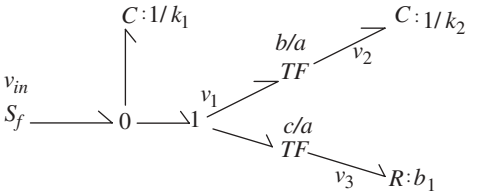
(a)



(b)



(c)



(d)

FIGURE 4.13. A pulley–lever mechanism.

other end of this spring goes around a pulley and causes vertical motion at one end of the massless lever. A spring with constant k_2 and a damper with constant b_1 are located on the right side of the lever with one end of each attached to inertial ground. The pivot for the lever is located a distance a from the left end, while the spring and damper are located at distances b and c to the right of the pivot. Figure 4.13*a* in addition shows positive velocity directions and indicates that both springs and the damper are assumed positive in tension. Notice that v_1 is labeled only on the end of the lever, but it is the same as the horizontal velocity at the point P.

Following the construction procedure, we first use 1-junctions to represent the distinct velocities. A good question is why these particular velocities were chosen as distinct and worthy of representation with their own 1-junctions. As mentioned in step 1 of the construction procedure, velocities are always assigned to distinct masses, as well as to other physical locations that might prove useful in establishing relative velocities across elements. You can never overspecify velocity components. If some 1-junctions turn out to not be needed or redundant, they can be eliminated during bond graph simplification.

Looking carefully at Figure 4.13*a*, we see that v_1 (which is the velocity at point p) will be useful for establishing the relative velocity across the horizontal spring, and velocities v_2 and v_3 are needed for the relative velocities across the vertical spring and damper. Figure 4.13*b* shows the 1-junctions with the important velocities labeled. The only elements that have any of these specific absolute velocities are the source elements that establish the input velocity, $v_{\text{in}}(t)$, and the reference velocity, $v_{\text{ref}} = 0$. These are shown in Figure 4.13*c*, where the relative velocities across the R- and C-elements have been established using 0-junctions.

Convince yourself that the half-arrows are correct for having all elements positive in tension. For example, with the velocities defined positive as they are, the spring with constant k_1 will be put into tension by the relative velocity, $v_{\text{in}} - v_1$, and this is how the flows add on the corresponding 0-junction. The other elements follow a similar argument. The rules of construction have been followed but the model is obviously not complete. There is no connection between the 1-junctions for v_1, v_2, and v_3.

From Chapter 3 we know that a massless lever is a transformer relating velocities and forces across it according to the constitutive relationships,

$$\frac{b}{a}\, v_1 = v_2 \quad \text{and} \quad \frac{c}{a}\, v_1 = v_3. \tag{4.3}$$

(We are assuming that the lever moves through a small enough angle such that the motion of the points at which other elements are attached is nearly in the vertical direction.)

Figure 4.13*d* shows the final model with the transformers installed, the reference velocity eliminated, and some simplifications performed.

One of the many valuable features of bond graph modeling is the relative ease with which models can be modified. Let's say that a mass that we neglected

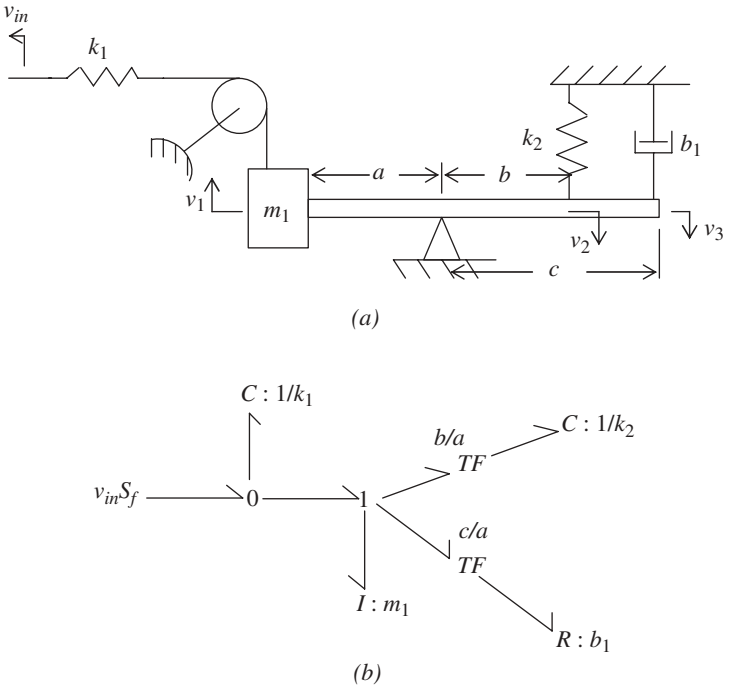


FIGURE 4.14. Modified pulley–lever mechanism.

in the schematic diagram of Figure 4.13*a* turns out to be important. The modified schematic is shown in Figure 4.14*a*. No additional velocities need be labeled, since v_1 was needed in the previous version of the model. Looking at Figure 4.13*c* where v_1 is exposed, the new mass element, m_1, simply becomes an —*I* element attached to the v_1 1-junction. Nothing else changes, and the final result is shown in Figure 4.14*b*.

A straightforward extension of translational mechanical systems is fixed-axis rotation, presented next.

4.2.2 Fixed-Axis Rotation

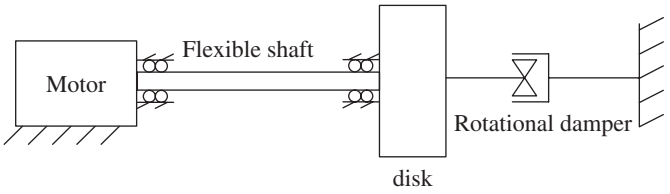
Fixed-axis rotation occurs in the mechanics of rotating machines, such as electric motors and internal combustion engines. Gear boxes, transmissions, transfer cases, differentials, and drive shafts are all examples of systems that have some aspect of fixed-axis rotation. The procedure for modeling mechanical translation need only be modified slightly to handle systems with fixed-axis rotation. Because it is a straightforward extension of what we just did, this section starts with the procedure for bond graph construction of fixed-axis rotation.

Construction Procedure for Fixed-Axis Rotation

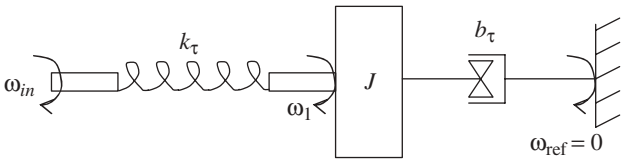
1. **On a schematic diagram of the physical system, use arrows and symbols to indicate the positive direction of absolute angular velocity for components.**
   These components include all individual rotational inertial elements, all prescribed input angular velocities, and the angular velocity of any other physical locations that may prove useful in establishing useful relative angular velocities. Decide on positive directions of twist for elements such as rotational springs and dampers. (The positive direction of relative rate of twist is hard to show in a schematic diagram but will be easy to see in a bond graph.)
2. **Use 1-junctions to represent each distinct angular velocity from step 1.**
   Label the 1-junction with the angular velocity symbol from the schematic diagram. This will help to remind you which junction is associated with which angular velocity component. You might use a 1-junction to represent the reference of zero angular velocity. This will later be eliminated, but it might be helpful for establishing relative angular velocities.
3. **Attach to each 1-junction any element that has that angular velocity.**
   In general, distinct rotational inertias are —I elements associated with distinct angular velocities represented by some of the 1-junctions. Remember, positive power is *always* directed into an —I, —R, or —C element.
4. **Use 0-junctions to establish proper relative angular velocities.**
   Use the sign conventions on the 0-junctions to make the relative velocities for the connecting elements conform with the twist directions assumed in the schematic diagram.
5. **Eliminate the bonds with zero angular velocity and reduce to the final model.**

As the first example, consider a model of a grinding wheel at the end of a torsionally flexible shaft, shown in Figure 4.15*a*. An electric motor is assumed to prescribe an input angular velocity, $\omega_{\text{in}}(t)$, at the left end of the shaft. The shaft is flexible and characterized by its torsional stiffness, k_τ, with units N-m/rad. The right end of the shaft has a disk of rotational inertia, J, with units kg-m^2. There is also rotational damping with a coefficient having b_τ the units N-m/(rad/s), attached between the disk and ground.

Figure 4.15*b* is a schematic diagram of the physical system with angular velocities labeled and positive directions given. The twist in the shaft is assumed to be positive in the clockwise direction as viewed looking into the ends of the shaft, and the rotational damper torque is positive for clockwise rotation as viewed looking into end of the damper. You will notice that it is harder to show positive angular velocities clearly in a two dimensional sketch than it is to show positive linear velocities.



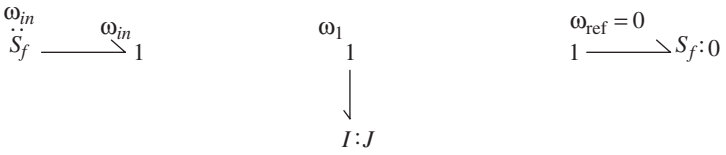
(a)



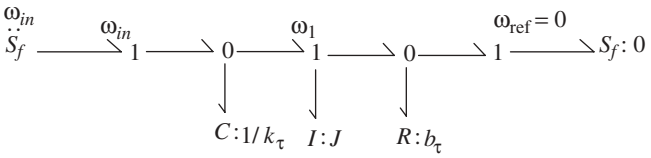
(b)



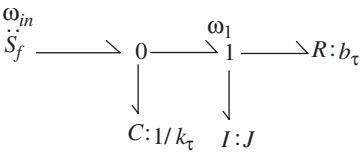
(c)



(d)



(e)



(f)

FIGURE 4.15. A grinding wheel model to demonstrate fixed-axis rotation. Example 1.

Figure 4.15*c* shows 1-junctions representing the distinct angular velocity components, including the reference of zero angular velocity. In Figure 4.15*d*, the flow source that prescribes $\omega_{\text{in}}(t)$ is attached to the appropriate 1-junction, and the rotational inertia is attached to its appropriate 1-junction. In Figure 4.15*e*, 0-junctions are used to establish relative angular velocities across the torsional spring and torsional damper. Note that the rate of twist of the torsion spring is $\omega_{\text{in}} - \omega_1$, which is hard to indicate on the schematic diagram without using equations. In Figure 4.15*f*, the inertial reference is removed and some simplifications are carried out to reduce to the final model.

A more complex example of fixed-axis rotation is shown in Figure 4.16. Disks of rotational inertias J_1 and J_2 are rigidly attached to each end of a flexible shaft

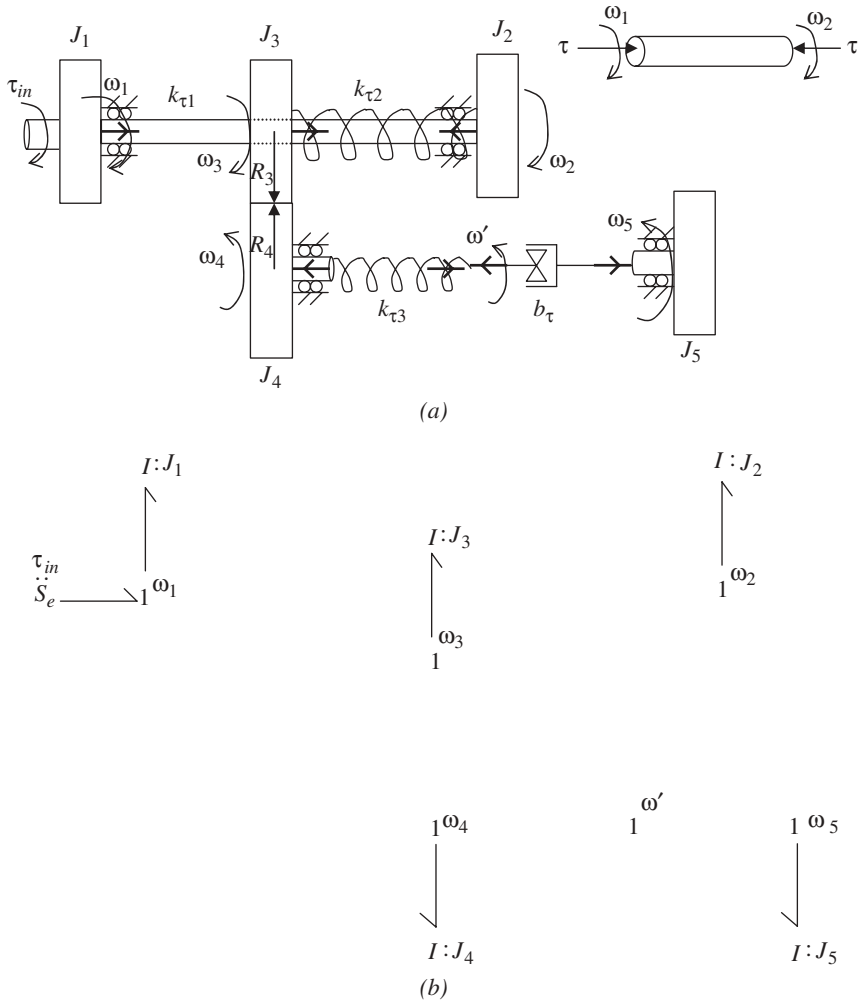
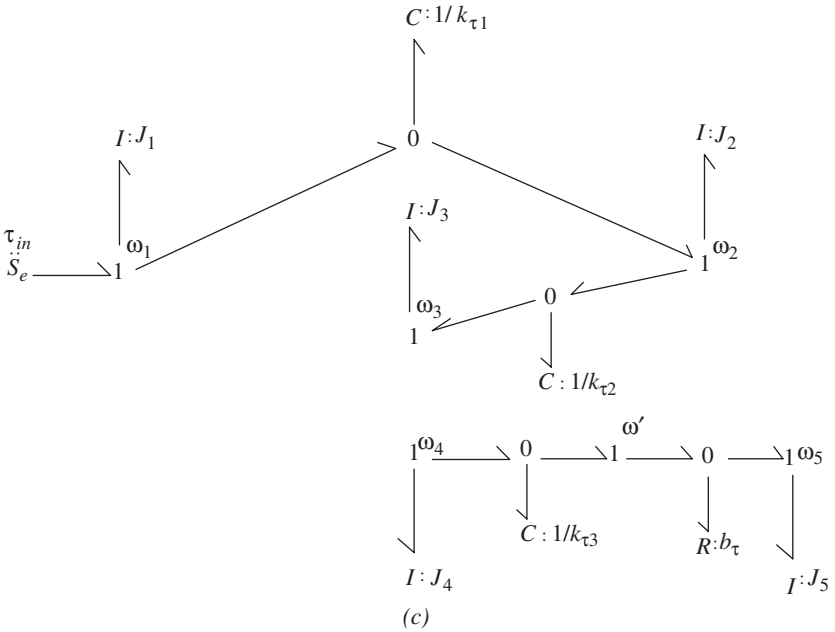


FIGURE 4.16. A system of shafts and gears. Example 2.



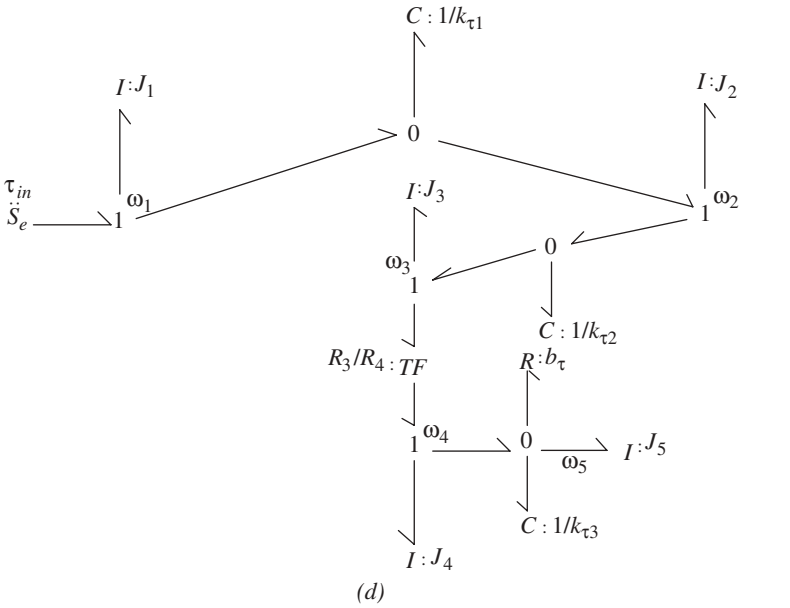


FIGURE 4.16. (*Continued*)

with torsional stiffness, k_{τ_1}. Disk 2 is attached at one end to a second torsional spring with constant k_{τ_2}, which is attached to a third disk, J_3, that is free to rotate on the shaft. This third disk is of radius R_3 and forms a gear set with the fourth disk with moment of inertia J_4 and radius R_4. Finally, disk 4 is attached through a torsional spring and damper, parameters k_{τ_3} and b_τ, to a fifth disk of rotational inertia, J_5. Note that an intermediate angular velocity, ω', has been labeled in between the rotational spring and damper. This was done for convenience in anticipation of the need to establish the relative angular velocities across these two elements. We would ultimately like to predict the motion–time behavior of this system for a specified input torque, $\tau_{\text{in}}(t)$.

On the physical schematic the positive angular velocity directions are indicated with arrows and labels, and the positive torque directions for the springs are indicated by the lines with full arrows using the right-hand-rule. The isolated shaft accompanying the schematic shows this notation.

In Figure 4.16*b*, 1-junctions have been used to represent all the distinct angular velocities from the schematic. Moreover, the bond graph elements that have these specific angular velocities have been attached directly to these 1-junctions. These include the rotational inertias and the input torque source that is moving at the angular velocity ω_1. In Figure 4.16*c*, 0-junctions have been used to establish relative angular velocities across the rotational spring and damper elements. The power convention for the 0-junctions was chosen to establish the positive torque directions from the schematic. Notice how useful the 1-junction for ω' is for establishing the relative angular velocities across the spring and damper with parameters k_{τ_3} and b_τ. Also notice that there is no connection between ω_3 and ω_4. The two disks of rotational inertias J_3 and J_4 are a gear set with the kinematic relationship

$$\frac{R_3}{R_4}\omega_3 = \omega_4. \tag{4.4}$$

This relationship is enforced by a transformer as shown in Figure 4.16*d*. In this figure, some final simplifications were done, including the removal of the unnecessary 1-junction representing ω'. After this 1-junction was removed, the two remaining 0-junctions were connected by a common bond enforcing that all connected bonds had the same torque (effort). All the bonds with a common effort can then be connected to a common 4-port 0-junction as was done in the final step, Figure 4.16*d*. This is a nice example of introducing 1-junctions for flows that might be convenient for model construction, and then having the modeling procedure dictate whether these 1-junctions remain in the final model. The modeler is free to define and introduce as many flows on 1-junctions (or efforts on 0-juntions) as desired. If these extra junctions are actually unnecessary, simplification can remove them.

We now go on to model systems containing rigid bodies that can move in translation as well rotation. Here we treat the important case of plane motion.

4.2.3 Plane Motion

For mechanical translation it was fine to imagine the mass elements as point masses, and for fixed-axis rotation it was fine to consider only the rotational inertia of a mass element. But, in general, mechanical bodies with finite mass both translate and rotate in real applications. Consider an automobile as it moves over roadway unevenness. The body of the vehicle moves forward, sideways, and vertically, and it rolls side to side, pitches front to rear, and rotates about a vertical axis as it turns. If we modeled the car body as rigid, we would have to characterize the body by its mass and the moments of inertia about three perpendicular axes. The motion of the vehicle would be quite complex and certainly not able to be described by mechanical translation alone or fixed-axis rotation alone. Instead, the dynamics are governed by a simultaneous combination of translation and rotation. To describe this motion in three dimensions is quite complicated, but is greatly facilitated by using bond graphs. Chapter 9 discusses this complex topic. When our attention is restricted to plane motion, model construction is straightforward.

Plane motion results when the inertial bodies of a physical system are constrained to translate in two dimensions and to rotate only about an axis perpendicular to the plane of motion. Figure 4.17*a* shows a rigid body of mass, m, and moment of inertia, J, about its center of mass. The X- and Y-axes are assumed to be inertial axes. The body is translating in the XY-plane and it is rotating about the inertial Z-axis, which points out of the page according to the right-hand-rule. Any forces acting on the body from attached devices are not shown in the figure.

It is useful to describe the motion of the body by describing the translation of the center of mass and the angular velocity of the body. The center of mass has an absolute velocity vector pointing in some direction in the XY-plane. This vector is resolved into two mutually perpendicular components, v_X, v_Y, aligned in the XY-directions as shown on the figure. The angular velocity is the angular rate of change of any line fixed to the body with respect to the XY-axes.

The kinetic energy T of a rigid body is characterized in a particularly simple way by the velocity of the center of mass and the angular velocity of the body,

$$T = \tfrac{1}{2}mv_X^2 + \tfrac{1}{2}mv_Y^2 + \tfrac{1}{2}J\omega^2 \tag{4.5}$$

in which it is notable that by focusing on the motion of the center of mass and using the moment of inertia about the center of mass the energies of translation and rotation are uncoupled.

Since bond graphs bookkeep energy, to account for all the energy of a rigid body in plane motion we must represent the translation velocity of the body in two perpendicular directions as well as the angular velocity. This is done in the bond graph fragment of Figure 4.17*b*. The 1-junctions are used to represent each center of mass velocity component and the angular velocity. Attached to the velocity 1-junctions are —I elements with inertia parameter equal to the mass, m, and attached to the 1-junction for the angular velocity is an —I element with the moment of inertia, J, as the parameter. This fragment characterizes all the

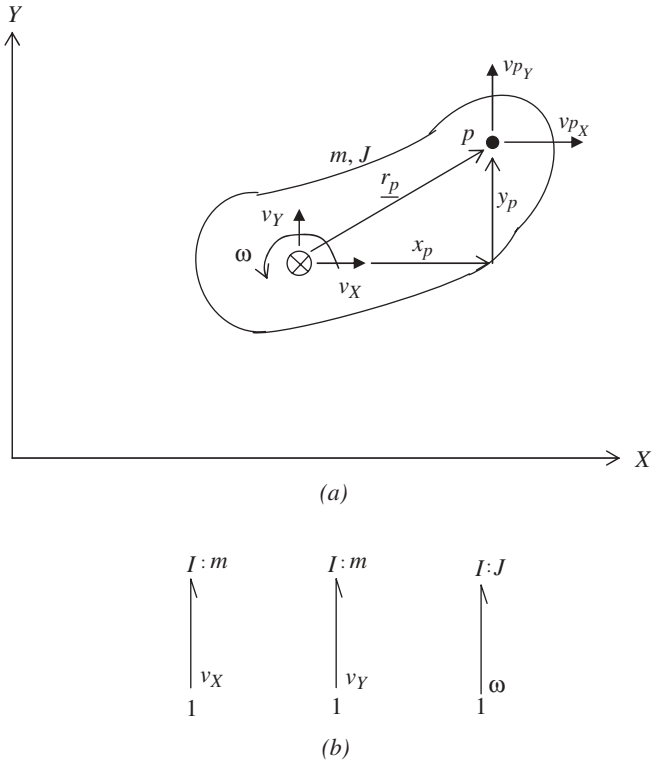


FIGURE 4.17. General plane motion of a rigid body.

energy of a body in plane motion. Any devices, such as springs and dampers, that interact with the rigid body will end up communicating with the 1-junctions of Figure 4.17*b* when incorporated into a bond graph model.

In general, devices will be attached to rigid bodies at different places on the body. For example, the point p indicated in Figure 4.17*a* might be such a point. We will discover that we need the velocity components of the attachment points indicated as v_{P_X}, v_{P_Y} in the figure. The attachment point is located with respect to the center of mass by the position vector $\vec{r}_P$ composed of the components x_P, y_P. The kinematic relationship (see Reference [2]) that allows determination of the velocity of any point on a rigid body is

$$\vec{v}_P = \vec{v}_{cm} + \vec{\omega} \times \vec{r}_P \tag{4.6}$$

where the cross-product term accounts for the contribution of ω to the velocity of the point in question. This relationship gets used over and over when modeling rigid bodies in plane motion. In plane motion, the angular velocity vector is always perpendicular to the position vector, thus making the use of Eq. (4.6) particularly simple.

For the geometry of Figure 4.17*a*, application of Eq. (4.6) yields

$$v_{P_x} = v_X - y_P\omega \quad \text{and} \quad v_{P_Y} = v_Y + x_P\omega. \tag{4.7}$$

It should be noted that for large angular motions of the body, x_P and y_P will change as the body moves. In this case, if r_P is the length of the vector $\vec{r}_P$, then $x_P = r_P \cos\theta$ and $y_P = r_P \sin\theta$, where is the angle between $\vec{r}_P$ and the axis with $\dot{\theta} = \omega$. Equation (4.7) then becomes

$$v_{P_x} = v_X - (r_P \sin\theta)\omega \quad \text{and} \quad v_{P_Y} = v_Y + (r_P \cos\theta)\omega, \tag{4.7a}$$

and one can recognize that *displacement modulated transformers* as discussed in Section 3.2 and shown in Figure. 3.10 will be involved in the bond graph of the system. This type of geometrically nonlinear mechanical system is discussed at length in Chapter 9.

As a simpler first example in which the angle of the rigid body is assumed not to experience large excursions, consider the heave–pitch vehicle model shown in Figure 4.18*a*. This is an extension of the quarter-car model done previously in Figure 4.10. This model is sometimes called the *half-car model*. We are looking at the side of the vehicle with the front at the right and the rear at the left. The front and rear suspensions each consist of a tire springs with stiffnesses k_{t_f}, k_{t_r}, unsprung masses m_{us_f}, m_{us_r}, and suspension springs and dampers with parameters k_{s_f}, b_{s_f}, and k_{s_r}, b_{s_r}. The inputs to the system are the prescribed input velocities at the front and rear, $v_{i_f}(t)$ and $v_{i_r}(t)$. The body is modeled as a rigid body with center of mass located a distance a from the front suspension connection and a distance b from the rear suspension connection. The body has mass, m, and centroidal moment of inertia, J. Gravity is acting vertically downward.

For this model the vehicle is not moving forward but is mounted in a test rig that supplies the input velocities to the tire springs as if a bumpy road were passing under the vehicle. The body is assumed to move only vertically and to rotate, or pitch, through small angles. Thus, the rigid body is characterized by the vertical center of mass velocity, v_g, and the pitch angular velocity, ω. Other important velocities are oriented and labeled in the figure, including the vertical velocity components at the front and rear of the body, v_f and v_r. These components are introduced for convenience in anticipation of construction of the relative velocity across the front and rear suspension units. As indicated in Figure 4.18*a*, all spring and damper forces are assumed positive in compression.

The procedure for constructing a bond graph model of a system with rigid bodies in plane motion is basically identical to the procedure for mechanical translation and fixed-axis rotation, but with the added complexity of dealing with kinematics described by Eq. (4.6). Figure 4.18*b* shows the use of 1-junctions to represent all the distinct velocities and angular velocities of the system. Attached to these 1-junctions are bond graph elements that have these specific velocity and angular velocity components. The front and rear suspension unsprung masses and input velocities should look familiar because of the earlier quarter-car example.

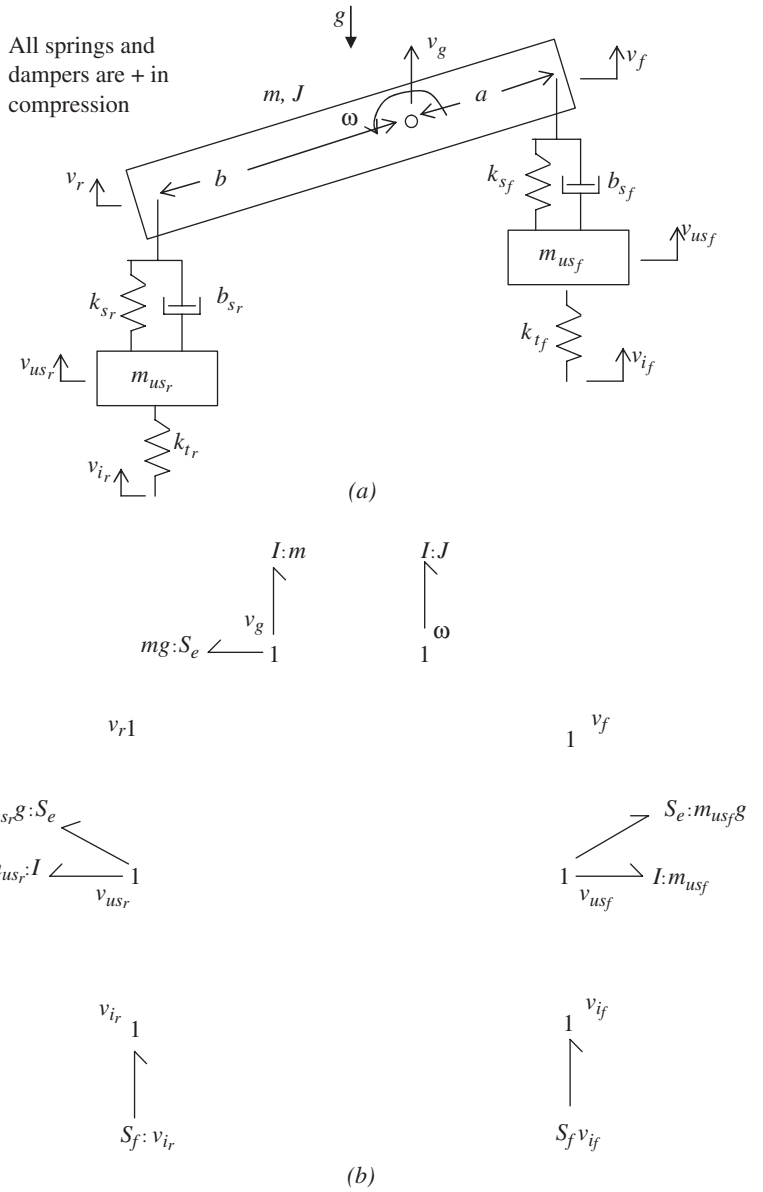


FIGURE 4.18. Heave–pitch vehicle model. Example 1 of plane motion.

The rigid body is characterized by the center-of-mass velocity and angular velocity, so —I elements for the mass, m, and centroidal moment of inertia, J, are attached to the appropriate 1-junctions. The weight of the body is an effort source attached to the center-of-mass vertical velocity 1-junction. Positive power is directed into the source owing to the velocity convention (positive upward) adopted for this example. The same is true for the effort sources representing the weights of the unsprung masses, front and rear.

In Figure 4.18*c*, 0-junctions have been used to establish the relative velocities across the springs and dampers. The power convention is such that all elements are positive in compression as indicated for this system. Some simplifications have already been done. Notice how the 1-junctions for v_f and v_r were used to help in constructing the relative velocities across the front and rear suspensions.

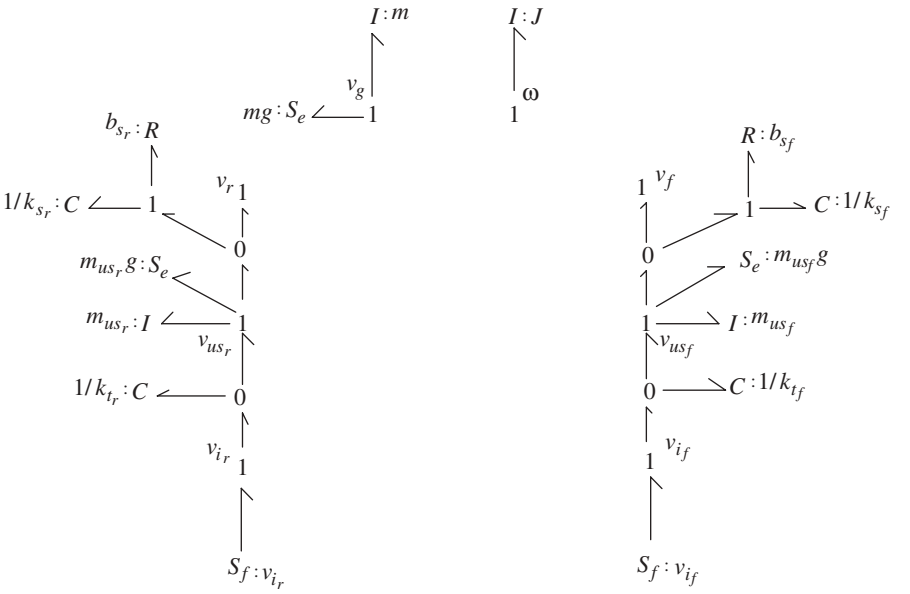


FIGURE 4.18. (*Continued*)

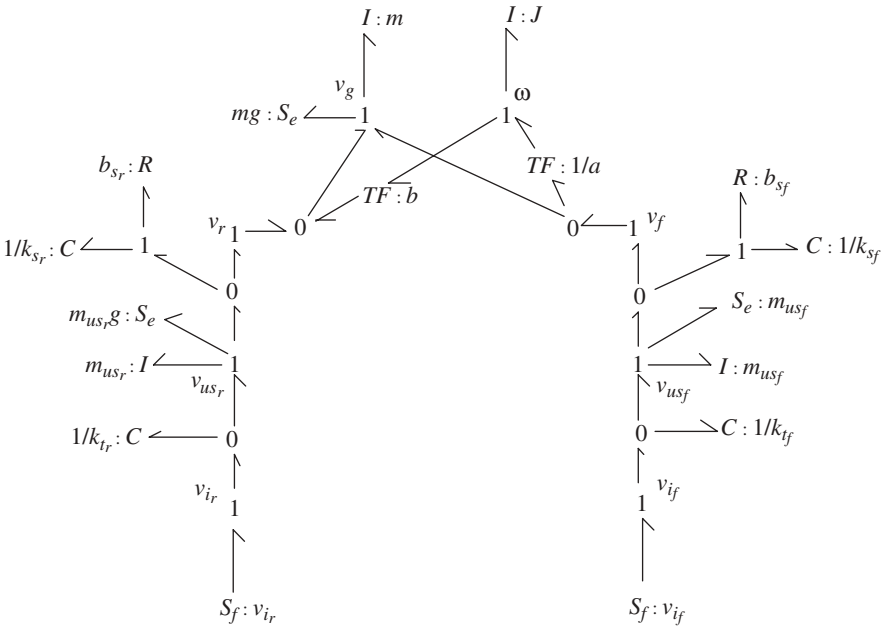


FIGURE 4.18. (*Continued*)

Also notice that for the suspension elements, the relative velocities were constructed only once, and then 1-junctions were used to enforce that the respective springs and dampers have the same relative velocity as in Figure 4.12. From Figure 4.18*c* it is obvious that the model is not complete, and yet all elements have been incorporated. This is where the kinematics plays a major role.

For small angular motions, using Eq. (4.6) and using the right-hand-rule for cross products, we can derive that

$$v_f = v_g + a\omega \quad \text{and} \quad v_r = v_g - b\omega. \tag{4.8}$$

(These relations may seem almost obvious but they are special cases of the vertical velocity relation in Eq. (4.7*a*) when remains near 0 for v_f and π for v_r.)

These *kinematic constraints* must be enforced to complete the model. Until now, it has been recommended that positive directions be shown on a schematic and then transferred to the bond graph model. If this is not done then the model may still be all right. But in a simulation of the system response, we might not know whether a positive velocity was up or down, or whether a positive force was compressive or extensive. This is important information to have, so it is always a good idea to have a power convention in mind. However, if the kinematic constraints are not enforced correctly, then the model is just plain wrong.

Kinematic constraints are always relationships among flow variables. To add the flows according to the constraints we use 0-junctions. We just need to pay particular attention to the power convention when adding up the flows to enforce the constraints. In Figure 4.18*d* the constraints from Eqs. (4.8) are enforced with two different half-arrow patterns using 0-junctions. The reader should check that in both cases the velocity components add correctly to produce Eqs. (4.8). Transformers are used to convert the angular velocity into velocity components, that is, ω into $a\omega$ or $b\omega$. On the transformers in the figure, the moduli are appended according to the definition of the transformer given in Chapter 3. Although either version in Figure 4.18*d* can be used, if the right-hand version of part (d) is used, as is done in Figure 4.18*e*, then a bond graph simplification can be done at the 1-junctions for v_f and v_r.

A second example of plane motion is the system shown in Figure 4.19*a*. The system is a cart of mass, m, acted upon by an input force, $F(t)$. Attached to the cart through a spring and damper, k and b, is a cylinder that rolls without slip on the cart. The cylinder is a rigid body of mass, m_c, with center of mass in

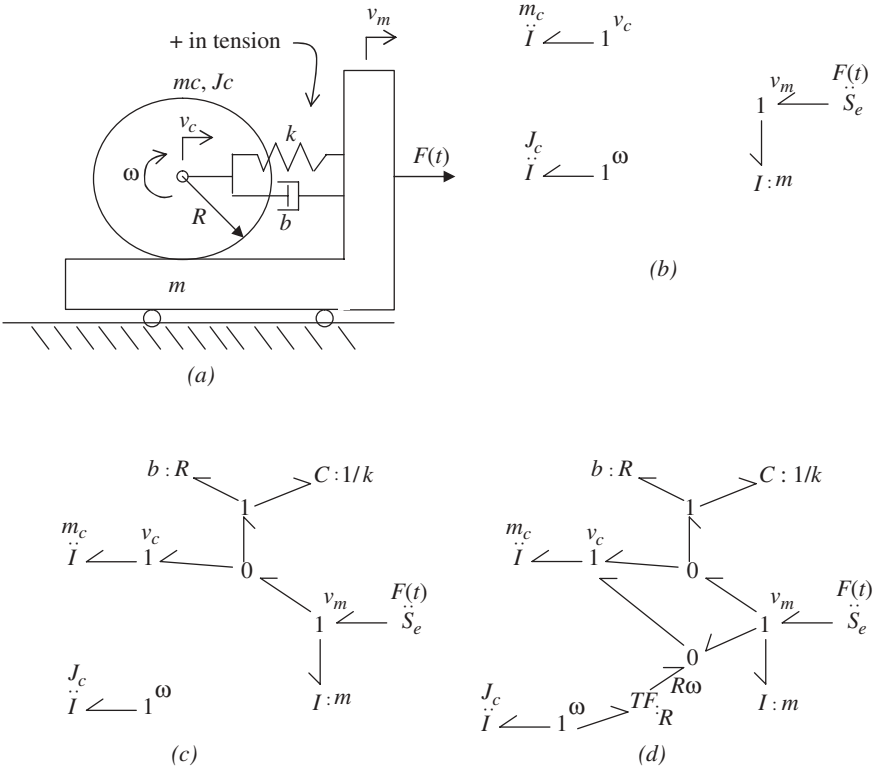


FIGURE 4.19. Cart with rolling cylinder. Example 2 of plane motion.

the center and moment of inertia about its center of mass, J_c, and is of radius R. We desire a model that, if solved analytically or simulated, will predict the motion–time history of this system for any prescribed input force.

The schematic shows positive velocity and angular velocity directions, and the spring and damper are assumed positive in tension. Figure 4.19*b* uses 1-junctions to represent the distinct velocities and angular velocities, and the elements that have these distinct flows have been attached to the 1-junctions. The rigid-body cylinder is a body in plane motion and is represented by its center-of-mass velocity, v_c, and its angular velocity, ω, and thus has the —I elements for its mass and moment of inertia appropriately attached. In Figure 4.19*c* the relative velocity is established across the spring and damper. Perusal of this figure clearly indicates that the model is not complete, even though all energy has been accounted for.

Again, there is a kinematic relationship that must be derived before we can complete the model. Application of Eq. (4.6) to the velocity of the point at the bottom of the cylinder, v_p, positive to the right, yields

$$v_P = v_c - R\omega. \tag{4.9}$$

Since the cylinder rolls without slip, the velocity at the bottom of the cylinder must be the same as the velocity of the cart at the same point. Thus, the kinematic constraint is

$$v_m = v_c - R\omega. \tag{4.10}$$

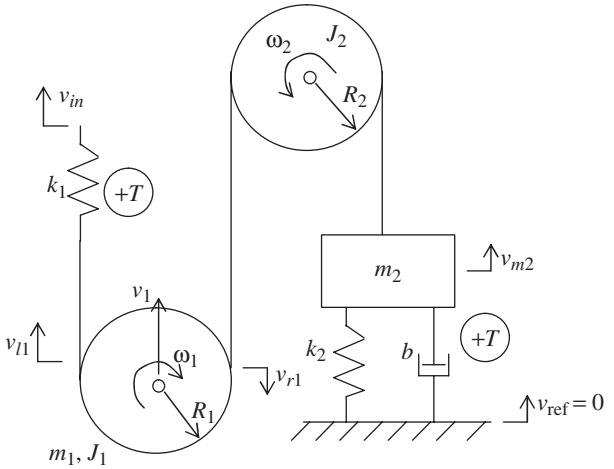
In Figure 4.19*d* this constraint is enforced using a 0-junction and the model is complete.

As a final example of systems with plane motion dynamics, consider the system in Figure 4.20*a*. An input velocity, $v_{\text{in}}(t)$, is prescribed on one end of a spring, k_1, and the other end is attached to an inextensible rope that wraps around a floating cylinder of mass m_1, radius R_1, and moment of inertia J_1. The rope then passes over a second cylinder that is pinned at its center and has radius R_2 and moment of inertia J_2. The rope finally is attached to a mass, m_2, which is attached to ground through the spring and damper, k_2 and b. The idea is to formulate a model that, if simulated, would predict the motion–time history of the system to a prescribed input velocity. Positive velocity and angular velocity directions are shown in the figure, and all springs and dampers are assumed positive in extension.

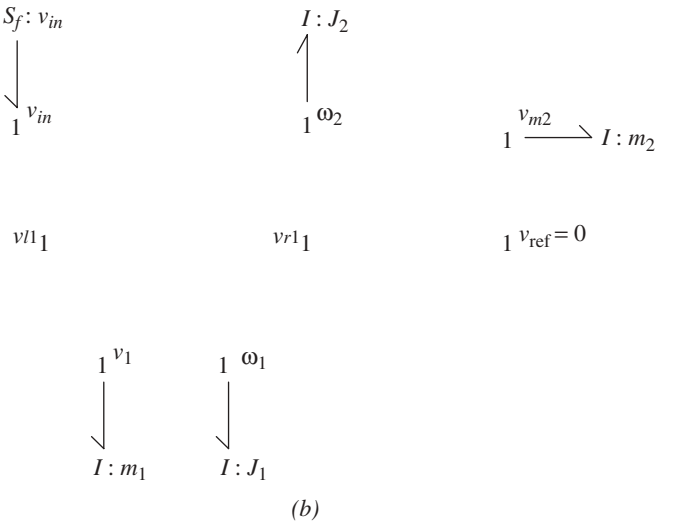
In Figure 4.20*b*, 1-junctions are used to represent the distinct velocities and angular velocities. The floating cylinder requires a 1-junction for its center-of-mass vertical velocity and a 1-junction for its angular velocity. Notice that the velocities at the left and right of the floating cylinder have been defined. The one at the left, v_{l_1}, has been introduced as a convenience for establishing the relative velocity across the spring with constant k_1. The velocity at the right of the cylinder, v_{r_1}, will be useful for relating to the angular velocity, ω_2, of the pinned cylinder. Also, we can see that, since the rope is inextensible, v_{r_1} is equal to the mass velocity, v_{m_2}. Attached to these 1-junctions are the bond graph elements that

have these distinct flows. The power arrow for the flow source, —S_f, is directed out of the source and into the system, because if the spring with constant k_1 is defined as positive in tension and if the top end of the spring is moving upward (defined as positive), then power would be flowing *into* the system.

In Figure 4.20*c*, 0-junctions are used to establish the proper relative velocities across the springs and dampers so as to put these elements positive in tension. The spring and damper attached to m_2 have been modeled using the "reducible

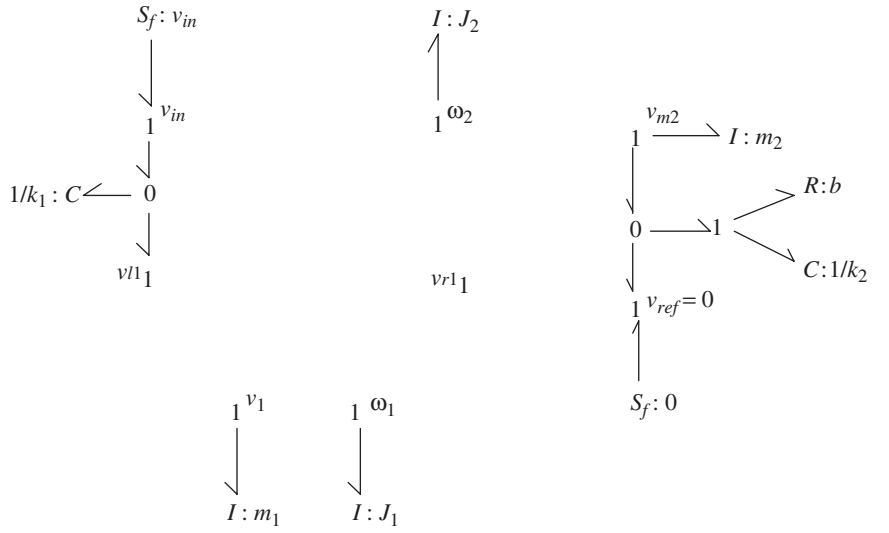


(a)

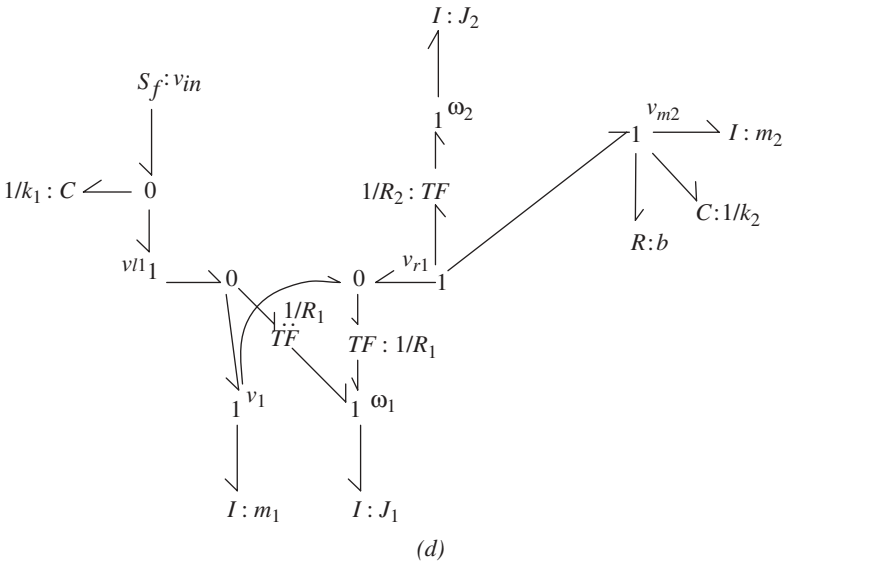


(b)

FIGURE 4.20. System with a floating pulley. Example 3 of plane motion.



(c)



(d)

FIGURE 4.20. (*Continued*)

loop" simplification of Figure 4.12. The reference velocity, $v_{\text{ref}} = 0$, can be erased in the final bond graph. Clearly, the model in Figure 4.20*c* is not complete.

For cylinder 1, application of Eq. (4.6) yields the kinematic constraints

$$v_{l_1} = v_1 + R_1\omega_1 \quad \text{and} \quad v_{r_1} = -v_1 + R_1\omega_1. \tag{4.11}$$

Moreover, since the rope is inextensible, the velocity at the rim of cylinder 2 is the same as v_{r_1}, thus,

$$v_{r_1} = R_2\omega_2. \tag{4.12}$$

Equations (4.11) are enforced using 0-junctions as shown in Figure 4.20*d*. The velocity v_{r_1} is related to ω_2, as required by Eq. (4.12). And v_{r_1} is set equal to v_{m_2} by simply connecting the 1-junctions with a bond. The model in Figure 4.20*d* is now complete. Some simplifications are possible, but this is left to the reader.

As has been demonstrated in all examples of mechanical systems with elements in plane motion, bond graph models are constructed by first establishing distinct velocities and angular velocities using 1-junctions that assign the same flow to each attached bond and add the efforts according to the power convention. This is followed by establishing the proper relative motions across compliances and resistances, using 0-junctions that add flows according to the power convention and assign the same effort to each attached bond. This part of the modeling procedure is straightforward and invariant. With some practice these steps will become automatic. The most difficult part of model construction is deriving and enforcing the kinematic constraints. There is no easy set of rules for recognizing and deriving these relationships. The constraints will always involve flow variables and will always involve application of the vector relationship, Eq. (4.6). Practice is the only way to become comfortable with this aspect of modeling.

A major advantage to the use of bond graphs is that if the velocities are properly constrained using power-conservative elements, then the force and torque relationships will automatically be correctly represented.

Body-fixed Coordinates. The concepts of plane motion were developed using the general motion of a rigid body, shown in Figure 4.17. In order to characterize all the energy of a rigid body in plane motion it is necessary to know the velocity of the center of mass and the angular velocity of the body. In Figure 4.17, the velocity vector of the center of mass was resolved into two mutually perpendicular components aligned in the inertial XY-directions. When there are several places where a rigid body interacts with other parts of a system and when the rigid body can execute large angular motions, it is convenient to introduce the concept of *body-fixed coordinates*. This is a coordinate frame that remains attached to the body at the center of mass and moves with the body as it translates and rotates under the action of whatever is attached to it. The virtue of such a frame is that the attachment points remain at fixed locations relative to this frame, and the inertial properties, that is, moments and products of inertia, remain invariant with respect to this frame. These properties are very important for three-dimensional rigid-body motion, which is discussed fully in Chapter 9. We introduce the concept here for plane motion.

The general rigid body of Figure 4.17 is shown again in Figure 4.21*a* with body-fixed coordinates. The xy-coordinate frame is attached to the body at the

*This section can be omitted without any loss of generality.

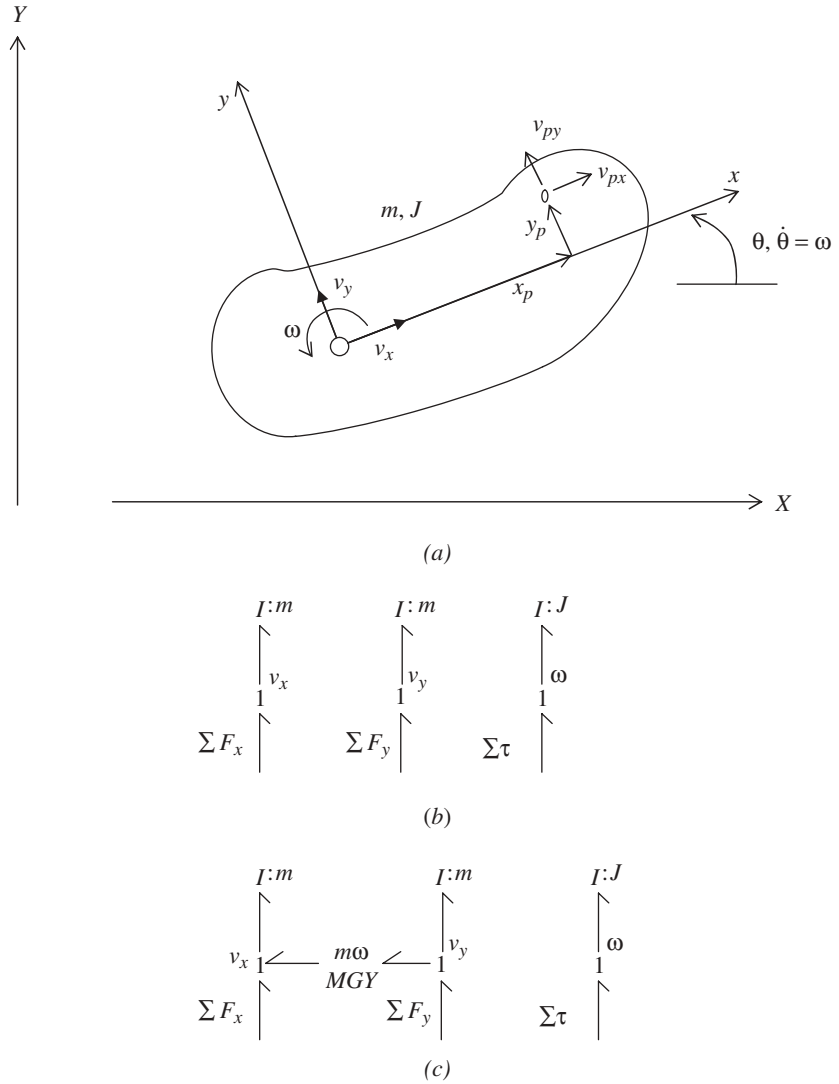


FIGURE 4.21. General plane motion with body-fixed coordinates.

center of mass, and it is aligned in principal directions, although this is not essential. The instantaneous velocity vector that is pointing in some direction in the plane of motion is resolved into two mutually perpendicular directions aligned along the body-fixed axes. These components, v_x and v_y, are different from the inertial components in Figure 4.17, since they change direction as the body rotates. This change in direction of velocity vector components results in acceleration components that must be accounted for. Also shown in Figure 4.21*a*

is the attachment point, P, that is located with respect to the xy-coordinates by the fixed lengths x_p, y_p.

In Reference [2], it is shown that for any vector, say the velocity vector, $\vec{v}$, referred to a rotating frame, the absolute time rate of change of the vector, or acceleration in this case, is

$$\frac{d}{dt}\vec{v} = \frac{\partial}{\partial t}\vec{v}|_{\text{rel}} + \vec{\omega}x\vec{v}, \tag{4.13}$$

where the first term on the right is the acceleration as observed relative to the moving frame, and the second term accounts for the contribution of the frame rotation to the absolute acceleration. Applying this relationship to the rigid body of Figure 4.21*a* yields the center of mass absolute acceleration, a_x, a_y, as

$$a_x = \dot{v}_x - \omega v_y \quad \text{and} \quad a_y = \dot{v}_y + \omega v_x, \tag{4.14}$$

where the components due to rotation are derived by carrying out the cross products using the right-hand-rule. For example, the ω-vector pointing out of the page, when crossed into the x-direction velocity component, v_x, yields a vector of length ωv_x pointing in the y-direction.

Newton's laws apply to the absolute acceleration of the center of mass, and become

$$\sum F_x + m\omega v_y = m\dot{v}_x \quad \text{and} \quad \sum F_y - m\omega v_x = m\dot{v}_y, \tag{4.15}$$

where the sum of forces in the x- and y-directions come from any attached systems to the body, and the rotational contributions to the acceleration have been brought to the force side of the equation after multiplying by the body mass, m.

Returning now to Figure 4.21*b*, it is tempting to represent the plane motion of the rigid body as shown, where the body-fixed velocity components, v_x, v_y, are used rather than the inertial components, v_X, v_Y from Figure 4.17. The kinetic energy of the body is accounted for, but the absolute acceleration is not. In other words, if the forces from external elements were added using the 1-junctions of Figure 4.21*b*, then the resulting equations of motion would not be correct. The fix to this problem lies in Eqs. (4.15).

The cross-product terms of Eqs. (4.15) have a remarkable symmetry. In the first equation, there is a force component in the x-direction that is equal to $m\omega$ times the velocity in the y-direction. And in the second equation there is a force component in the y-direction equal to $m\omega$ times the velocity in the x-direction. Since gyrators relate efforts to flows across them, consider Figure 4.21*c* where a modulated gyrator has been inserted between the 1-junctions representing v_x and v_y. A modulated gyrator is used because the modulus, $m\omega$, is not constant but varies as the body moves and changes. Convince yourself that the forces on the —I elements exactly duplicate the left-hand side of Eqs. (4.15). When body-fixed coordinates are used, the —MGY— must be part of the bond graph.

For the attachment point in Figure 4.21*a*, the velocity components in the body-fixed directions are related to the center of mass velocity components using Eq. (4.6), with the result

$$v_{p_x} = v_x - y_p\omega \quad \text{and} \quad v_{p_y} = v_y + x_p\omega. \tag{4.16}$$

As an example of the use of body-fixed coordinates, consider the system in Figure 4.22*a*. The rigid body of mass, m, and centroidal moment of inertia, J, is attached to ground at points 1 and 2 through springs and dampers as shown. The ground side ends of the springs and dampers are on frictionless carts that ensure that horizontal elements remain horizontal regardless of the motion of the body. Horizontal spring 1 has a velocity input at its ground side end. The body is shown displaced from its starting orientation, and body-fixed coordinates are attached at the center of mass. The attachment points are located with respect to the body-fixed coordinate frame by the fixed distances x_1, y_1 for point 1 and x_2, y_2 for point 2, where x_2 is a negative quantity. In Figure 4.22*b* the body-fixed velocity components at the attachment points are indicated, and in Figure 4.22*c*

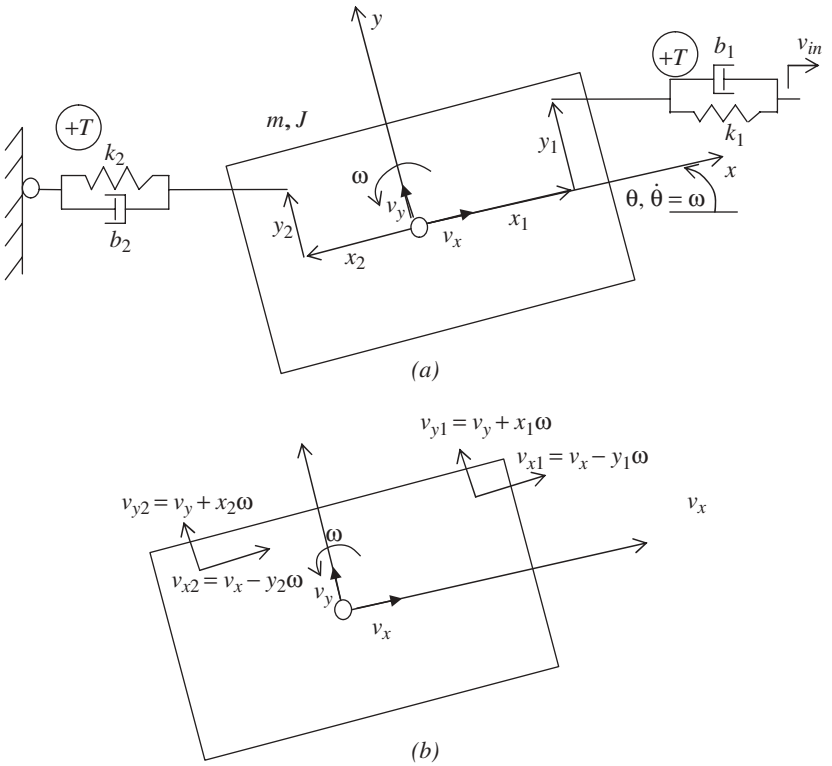


FIGURE 4.22. A rigid body attached to a frame undergoing large motions.

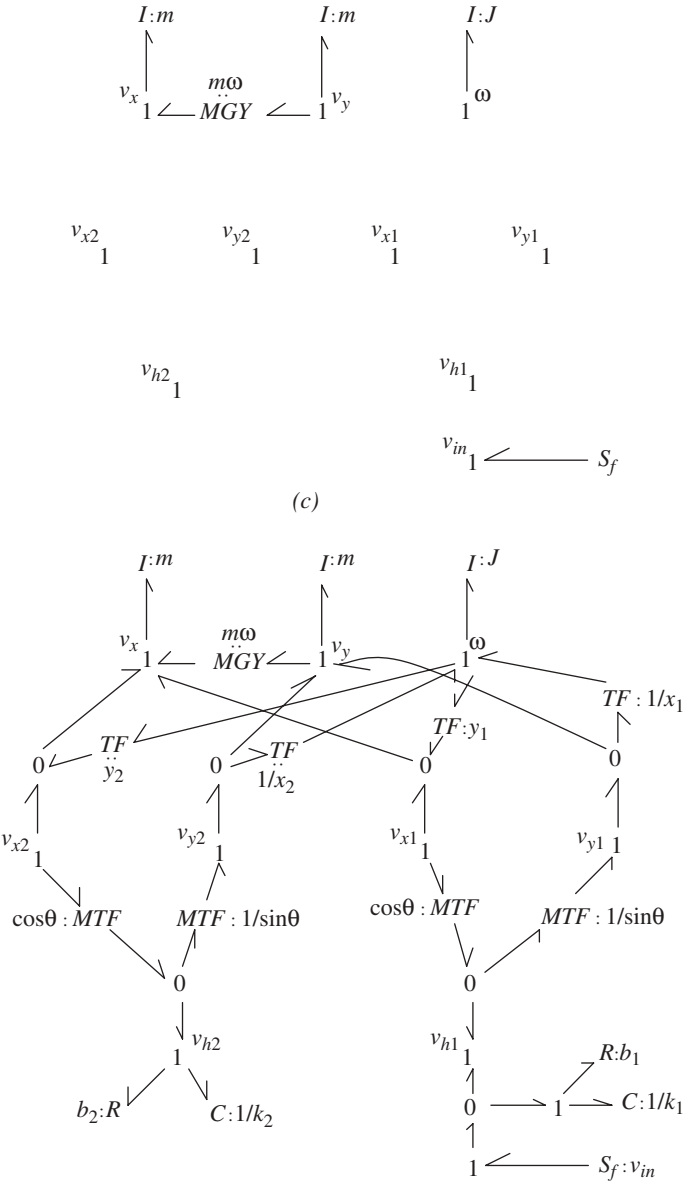


FIGURE 4.22. (*Continued*)

1-junctions are used to represent all distinct velocities. Attached to these 1-junctions are elements that have these absolute velocities. Notice the representation of the body when using body-fixed coordinates. The body-fixed velocity components at the attachment points are indicated using 1-junctions. These are defined as a convenience for ultimately determining the horizontal velocity at the attachment points for use as inputs to the horizontal spring–damper elements. As noted in Figure 4.22*b*, the body-fixed velocities at the attachments are

$$\begin{aligned} v_{x_1} &= v_x - y_1\omega, \\ v_{y_1} &= v_y - x_1\omega, \\ v_{x_2} &= v_x - y_2\omega, \\ v_{y_2} &= v_y + x_2\omega, \end{aligned} \tag{4.17}$$

and the horizontal velocity at the attachments are

$$\begin{aligned} v_{h_1} &= v_{x_1}\cos\theta - v_{y_1}\sin\theta, \\ v_{h_2} &= v_{x_2}\cos\theta - v_{y_2}\sin\theta. \end{aligned} \tag{4.18}$$

These kinematic relationships are enforced using 0-junctions as shown in Figure 4.22*d*. The transformers used to enforce constraints from Eqs. (4.18) are modulated transformers, —*MTF*—, because the moduli of these elements vary as the body moves. Figure 4.22*d* is the final bond graph for this example. The bond graph does appear a bit busy, but it contains a lot of information. In Chapter 9, some shorthand notation is introduced for complex systems that allows the bond graph to be constructed while showing fewer bonds.

4.3 HYDRAULIC AND ACOUSTIC CIRCUITS

In this section, special but important classes of fluid-flow systems will be modeled using bond graphs. The models to be used exhibit a close analogy to the mechanical and electrical systems studied in the previous section of this chapter. The variables to be used here have been discussed in Chapter 2 (Table 2.4), and the basic fluid system elements have been discussed in Chapter 3. Tables 3.1, 3.2, 3.3, and 3.4 and Figures 3.1, 3.2, and 3.4 show the type of 1-port fluid elements to be used in this section. Figure 3.8 shows a 2-port transformer with a fluid port, and Figure 3.11 shows 0- and 1-junction fluid elements. So far the discussion has been fairly general, but it is now time to be more specific about the modeling of fluid systems.

One category of systems that is important in engineering and can be modeled using the elements shown previously is commonly called *hydrostatic*. These are systems composed of pumps, motors, pipes, pistons, valves, filters, and accumulators that use nearly incompressible fluids, such as water or hydraulic oil. Such systems are found in machine tools, earth-moving equipment, power

transmissions, and aircraft control surface servomechanisms. Generally, these systems have high pressures and low fluid velocities so that the static pressure dominates the dynamic pressure. This is the origin of the name *hydrostatic*, even though the *dynamics* of such systems are typically of great interest.

A second category of fluid systems that can be treated with the elements to be used here involves compressible gases (such as air) but in which the pressure deviations are small enough that the *acoustic approximation* applies. Acoustic circuits can often be used to design some types of mufflers to reduce noise in air-conditioning systems, for example. However, pneumatic systems in which large pressure variations are expected need a more complicated modeling treatment since they are general thermofluid systems, as discussed in Chapter 12.

Most engineers are familiar with lumped parameter elements of mechanical and electrical systems such as rigid bodies, springs, dampers, resistors, capacitors, and inductors. The equivalent fluid elements are generally less familiar and less obvious. Thus, it is worthwhile to describe some 1-port elements that can be used to model hydraulic and acoustic circuits before showing how to assemble bond graph models of these systems.

4.3.1 Fluid Resistance

Figure 3.1 showed a 1-port fluid resistor that relates an effort, a pressure drop, P, to a flow, a volume flow rate, or volume velocity, Q. Figure 4.23 shows how resistors typically appear in a bond graph, together with example devices represented by a resistor. Part *a* of Figure 4.23 is a combination of a resistor and a 1-junction. This combination implies that the pressure drop, P_3, is related to the pressures P_1 and P_2 by the equation

$$P_3 = P_1 - P_2, \tag{4.19}$$

and that all volume flow rates are equal,

$$Q_1 = Q_2 = Q_3. \tag{4.20}$$

Following Table 3.1, the resistor implies that the pressure drop is related to the volume flow rate by a nonlinear function,

$$P_3 = \Phi_R(Q_3), \tag{4.21}$$

or, if a linear relation is assumed, by a resistance coefficient R_3

$$P_3 = R_3 Q_3. \tag{4.22}$$

It is an unfortunate fact that in many hydraulic and acoustic systems it is not easy to predict resistance functions or coefficients for parts of the system before it is constructed. However, experimentally it not difficult to measure pressure drops

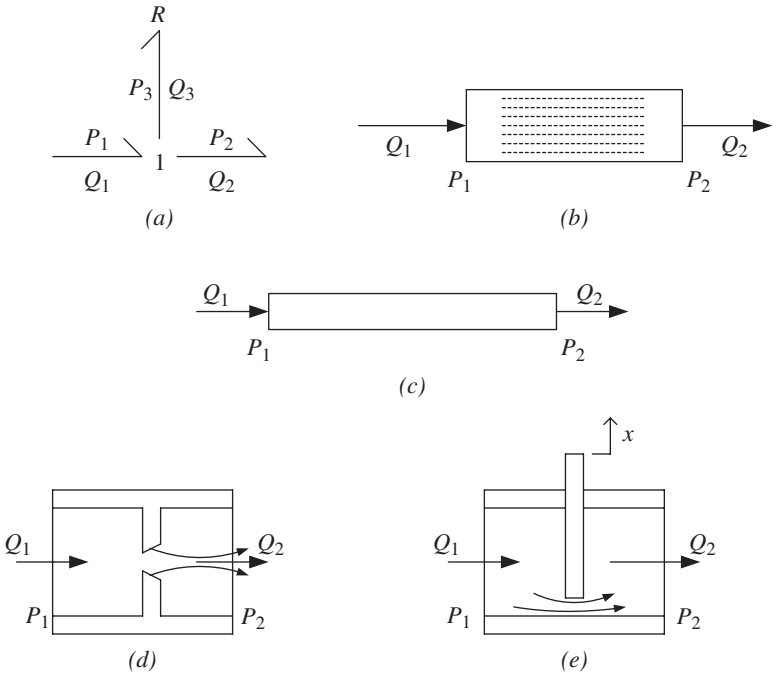


FIGURE 4.23. Fluid resistors: (*a*) bond graph; (*b*) porous plug; (*c*) long pipe; (*d*) orifice; (*e*) valve with variable area, $A(x)$.

and flow rates under steady conditions and to characterize resistance effects. This means that there are at least guidelines in the engineering literature for suitable resistance law assumptions.

Part *b* of Figure 4.23 is intended to represent a porous plug in a pipe for which viscous forces on the fluid from the plug might be assumed to dominate. In this case it would be logical to use Eq. (4.22). In the absence of experimental data relating to the plug, one would have to experiment with values of the coefficient R_3 in studying a system model.

Another case in which a linear resistance can be assumed is shown in part *c* of Figure 4.23, which is supposed to represent a long, thin tube in which laminar flow of an incompressible fluid develops. In this case a theoretical value for the resistance can be given, (see Reference [4], Section 7.4):

$$R_3 = 128\mu l/\pi d^4, \tag{4.23}$$

where μ[Pa · s] is the fluid viscosity coefficient, l [m] is the length and d [m] is the inside diameter. (Note that the Appendix lists typical property values, such as viscosity, for a number of materials. These are particularly useful for estimating 1-port element parameters and functions for the types of systems treated in this section.)

For incompressible flow in long pipes, it is useful to compute a Reynolds number, Re, defined as

$$\mathrm{Re} = 4\rho Q/\pi d\mu, \tag{4.24}$$

where ρ [kg/m^3] is the fluid mass density. When the Reynolds number is low, say about 200 or less, viscous forces predominate and Eqs. (4.22) and (4.23) can be used. At a higher Reynolds number, the flow becomes turbulent and the pressure–flow relationship under steady-flow conditions is nonlinear. The transition to turbulent flow depends on the pipe dimensions l and d, the pipe surface roughness, and fluid properties.

For Reynolds numbers greater than about 5000, the flow is likely to be turbulent, and a general function for the nonlinear relation from Eq. (4.21) is

$$P_3 = a_t Q_3 |Q_3|^{3/4}. \tag{4.25}$$

In this formula, the absolute value sign is necessary to make sure that P_3 is negative if Q_3 is negative. The constant a_t is often determined experimentally (see Reference [5]). Even when Eq. (4.25) is known to be valid for steady flow, its use in studying transient conditions is not necessarily valid. For oscillatory flows or other dynamic conditions, turbulence may not develop fully, so in these conditions, Eq. (4.25) should be regarded only as an approximation.

Parts *d* and *e* of Figure 4.23 show two important cases in which a pressure drop occurs over short lengths. The orifice is assumed to have a fixed area A_0, while the valve has a variable area $A(x)$, where x is a position coordinate. Although the valve is shown as if it were a gate valve, the basic equation for the pressure drop applies to a variety of other valve configurations. It is a standard exercise in fluid mechanics texts to derive the laws of an orifice using considerations of energy, momentum, and continuity. The main result is that the pressure drop is proportional to the square of the volume flow rate. One form of the law is (see Reference [6], Section 3.8)

$$P_3 = (\rho Q_3 |Q_3|)/2C_d^2 A_0^2, \tag{4.26}$$

where again the use of the absolute value sign corrects the sign of the pressure drop P_3 if Q_3 happens to be negative and C_d is a discharge coefficient. For a round, sharp-edged orifice, C_d can be predicted to have a value of 0.62, but for other shapes of holes, the values of C_d vary somewhat.

For valves, the area depends on a position coordinate and the discharge coefficient may also vary, so the relation equivalent to Eq. (4.26) can be written

$$P_3 = (\rho Q_3 |Q_3|)/2C_d^2(x) A^2(x). \tag{4.27}$$

In case it turns out that a causal analysis of the system model ultimately requires that the flow must be expressed in terms of pressure drop, the

relationships of Eqs. (4.26) and (4.27) must be inverted. For example, the inverse version of Eq. (4.26) is

$$Q_3 = C_d A_0 (2|P_3|/\rho)^{1/2} \operatorname{sgn} P_3. \tag{4.28}$$

This form is slightly more complicated than the form usually given to allow the pressure drop and volume flow rate to take on both positive and negative values. (The *signum* function "sgn" is just $+1$ if P_3 is positive and -1 if it is negative.) In dynamic systems in which oscillating flows may exist, one must make sure that both plus and minus values of the variables are computed correctly.

Equation (4.28) is the inverse form of a nonlinear relation as indicated by Eq. (4.21). When a linear relation can be assumed as shown in Eq. (4.22), the inverse relation is much simpler:

$$Q_3 = P_3/R_3. \tag{4.29}$$

Chapter 5 will present methods for deciding which causal form of resistance laws are required for formulating the state equations of a system to be analyzed or simulated.

4.3.2 Fluid Capacitance

As shown in Table 3.2, a fluid capacitor imposes a relationship between a pressure (an effort variable) and the integral of the volume flow rate (a displacement variable). This integrated variable is a volume of fluid in the cases at hand. Figure 4.24 shows the common case in which capacitors are combined with common effort junctions (0-junctions). The nonlinear form of the capacitance constitutive relation for the bond graph of Figure 4.24*a* is

$$P_3 = \Phi_C^{-1}(V_3), \tag{4.30}$$

where

$$V_3 = \int_0^t Q_3 dt. \tag{4.31}$$

If the capacitor can be considered to be linear, a capacitance parameter, C_3, can be defined and the capacitor law of Eq. (4.30) simplifies to

$$P_3 = V_3/C_3. \tag{4.32}$$

The 0-junction laws are

$$P_1 = P_2 = P_3 \quad \text{and} \quad Q_3 = Q_1 - Q_2, \tag{4.33}$$

so V_3 in this case is the integral of the difference in the volume flow rates Q_1 and Q_2.

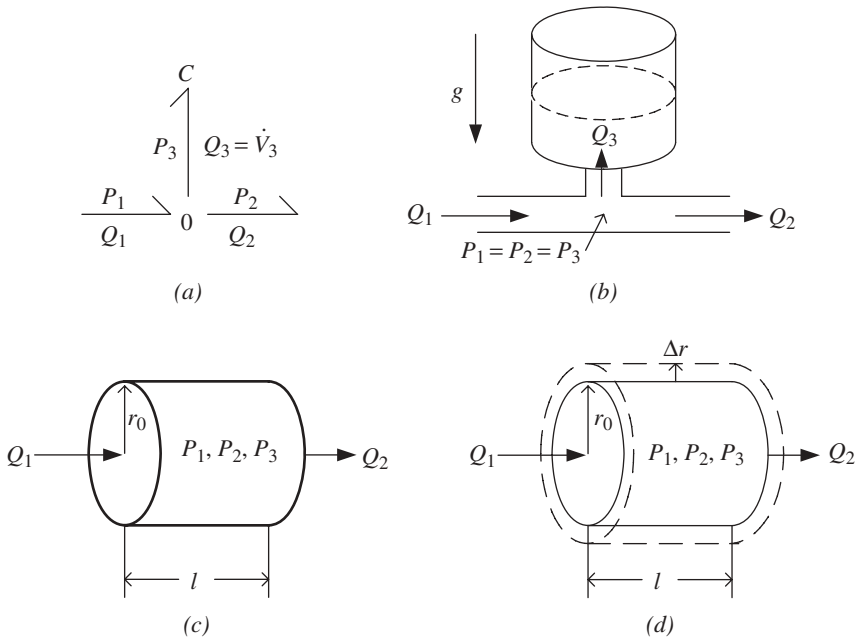


FIGURE 4.24. Fluid capacitors: (*a*) bond graph; (*b*) water tank; (*c*) rigid pipe segment; (*d*) elastic pipe segment.

The first example, shown in Figure 4.24*b*, is a straight-walled water tank in a gravity field. In this case V_3 is the volume of water that has been stored in the tank. If h represents the height of the water above the horizontal connecting pipes, then the pressure at the bottom of the tank is just the hydrostatic value of ρgh. (Strictly speaking, this is the *gage pressure*, counting atmospheric pressure to be zero. We also assume that the bottom of the tank is essentially at the location of the EET junction connecting the pipes. The little stub shown entering the tank bottom is supposed to be very short.).

If the tank area is A_3, the volume of water in the tank, V_3, is A_3h, and the pressure at the bottom of the tank is therefore

$$P_3 = \rho g V_3 / A_3. \tag{4.34}$$

Comparing Eqs. (4.34 and 4.32) we see that the capacitance of the tank is

$$C_3 = A_3 / \rho g. \tag{4.35}$$

A gravity tank with straight sides can be represented as a linear fluid capacitor. The next example shown in Figure 4.24*c* is a segment of a rigid pipe that has a physical volume of

$$V_0 = \pi r^2 l. \tag{4.36}$$

For a truly incompressible fluid, the volume flow rate into the pipe, Q_1, and the flow rate coming out of the pipe, Q_2, would have to be equal, and thus the flow Q_3, according to Eq. (4.33), would be zero. On the other hand, if Q_1 and Q_2 were not exactly equal, Q_3 would represent the rate of compression of the fluid in the pipe. More important, V_3 according to Eq. (4.31), would represent the *decrease* in volume of the fluid in the pipe segment.

For liquids such as water or hydraulic oil that are only slightly compressible, a *bulk modulus* B is defined by the relation (see Reference [6])

$$dP = -B(dV|V). \tag{4.37}$$

The minus sign in Eq. (4.37) is due to the fact that when the volume of the fluid increases incrementally by an amount dV, the pressure decreases by the amount dP. For the capacitance model we will extend the relation of Eq. (4.37) to small but finite changes and note that ΔV_3 represents a small *decrease* in the volume of fluid in the pipe segment, so the minus sign in Eq. (4.37) disappears and the capacitance law for the pipe segment is

$$\Delta P_3 = (B/V_0)\Delta V_3, \tag{4.38}$$

where V_0 comes from Eq. (4.36) and the Δ may be eliminated if it is remembered that P_3 and V_3 in the bond graph of Figure 4.24*a* represent deviations from steady-state values in this case.

Comparing Eq. (4.32) with Eq. (4.38), we see that the capacitance for the rigid pipe segment is

$$C_3 = V_0/B. \tag{4.39}$$

The Appendix gives values for the bulk modulus for some common liquids.

If the fluid is a gas and if the pressure and volume variations are small enough that the so-called *acoustic approximation* is valid, another relation can be given relating pressure change to volume change (Reference [7]):

$$\Delta P_3 = \rho_0 c^2 V_3/V_0, \tag{4.40}$$

where ρ_0 is the density of the gas at the reference pressure and c is the speed of sound. This means that the capacitance for this case is

$$C_3 = V_0/\rho_0 c^2. \tag{4.41}$$

Densities and speeds of sound for common gases can be found in the Appendix.

Figure 4.24*d* shows a case in which the capacitance in a fluid system is due to the elasticity of the vessel containing the fluid, rather than the compressibility of the fluid itself. The pipe segment has thin walls and is made of an elastic material. (It might represent a nonmetallic hydraulic brake line, for example.)

An approximate analysis will be made in which ΔV_3 represents the *increase* in volume of the pipe segment due to an increase in pressure with the fluid being assumed incompressible. The stress and strain effects in the longitudinal direction will be neglected, and only the volume change due to the hoop stress in the pipe will be computed.

An elementary analysis yields the hoop stress in the pipe:

$$\sigma = r_0 P / t_w, \tag{4.42}$$

where r_0 is the nominal pipe radius and t_w is the (small) wall thickness. The circumferential strain is then

$$\varepsilon = \sigma / E = r_0 P / t_w E = \Delta(2\pi r)/2\pi r_0 = \Delta r / r_0, \tag{4.43}$$

where E is the elastic modulus (see the Appendix). The change in volume is

$$\Delta V = \Delta(\pi r^2 l) = \pi l 2 r_0 \Delta r. \tag{4.44}$$

Combining Eqs. (4.43) and (4.44), yields

$$\Delta V = (2\pi l r_0^3 / t_w E)\Delta P = (2 r_0 V_0 / t_w E)\Delta P. \tag{4.45}$$

In terms relating to the bond graph of Figure 4.24*a* the result is

$$\Delta P_3 = (t_w E / 2 r_0 V_0)\Delta V_3.$$

This means that the capacitance due to pipe elasticity is

$$C_3 = 2 r_0 V_0 / t_w E. \tag{4.46}$$

Combining Capacitances. It sometimes happens that both fluid compressibility and pipe elasticity contribute to a compliance effect. One possibility is to use two separate bond graphs like Figure 4.24*a* in cascade with parameters for the two cases of Figure 4.24*c* and *d*, but another possibility is to construct a single equivalent compliance. Figure 4.25 shows another way to think about the two sources of compliance. Bond 3 in the first bond graph has a C-element dealing with fluid compressibility with capacitance from Eq. (4.39), and bond 4 has a compliance dealing with pipe wall elasticity and using a capacitance from Eq. (4.46). The 4-port 0-junction implies that all bonds have the identical pressure and that the algebraic sum of the flows (considering the sign half-arrows) add to zero. The bond graph implies that the difference between Q_1 and Q_2 is the sum of Q_3 and Q_4, that is, the difference between inlet and outlet flows is partly due to fluid compression and partly to pipe expansion. After integration, the volume variables are related as follows:

$$\Delta V_3 + \Delta V_4 = \int_0^t (Q_1 - Q_2)\, dt. \tag{4.47}$$

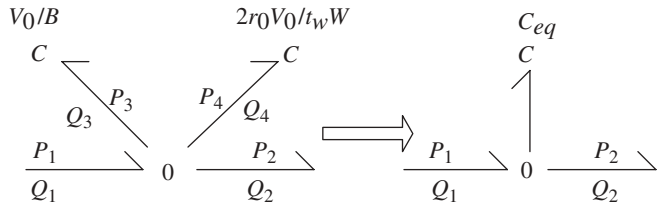


FIGURE 4.25. Single equivalent capacitance for liquid compressibility and pipe elasticity.

Using the capacitances from Eqs. (4.39) and (4.46) and calling the common pressure ΔP, Eq. (4.47) becomes

$$(V_0/B + 2r_0V_0/t_wE)\Delta P = \int_0^t (Q_1 - Q_2)\,dt, \tag{4.48}$$

which implies that capacitances add on a 0-junction. The equivalent single capacitance shown in Figure 4.25 is just

$$C_3 = V_0(1/B + 2r_0/t_wE). \tag{4.49}$$

Note that this formula enables one to decide whether fluid compressibility or pipe wall elasticity is the more important factor or whether both effects should be considered together. Both effects have to do with the volume of the pipe segment, but other physical parameters and dimensions are also involved.

Nonlinear Capacitances. Although much of the dynamic modeling of fluid systems can be usefully accomplished using linearized component laws, there are circumstances in which nonlinear models are essential. This is the case for some types of fluid resistance and for some fluid capacitance devices. Consider, for example, the compressed gas *accumulator* shown in Figure 4.26. It consists of a pressure vessel in which a flexible bladder contains a compressed gas. Such devices are often used to smooth out pressure and flow fluctuation in hydraulic systems or to provide short bursts of power under transient conditions.

The incompressible fluid flow, Q_3, compresses the gas in the bladder, and the volume of the compressed gas determines the pressure in the accumulator and at the 0-junction. Under dynamic conditions, the gas usually does not have time to exchange much heat with its surroundings. If this is the case, the isentropic pressure–volume law

$$PV^\gamma = P_0V_0^\gamma = \text{constant} \tag{4.50}$$

is a good approximation. In this equation, γ is the ratio of specific heats at constant pressure and volume (see the Appendix), P and V are the instantaneous pressure and volume of the gas, and P_0 and V_0 are values at some initial time.

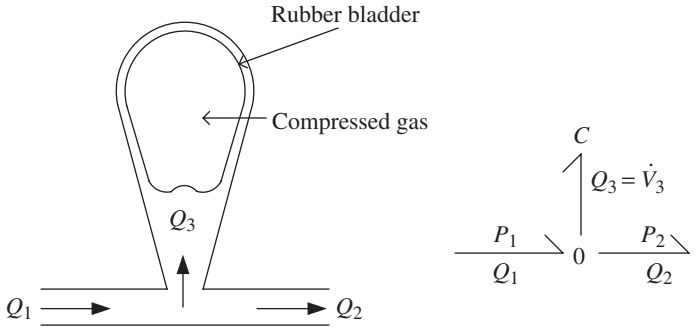


FIGURE 4.26. Compressed gas accumulator.

The bond graph of Figure 4.26 indicates that the gas volume is

$$V = V_0 - V_3 = V_0 - \int_0^t Q_3\, dt = \int_0^t (Q_1 - Q_2)\, dt. \tag{4.51}$$

This means the nonlinear capacitance relation between P_3 and V_3 is

$$P_3 = P_0 V_0^{\gamma} / (V_0 - V_3)^{\gamma}, \tag{4.52}$$

upon use of Eq. (4.50).

4.3.3 Fluid Inertia

The final passive 1-port element that needs to be discussed for fluid circuits concerns the inertia associated with fluid in a pipe segment. Table 3.3 shows that fluid systems have inertia effects analogous to those of mechanical and electrical systems, but for most people the variables involved, *volume flow rate* and *pressure momentum* are much less familiar than the corresponding mechanical or electrical variables. Also, as we will see, the inertia coefficient for a fluid circuit element is related to the parameters of a pipe segment in a counterintuitive way.

Figure 4.27 shows the simplest case of a straight pipe section of area A and length l filled with a fluid of density ρ. The bond graph shows the common situation in which the 1-port appears connected to a 1-junction. In this case, the volume flow rates at both ends of the pipe are equal and the pressure difference acting on the inertia element is

$$P_3 = P_1 - P_2. \tag{4.53}$$

The pressure momentum variable for the inertia element, as defined in Table 2.4, is

$$P_{P_3} \equiv \int^t P_3\, dt = \int^t (P_1 - P_2)\, dt. \tag{4.54}$$

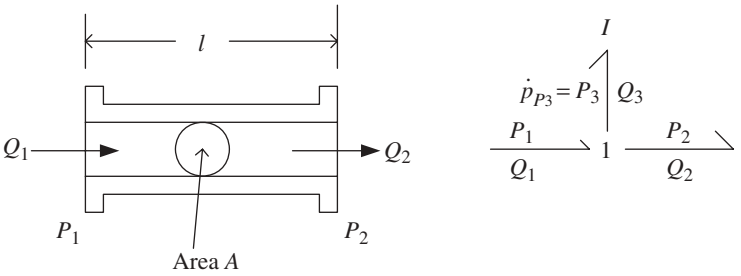


FIGURE 4.27. Fluid inertia for a pipe segment.

A linear inertia coefficient for this I-element is defined by a relation between the volume flow rate and the pressure momentum:

$$I_3 Q_3 = p_{P_3} \text{ or } Q_3 = p_{P_3}/I_3. \tag{4.55}$$

A simple derivation of the inertia coefficient for this case involves writing Newton's law for the fluid in the pipe. (A more thorough analysis is given in Section 12.4.1.) Assuming that the slug of fluid in the pipe moves as a rigid body, its velocity is Q_3/A, its acceleration is $\dot{Q}_3/A$ and its mass is ρAl. The net force accelerating the slug involves the pressure times the area at the two ends, $P_1A - P_2A$. Then Newton's law becomes

$$(\rho Al)\dot{Q}_3/A = (P_1 - P_2)A. \tag{4.56}$$

Using Eqs. (4.53), (4.54), and (4.55), the results are

$$(\rho l/A)\dot{Q}_3 = P_3 \tag{4.57}$$

or

$$(\rho l/A)Q_3 = p_{P_3}, \tag{4.58}$$

which means that the inertia coefficient is

$$I_3 = \rho l/A. \tag{4.59}$$

This coefficient for the inertia of fluid in a pipe segment is analogous to the mass for a body in translation or the inductance in an electric circuit.

It is no surprise that the inertia coefficient is proportional to the mass density and length of the fluid slug, but that it is inversely proportionate to the area seems counterintuitive. The reason is that we are using pressure and volume flow rate as effort and flow variables instead of force and velocity. Many engineers have assumed that narrow tubes would exhibit little inertia effect since they contain

little fluid mass, but, in fact, the narrower the tube, the more inertia the fluid in the tube will have when it is used in a fluid circuit. Oddly enough, fluid circuits with long, small-diameter pipes react sluggishly, as would mechanical systems with large masses.

As discussed in Section 12.4.1, the inertia coefficient for a pipe whose area varies with distance s along the pipe, $A(s)$, is given by the formula

$$I_3 = \int_0^t \rho ds / A(s), \tag{4.60}$$

which shows that it is the parts of the pipe where the area is the smallest that contribute the most to the inertia coefficient.

It should be noted that although the inertia formula, Eq. (4.59), applies to both hydraulic systems and acoustic systems, there is a subtlety connected with the inertia coefficient for an open-ended pipe in an acoustic circuit. If the pipe is unflanged, the length in Eq. (4.59) should be corrected by adding about 0.6 times the radius to the physical length at both ends. This takes into account the radiation impedance associated with the connection to the atmosphere (see Reference [7], Section 9.2).

4.3.4 Fluid Circuit Construction

Hydraulic and acoustic circuits have a lot in common with electrical circuits. Remember that for electrical circuits, it was easy to establish a 0-junction for each location in a circuit at which a voltage could be defined. The passive 1-port elements were then inserted between the appropriate 0-junctions using 1-junctions to establish the voltage differences acting on the elements. Finally, one of the 0-junctions was picked as representing a reference or ground voltage and was eliminated together with all the bonds emanating from it. For fluid circuits, pressures act like voltages and volume flow rates act like electrical currents, so a short version of the circuit construction procedure is as follows:

1. **Show positive directions for fluid flows and pressures on a schematic diagram.**
2. **For each distinct absolute pressure establish a 0-junction**.
   Attach elements if they happen to react to one of the absolute pressures. If necessary, include a 0-junction for atmospheric pressure.
3. **Insert *R*-, *C*-, and *I*-elements between appropriate 0-junctions using 1-junctions.**
   The 1-junctions, (with through power signs) establish pressure differences, which act on passive elements.
4. **The 0-junction for atmospheric pressure can be eliminated if "gage" pressures are to be used.**

5. **Simplified the graph.**
   If there are any 2-port 0- or 1-junctions with through-sign-convention half-arrows or branches that resulted in reducible loops, it will be possible to simplify the graph.

Figure 4.28*a* shows a hydraulic pump–motor unit with a pressure-relief-valve bypass. We shall assume the pump to be a flow source, the relief valve and motor both to be resistances, and the reservoir to be at atmospheric pressure. For now we shall not deal with the fact that both the pump and the motor are transducers relating hydraulic and mechanical powers. We also shall ignore the filter. In Figure 4.28*b* the two distinct pressures are identified. In part *c* the pump, valve, and motor are inserted. This is a correct bond graph model. Next we define the pressure at B with respect to atmospheric pressure and eliminate the P_A junction. Simplification yields the system bond graph of Figure 4.28*d*. The model now deals with "gage" pressures.

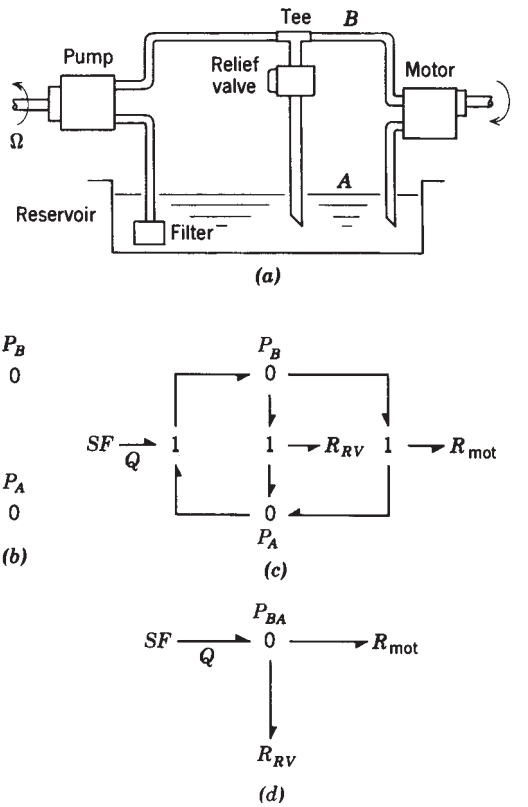


FIGURE 4.28. A hydraulic pump–motor unit.

Frequently, hydraulic elements must be represented that do not show their pressure points quite so distinctly as in the previous example. Consider a length of flexible line connecting a pump to a load. If the line is long, if the pressure varies rapidly or widely, or if great accuracy in prediction is required, it may be necessary to include the effects of fluid inertia in the line, the flexibility of the line walls, and the internal line resistance to flow. In fact, all these effects are distributed throughout the line. Fortunately, it is often possible to get good predictive models by making a few simple assumptions. A long, flexible line is shown in Figure 4.29*a*. In part *b* a simple dynamic model indicating inertia, resistance, and compliance is shown; notice the pressure points defined in the process. The inertia acts to generate a flow (Q_{12}) in dynamic response to a pressure difference ($P_1 - P_2$). The resistance also generates a flow (Q_{23}) in response to a pressure difference ($P_2 - P_3$). Wall compliance generates a pressure (P_3) (relative to atmosphere) in dynamic response to a net flow. These pressures are

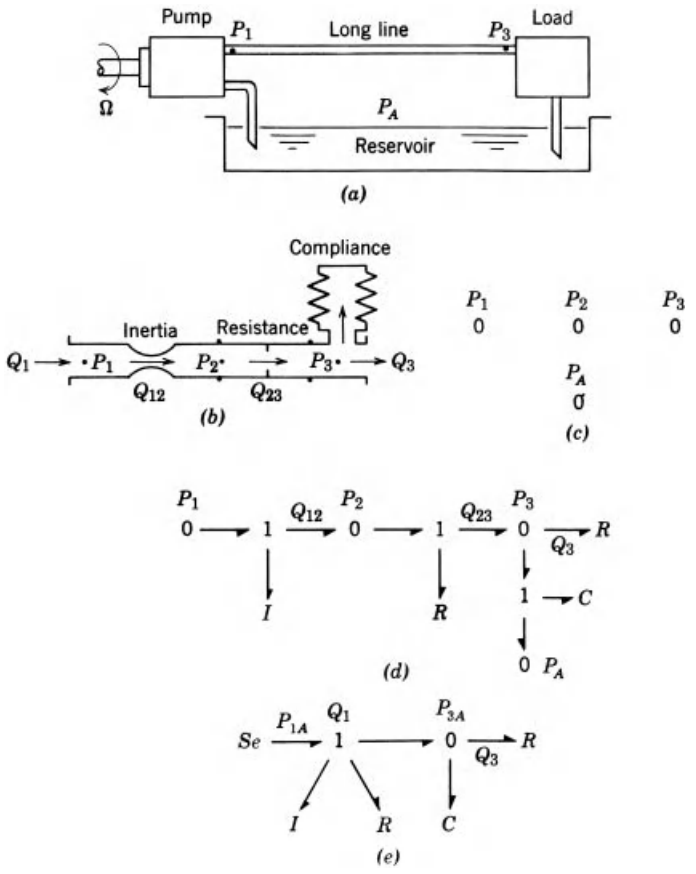


FIGURE 4.29. Hydraulic system with a long, flexible line.

displayed in Figure 4.29*c*. Elements are inserted as shown in part *d*. The inertia and resistance effects are inserted directly; the compliance goes between local pressure and atmospheric. Finally, the pump is added as a pressure source, the load as a resistance, and the graph is simplified. All pressures are measured relative to atmospheric pressure.

Inspection of Figure 4.29*e* with regard to the line model itself shows that a simple 1–0 structure with I, R, and C attached is the basic 2-port line representation. By subdividing the line, making such a model for each unit, and then adjoining them port to port in cascade, a model of greater accuracy (albeit greater complexity) can be constructed. This approach applies equally well to many other types of devices that are distributed, such as rods, shafts, plates, and electrical lines. It is discussed in more detail in Chapter 10.

4.3.5 An Acoustic Circuit Example

Figure 4.30 shows a reactive type of muffler that works primarily through the interaction of inertia and capacitive effects. (This is why it is called a *reactive* muffler.) Part *a* of the figure represents a cross section of a straight-through design consisting of a shell divided into two cavities, of volumes V_c and V_b, and by an internal plate and having two tubes allowing a steady flow to pass through. An imposed flow, having both steady and pulsating components, is imposed at the left side and a flow exits at the right into the atmosphere. If the muffler is effective, the pulsating components of the flow at the exit should be reduced, as compared with the input fluctuations for the frequency range of interest.

For this example, we will assume that the frequencies of interest are sufficiently low that the tubes can be represented as a single I-element and the cavities can be represented by C-elements. For higher frequencies, a much more complex model might be required similar to the long line model shown in Figure 4.29. (It should also be noted that this model applies better to an air-conditioning duct muffler in which the acoustic approximation is valid than to an automobile muffler that has very large pressure fluctuations from the engine exhaust ports.).

The bond graph shown in Figure 4.30*b* can be constructed by following the procedure outlined above. First, we establish three 0-junctions for the pressures P_a, P_b, and P_c. The I-elements representing the tubes are inserted between the appropriate 0-junctions using 1-junctions, as in Figure 4.27. A minor complication at the exit is the R-element representing the radiation impedance of the flow as it enters the infinite atmosphere. It is a standard exercise in acoustics to derive an expression for this impedance and to show that for low frequencies, the impedance can be represented as a resistance and an added inertia effect. The inertia effect can be included by assuming that the final tube is longer than its physical length by an amount called the *end correction*, which has been mentioned previously. (For a more thorough discussion of these effects, see Reference [7], Sections 8.12 and 9.3.)

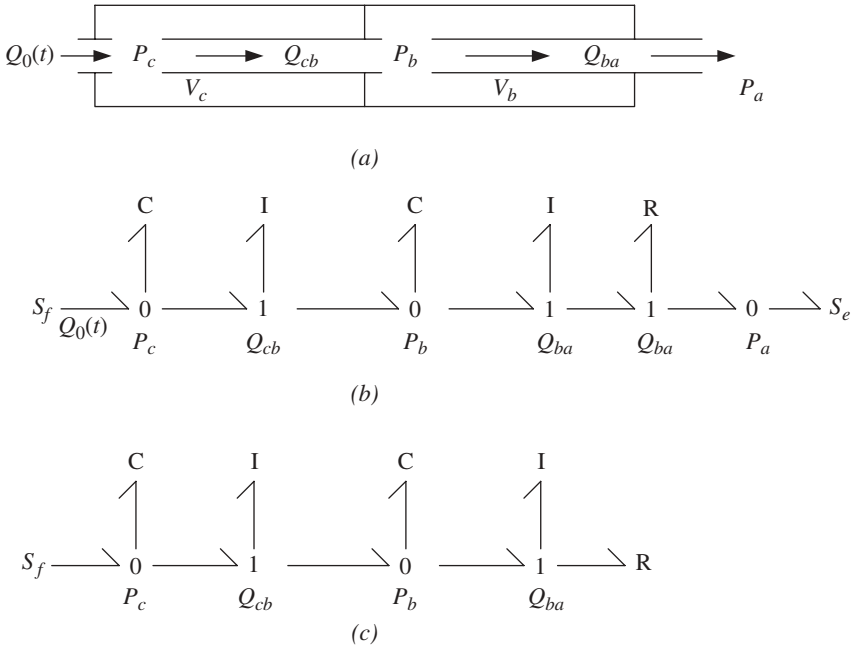


FIGURE 4.30. Low-frequency shell and tube reactive muffler model.

The flow through the radiation impedance R-element is the same as the flow through the last tube, so an extra 1-junction is used. Equivalently, a single 4-port 1-junction could be used for the I- and R-elements.

The atmospheric pressure is set by an effort source, and the internal pressures are set by capacitances, as in Figure 4.24. The input flow is determined by a flow source. The capacitance parameters are evaluated using Eq. (4.41), and the inertia parameters using Eq. (4.59).

A simplification is possible if all pressures are defined as deviations from atmospheric pressure. This means that $P_a \equiv 0$ and allows all bonds on which atmospheric pressure appears to be eliminated. When this is done and a 2-port 1-junction with a through sign convention to the R-element is eliminated, the bond graph of Figure 4.30*c* results.

Although the techniques and limitations of modeling hydraulic and acoustic circuits are quite complex in detail, it should be clear that there is a useful analogy between fluid and electrical circuit models.

4.4 TRANSDUCERS AND MULTI-ENERGY-DOMAIN MODELS

Transducers are devices that couple subsystems of distinct energy domains. For example, there are electromechanical devices such as motors, generators, and relays, there are hydraulic-mechanical devices such as pumps, motors, and rams,

and there are even devices such as electrohydraulic valves that involve electrical, mechanical, and hydraulic power. Chapter 8 deals in some detail with various types of transducers, so here we will discuss only some relatively simple transducers based on 2-port transformers and gyrators. Some of these devices have already been shown schematically in Figures 3.8 and 3.9. These common transducers will allow the modeling of multi-energy-domain systems using bond graphs.

A major reason for studying bond graphs is that they provide a uniform and precise way to represent system models when several forms of power and energy are involved in a system to be modeled. Bond graphs practically force the modeler to consider how power and energy are conserved across an ideal transducer and how added R-elements modeling dissipative effects account for efficiencies of less than 100% for real devices. Finally, it is straightforward to add I- and C-elements to a bond graph to model the dynamics of transducers.

4.4.1 Transformer Transducers

Some transducers are quite easy to understand. Consider, for example, the hydraulic ram shown in Figure 3.8*e* and again in Figure 4.31.

From the schematic diagram in Figure 4.31*a*, it is clear that the pressure in the hydraulic fluid acts on the piston face of area A to create a force. It should also be clear that the velocity of the piston times the area equals the volume flow rate, Q. Thus, the ideal transducer is represented by the bond graph in part *b*, which implies the Eqs. (3.10) given previously in Chapter 3:

$$F = AP, \quad AV = Q, \quad FV = PQ, \tag{4.61}$$

Note that the sign-convention half-arrows are consistent with the schematic diagram. If at some instant the force and velocity are positive in the directions

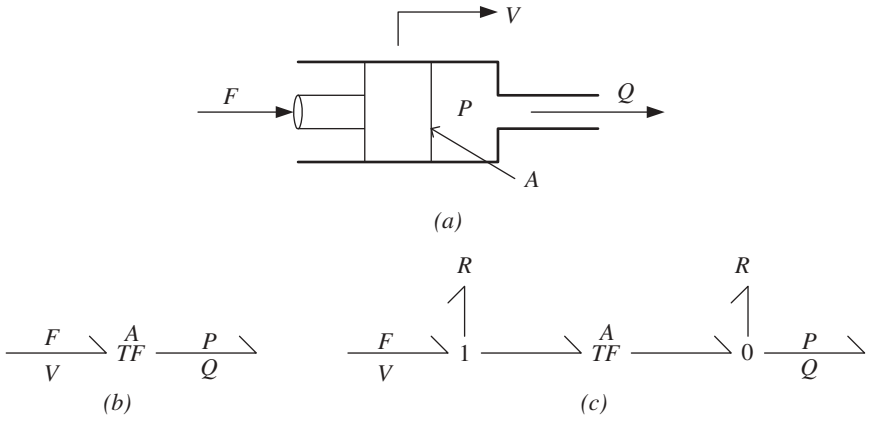


FIGURE 4.31. Hydraulic ram transducer: (*a*) schematic diagram; (*b*) ideal transformer model; (*c*) friction and leakage resistors added.

indicated, then mechanical power is being supplied to the device and hydraulic power is being delivered to some other device. This "through" power convention for transformers ensures that the area A that appears twice in the equations has the same sign both times. (When the sign half-arrows do not flow through the transformer, then $-A$ must appear in one of the equations and the mechanical power would always be the negative of the hydraulic power.)

Actually, the ideal transformer transducer really applies only to the effects at the face of the piston. All real devices contradict the idea that mechanical power can be converted with no loss to hydraulic power. As pointed out in Chapter 8, the modeling of a real transducer involves adding a number of elements to the ideal transducer to account for power losses and, possibly, for inertia and capacitance dynamic effects as well. Part c of Figure 4.31 shows how loss effects could be added to the simple ideal model.

On the mechanical side, the added 1-junction and resistor are used to include mechanical friction effects. Typical pistons have piston rings or packing that introduce frictional forces on the piston. This means that the net force on the piston rod is not just AP but also includes the frictional force. The 1-junction adds the forces and also enforces that the piston rod, the friction force element, and the piston all have the same velocity, V. Since the piston ring R-element can only absorb power, the hydraulic power will now be less than the mechanical power at the piston rod. On the hydraulic side, the combination of the 0-junction and resistance represent possible leakage past the piston. The pressure at the hydraulic outlet, that on the piston face, and that at the piston ring leakage location are assumed to be equal. The 0-junction makes the volume flow rate out of the ram to be AV minus the leakage flow through the leakage resistor. This hydraulic resistor also reduces the power efficiency of the transducer.

Positive displacement hydraulic pumps and motors consist of a number of pistons and a mechanism for moving them back and forth as a function of the angular position of a shaft. When there are a fairly large number of pistons, say seven or nine, the angular speed of the shaft ω is related to the volume flow rate Q by a nearly constant coefficient, T. (The coefficient actually varies a little with the position of the shaft. This makes the flow rate have a little ripple on top of a steady flow when the shaft rotates steadily.) Considering that the hydraulic ram can be represented by a transformer, it is not surprising that positive displacement machines can also be modeled using transformers, as shown in Figure 4.32.

In part a, the schematic diagram shows the torque and the angular speed of the shaft and the pressures at the inlet and outlet ports P_A and P_B. When the leakage is negligible, the same flow Q enters and leaves the device. The hydraulic power has to do with the flow Q and the pressure difference $P_B - P_A$.

The ideal model of the pump or motor is shown in bond graph form in Figure 4.32b. You should check to see that the 1-junction half-arrows correctly relate the pressures P_A, P_B and $(P_B - P_A)$. If the transduction coefficient relating angular speed and volume flow rate is defined by the relation

$$TQ = \omega, \tag{4.62}$$

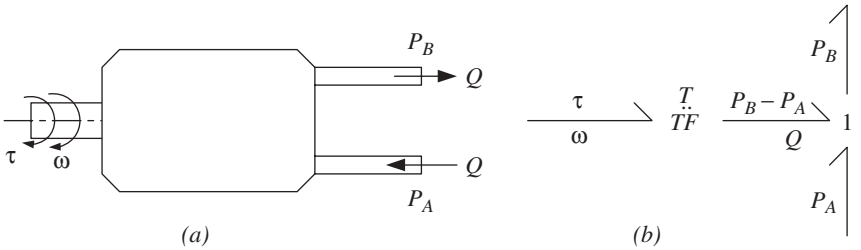


FIGURE 4.32. Hydraulic pump or motor modeled using a transformer.

then it follows that the torque–pressure relation must be

$$P_B - P_A = T\tau, \tag{4.63}$$

since the power relation for the ideal case must be

$$(P_B - P_A)Q = \tau\omega. \tag{4.64}$$

The choice of signs on the transformer may seem to indicate that the model represents a pump, but even with these sign choices, the mechanical and hydraulic powers could both be negative, indicating that the device model could as well represent a hydraulic motor.

Although the bond graph in its present form is ideal with no losses and no dynamic behavior, it should be clear that the model could be augmented by adding resistors to model losses, as was done in Figure 4.31, and one could add an I-element on a 1-junction on the mechanical side to represent the moment of inertia of the rotating parts. Chapter 8 discusses the modeling procedure based on starting with an ideal model and then adding loss and dynamic elements as appropriate.

4.4.2 Gyrator Transducers

While many types of transducers that can be represented as transformers are easily understood, those represented by gyrators often seem harder to understand. The importance of electromechanical transducers, such as rotary and linear motors or voice coils in electrodynamic loudspeakers, justifies a short discussion of why gyrators are used to describe a number of useful devices.

Figure 4.33 shows the basis of the gyrator models. A current-carrying conductor is moving in a magnetic field under the action of an applied force that is equal but opposite to a magnetic force. Because of the motion of the conductor, a voltage is induced in the conductor.

Because all the various vectors are at right angles to each other, the laws governing this situation take particularly simple forms. Faraday's law of magnetic

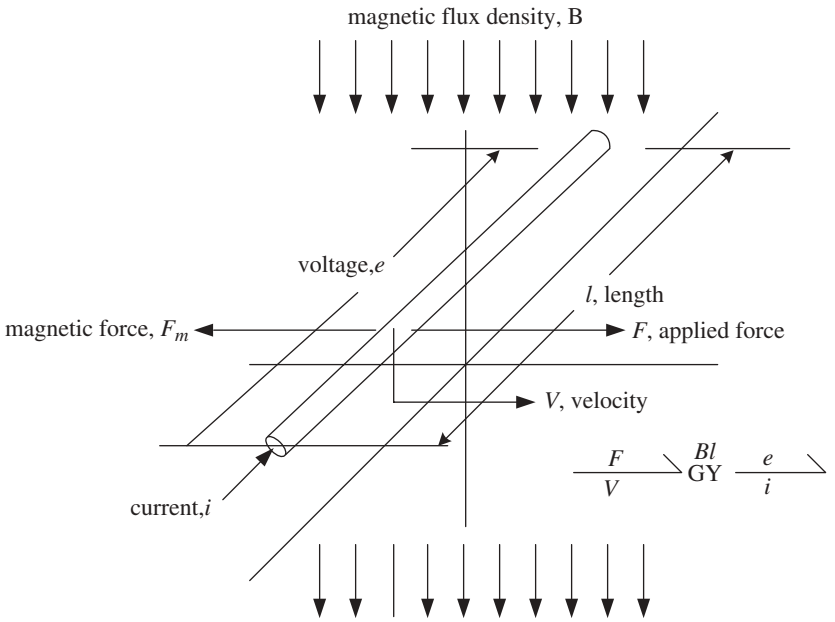


FIGURE 4.33. Gyrator model of a current-carrying conductor moving in a magnetic field.

induction relates the induced voltage to the velocity:

$$e = BlV. \tag{4.65}$$

The Lorentz force law relates the applied force to the current:

$$Bli = F. \tag{4.66}$$

Note that, by multiplying Eq. (4.65) by Eq. (4.66), the electrical and mechanical powers are shown to be always equal,

$$ei = FV, \tag{4.67}$$

and in both Eqs. (4.65 and 4.66) an effort at one port is related to a flow at the other port. Thus, the power-conserving 2-port transducer of Figure 4.33 is properly represented as a gyrator.

Just as the simple case of the hydraulic ram leads to the representation of hydraulic pumps and motors as transformers, so too does the case of a conductor in a magnetic field lead to the representation of electric motors and generators as gyrators. Figure 4.34 shows several representations of dc motors or generators. Internally, most motors have complex arrangements of coils of wires carrying

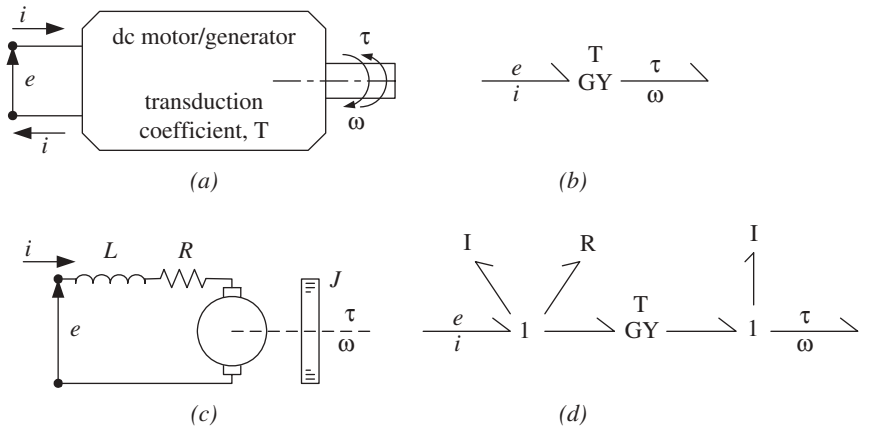


FIGURE 4.34. Motor/generator models: (*a*) sketch of device; (*b*) bond graph of ideal model; (*c*) schematic diagram of extended model; (*d*) bond graph of extended model.

currents moving in magnetic fields. In the case shown, it is assumed that the field is due to permanent magnets and a commutator or other means switches the coils of wire in such a way that the voltage at the terminals induced by the rotary motion is essentially proportional to the angular speed of the rotor. In addition, the torque produced is proportional to the current at the terminals of the device. (More detailed models of motors are able to predict minor ripples in the torque of actual dc motors as well as small voltage fluctuations in dc generators.)

Using the symbol T as a transduction coefficient, the equations for the ideal model of Figure 4.34*b* are

$$e = T\omega, \quad Ti = \tau, \quad ei = \tau\omega. \tag{4.68}$$

The gyrator correctly represents the essence of the transduction process, but in many cases it needs to be supplemented to model real devices that are never completely ideal. The schematic diagram of Figure 4.34*c* indicates that the windings of the motor have inductance and resistance and that the rotor has a moment of inertia, J. Part *d* of the figure shows an extended bond graph including these effects. The 1-junction on the electrical side enforces the fact that the same current flows from the terminals through the inductance and the resistance and participates in the transduction.

On the mechanical side, the 1-junction indicates that there is a single angular velocity for the output shaft, the rotary inertia, and the transduction process. In this model, the electrical power is not instantaneously equal to the mechanical power, as it was for the ideal gyrator model. Not only is there now a loss element, R, but also the I-elements can store and release energy during transient

conditions. The idea of starting with ideal transducers and adding realistic effects is developed further in Chapter 8.

4.4.3 Multi-Energy-Domain Models

Figure 4.35*a* shows a system for positioning a machine tool table bed. It includes an electric dc motor connected to a power supply driving a hydraulic pump. The power supply is modeled as a voltage source. The motor and pump may run continuously, because the combination valve (with handle position variable, x) can recycle some of the hydraulic fluid to the sump and can divert some of the flow to the hydraulic motor. The hydraulic motor in turn drives a rack and pinion arrangement that moves the table bed.

There are a number of transducer devices in this system. The dc motor changes electrical power to mechanical power, and the hydraulic pump and motor exchange mechanical and hydraulic power. One could even consider the rack and pinion as a transducer between rotary and translational mechanical power. (A transformer represents the rack and pinion, because the contact radius of the

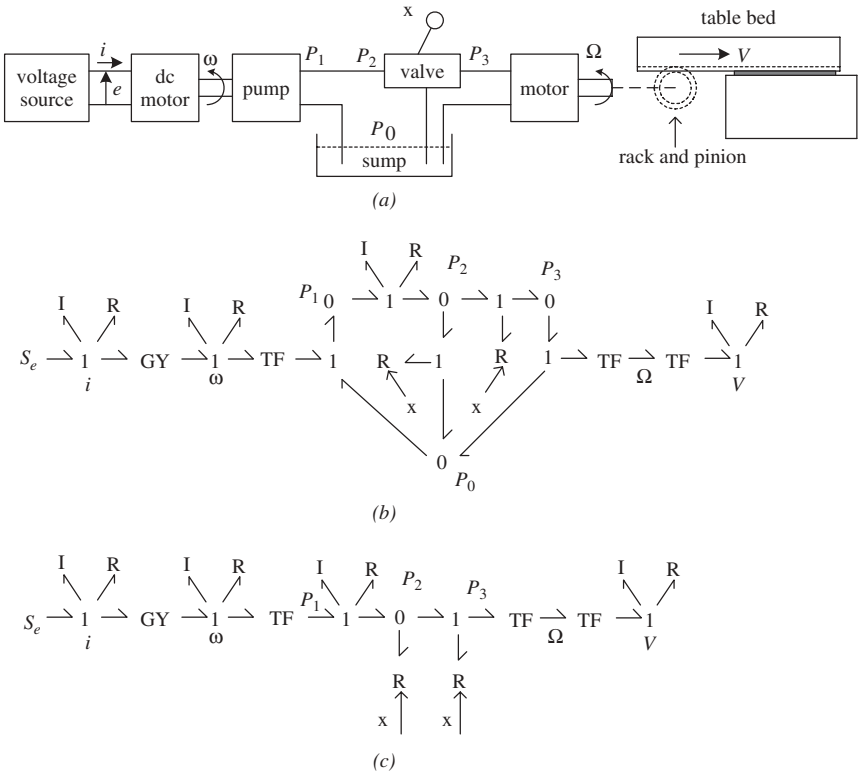


FIGURE 4.35. An electrical–hydraulic–mechanical system: (*a*) schematic diagram; (*b*) bond graph including atmospheric pressure; (*c*) bond graph using gage pressures.

pinion relates the angular velocity of the pinion to the linear velocity of the table, as well as the force on the table to the torque on the pinion.) This system is intended to demonstrate that bond graph models are capable of representing multi-energy-domain systems in a uniform and precise manner.

The schematic diagram of the system shown in part *a* of Figure 4.35 shows a number of important system variables, such as pressures and angular velocities, but it does not make clear just how the various devices are to be modeled. The first stage in the construction of the bond graph for the system shown in part *b* of Figure 4.35 begins to make clear the assumptions that are being made to model the various devices.

The left side of the bond graph in part *b* represents the voltage source and the dc motor. The motor model of Figure 4.34 has been used with some modifications. The first I- and R-elements clearly represent the inductance and resistance of the armature windings, but the next I–R combination attached to the 1-junction, with ω as the common flow, requires some explanation. The I-element represents not only the rotary inertia of the motor, as in Figure 4.34, but also the inertia of the pump, since both devices are assumed to rotate with the same angular speed ω. The R-element represents frictional losses associated with both the electric motor and the pump.

Transformers represent the hydraulic pump and motor, as in Figure 4.32. The hydraulic pressures P_0, P_1, P_2, and P_3 are represented as existing on 0-junctions, as in the hydraulic circuit construction method described in Section 4.3.4. A study of the sign half-arrows will reveal that the pressure difference acting on the pump is $P_1 - P_0$ and the pressure difference acting on the motor is $P_3 - P_0$.

The line between the pump outlet and the combination valve has been modeled as an I–R combination that includes fluid inertia, as in Figure 4.27, and fluid resistance, as in Figure 4.23. Evidently, no fluid or pipe compliance effects are included in the model. This is a modeling decision that may or may not be justified in a particular case. In any case, the bond graph is absolutely clear about which possible effects are included in the model and which are not.

The combination valve has a variable resistance reacting to the pressure difference $P_2 - P_0$ and another variable resistance reacting to the pressure difference $P_2 - P_3$. Both these valve resistances change with the handle position, x. The double-arrow active bond carrying the variable x indicates this. It is thus assumed that moving the handle does not transfer any significant power to the system, although the handle position does change the valve resistance laws relating pressure and volume flow rate.

Finally, the rack and pinion transformer connects the hydraulic motor to the table bed model, which is represented by an I-element for the table mass and an R-element for the table friction.

The last step to arrive at the bond graph of Figure 4.35*c* is to consider that the atmospheric pressure P_0 is zero. This allows three bonds to be removed and several 2-port 0- and 1-junctions with "through" sign half-arrows to be eliminated. This procedure is similar to the procedure for electrical circuits, but in this case

the result is that the efforts representing pressures are to be considered gage pressures.

As this example should make clear, there is an art, as well as a science, to making a useful mathematical model. There is a scientific basis for the ideal models that involve power or energy conservation, and these are elegantly incorporated in bond graph elements. However, there are a number of effects that can be added to the ideal elements to account for effects that occur in real systems. These extra loss and dynamic elements must be added with restraint, since an unduly complicated model that is hard to understand is often just as bad as an oversimplified model. In every case the model should be adjusted to provide guidance for the specific problems at hand. The best model is generally the simplest model capable of demonstrating those aspects of the behavior of the system that need to be understood.

Good modelers are always ready to modify a preliminary model by simplifying it, by removing elements, or by adding elements to bring in new effects. Often the changes are suggested based on the results of analysis or simulation of the model, particularly if there is a lack of agreement between the model predictions and experimental results.

REFERENCES

[1] D. K. Schroder, *Semiconductor Material and Device Characterization*, New York: Wiley, 1998.

[2] J. L. Merriam and L. G. Kraige, *Engineering Mechanics: Statics and Dynamics*, 4th ed., New York: Wiley, 1997.

[3] A. G. Bose and K. N. Stevens, *Introduction to Network Theory*, New York: Harper & Row, 1965.

[4] R. H. Sabersky, A. J. Acosta, and E. G. Hauptmann, *Fluid Flow: A First Course in Fluid Mechanics*, 2nd ed., New York: Macmillan, 1971.

[5] J. F. Blackburn, G. Reethhof, and J. L. Shearer, *Fluid Power Control*, New York: Wiley, 1960.

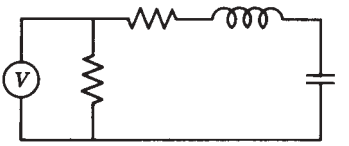
[6] J. Thoma, *Modern Hydraulic Engineering*, London: Trade and Technical Press, 1970.

[7] L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders, *Fundamentals of Acoustics*, 3rd ed., New York: Wiley, 1982.

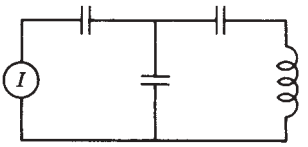
PROBLEMS

4-1. Make a bond graph model for each of the following electrical circuits. Use the inspection method whenever possible. If the circuit has voltage and current orientations, make the bond graph have equivalent power directions.

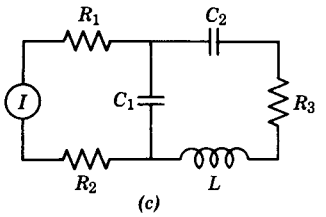
Note: The E-meter may be modeled as an open-circuit pair, and the auditor may be modeled as a resistance, R_A.


V

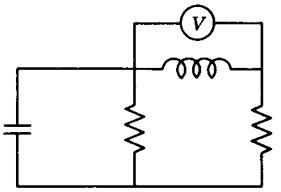
(a)


I

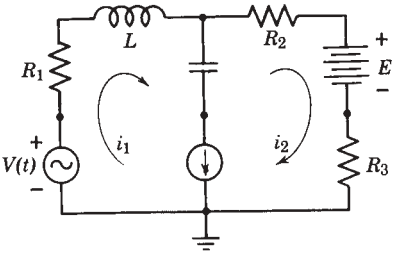
(b)


R1
C2
I
C1
R3
R2
L

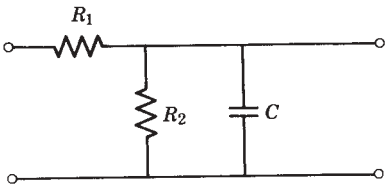
(c)


V

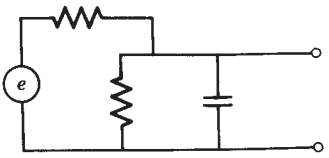
(d)


L
R2
R1
+
E
−
+
V(t)
−
i1
i2
R3

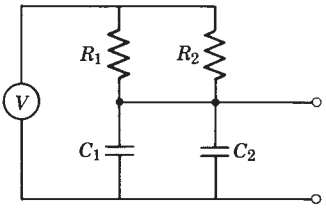
(e)


R1
R2
C

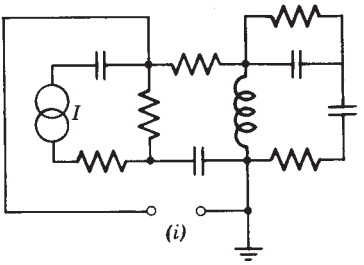
(f)


e

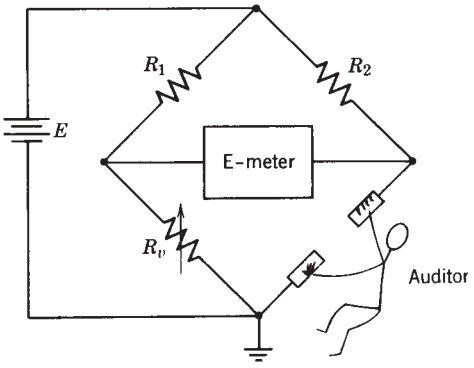
(g)


R1
R2
V
C1
C2

(h)

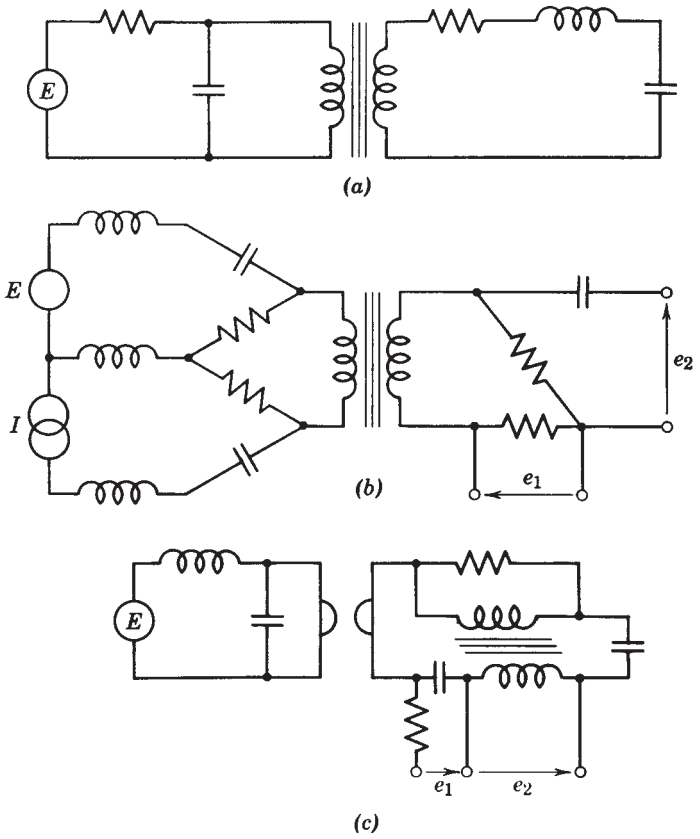

I

(i)

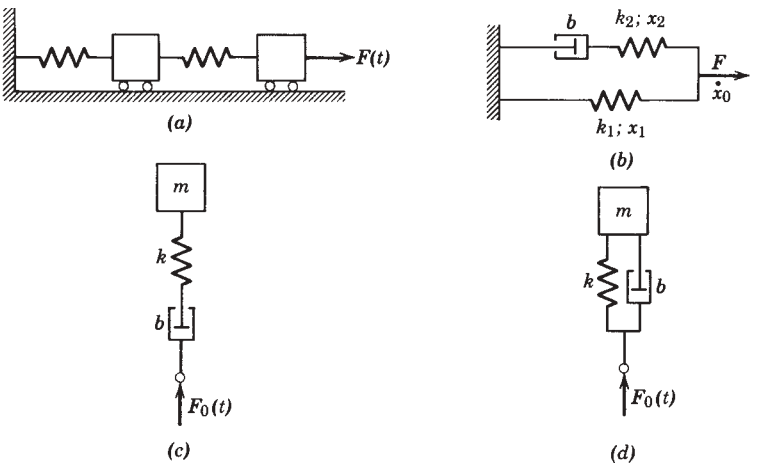

R1
R2
E
E-meter
Rv
Auditor

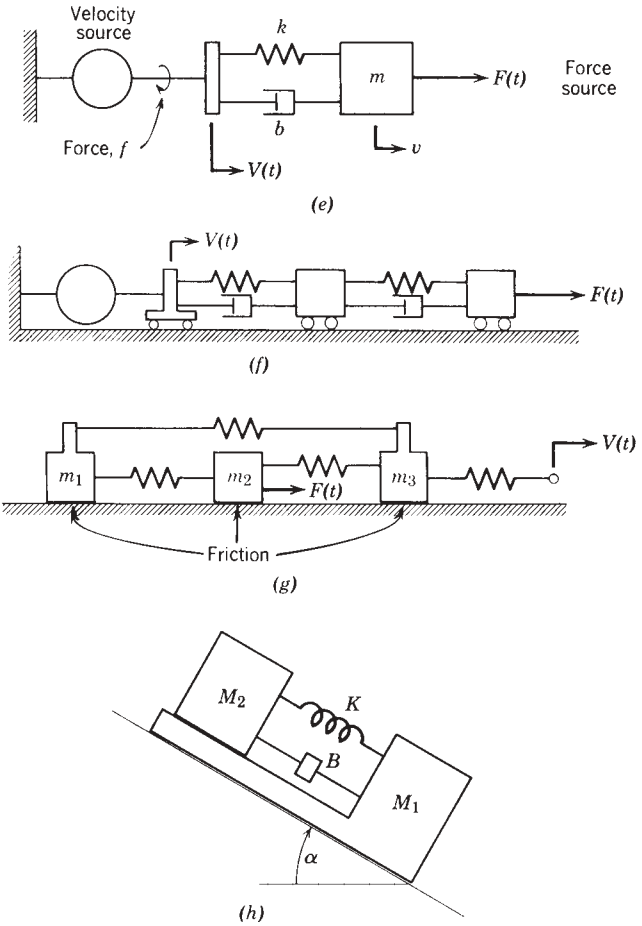
(j)

4-2. Make a bond graph model of each of the electrical networks shown below.

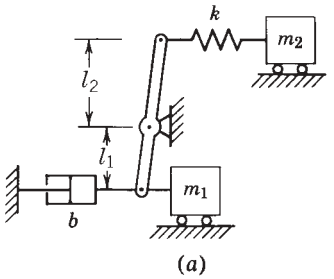


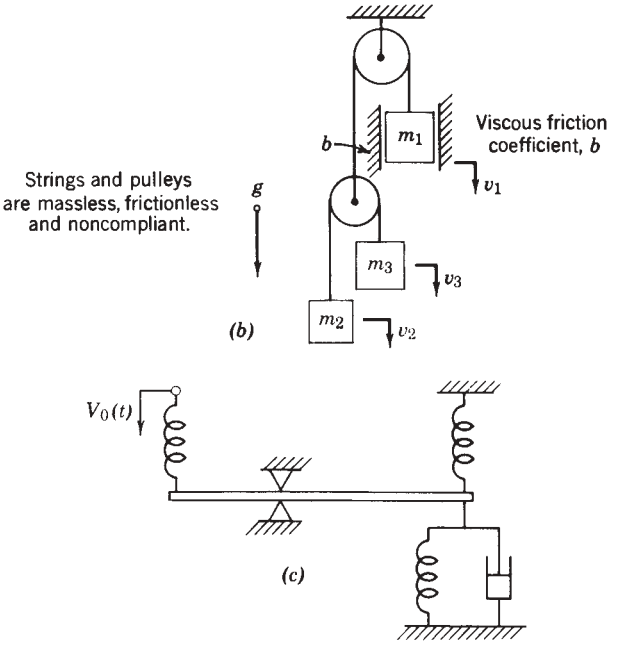
4-3. Make a bond graph model of each of the translational mechanical systems shown below.



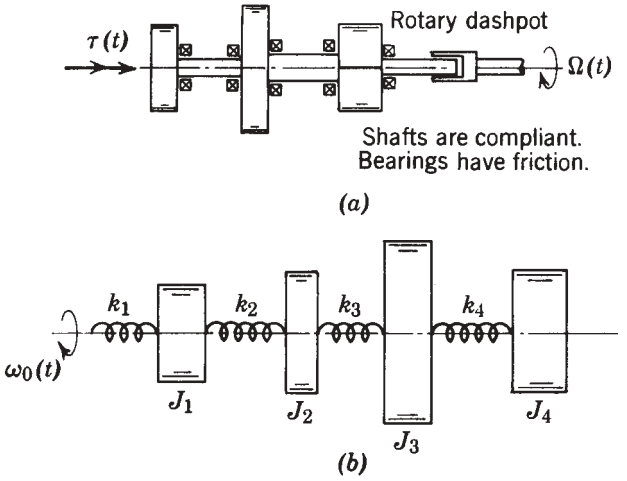


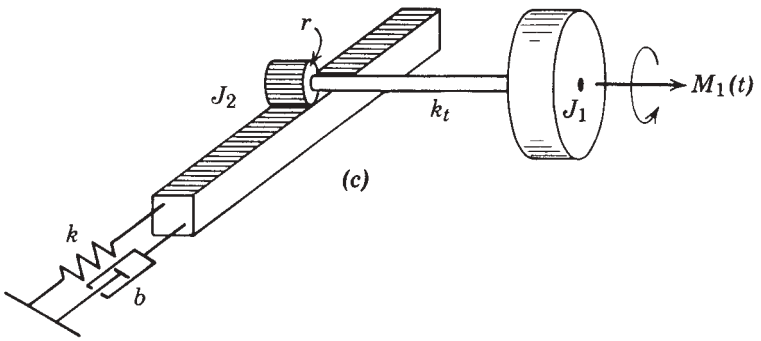
4-4. Find a bond graph model for the following mechanical networks.

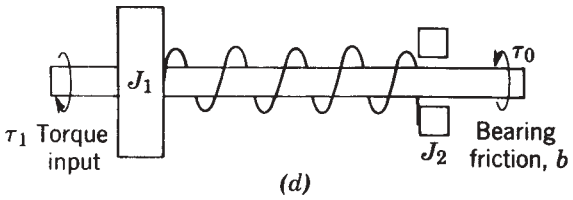


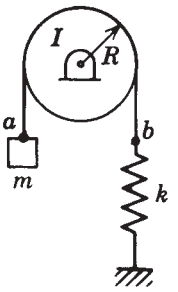


4-5. For the following mechanical systems involving fixed-axis rotation, find a bond graph model in each case. In the rack-and-pinion system shown, several frictionless guides and bearings are omitted for clarity.

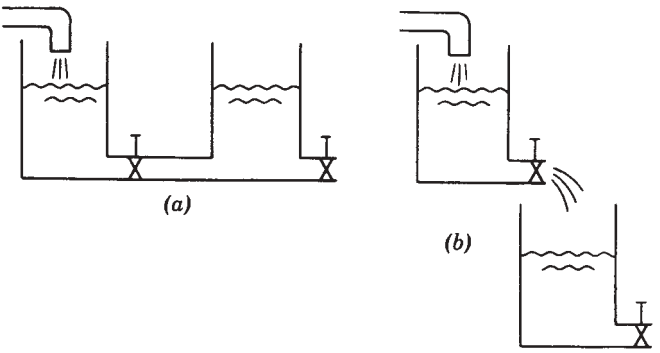








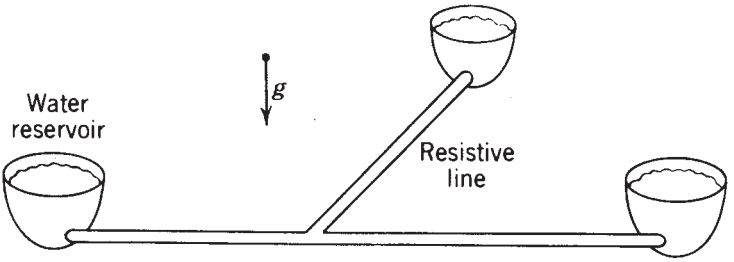
4-6. Two hydraulic systems that are closely related but not identical are shown below. Make a bond graph model for each.



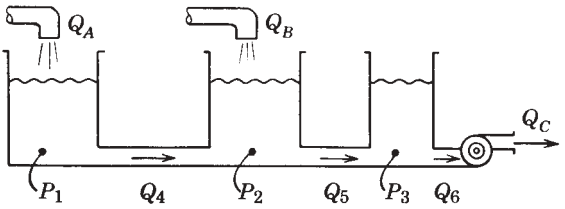
4-7. A simplified model of a water storage system is shown below. Assume the three tanks behave as (nonlinear) capacitances. Assume the three conduits have resistance only.

(a) Make a bond graph model.

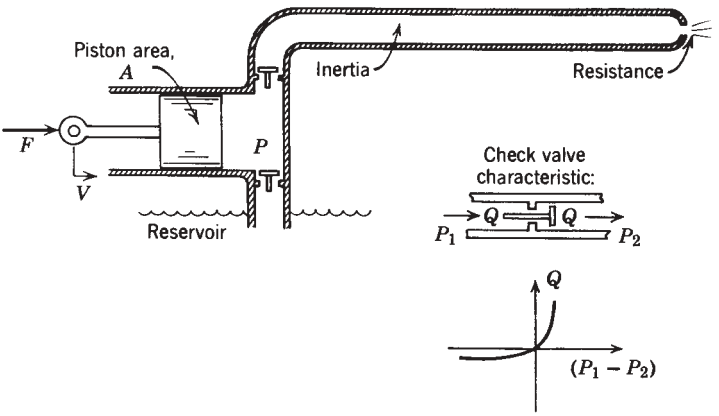
(b) Introduce inertance effects for each of the conduits into your model for (a).



4-8. In the hydraulic system shown below, the pipes containing flows Q_4 and Q_5 have both inertia and friction effects present. The outflow line (Q_6) has only friction effects. The control pump maintains a desired flow, Q_C, independent of pressure. Make a bond graph model of the system.



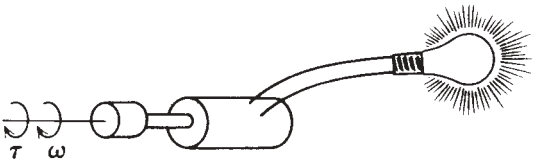
4-9. In the positive-displacement pump shown below, the piston moves back and forth and the check valves act like unsymmetrical resistances, allowing relatively free forward flow while impeding the backflow greatly. Inertia effects are important in the outlet line, while the nozzle is a resistive restriction.



(a) Make a bond graph model that can be used to relate the mechanical port variables (F, V) to the fluid port variables (P, Q).

(b) Noting that an electrical diode is analogous to a fluid check valve, use the bond graph of part (a) to find an equivalent electrical network to the pump system.

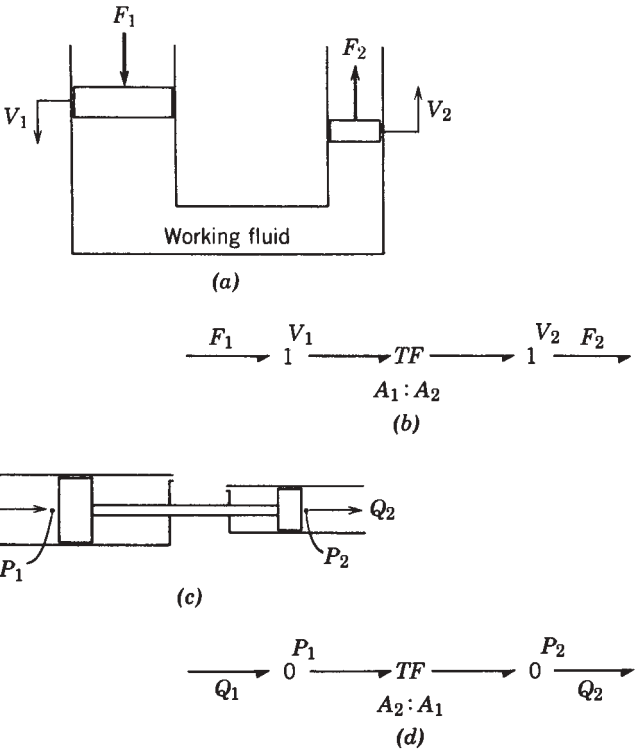
4-10. Consider a permanent-magnet generator of the type often used to power bicycle lights:



(a) Suppose the generator to be an ideal gyrator in its transducer action. Make a bond graph model of the system.

(b) Suppose the lamp shortcircuits. Is your model from (a) still useful? If not, modify it so that it can handle this situation.

4-11. Two related hydraulic devices are shown below, a hydraulic jack and a ram.



(a) If the inertia of the sliding parts and friction at the seals are important, modify the models given in the figure for the devices to include these effects.

(b) If compliance of the working fluid for the jack and of the shaft for the ram is important, modify your models from (a) to include these effects.

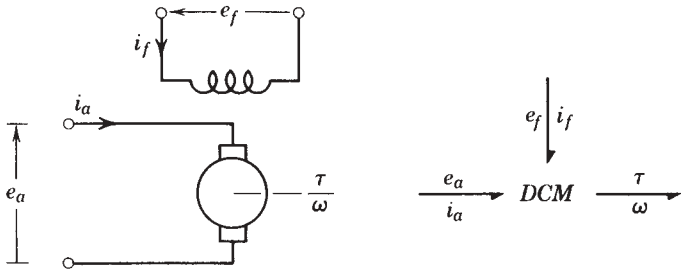
4-12. A schematic diagram of a basic dc machine is shown below. The principal electrical effects are armature resistance and inductance and field resistance and inductance. The coupling between armature and shaft powers is modulated by the strength of the magnetic field, set by i_f. A simple assumption for coupling is

$$\tau = T(i_f)i_a = (Ki_f)i_a$$

and

$$e_a = T(i_f)\omega = (Ki_f)\omega$$

where $T(i_f)$ is the modulating value, assumed linear with i_f.



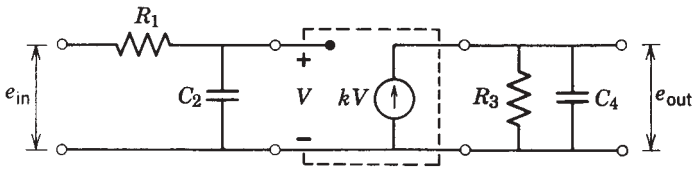
(a) Make a bond graph model of the basic dc machine using a modulated gyrator (*MGY*) as the heart of the coupling.

(b) Introduce mechanical rotational inertia and dissipation effects into your model from (a).

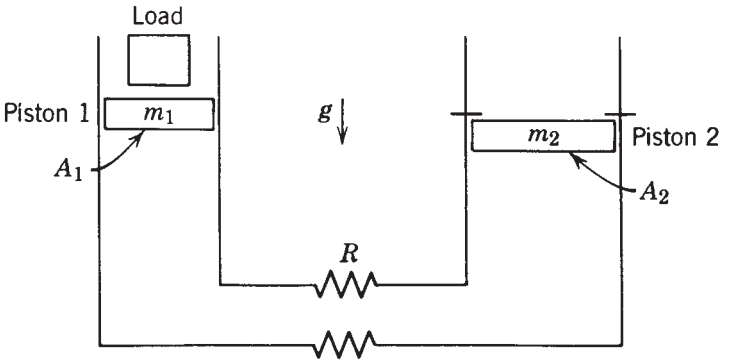
(c) Operate your model from (b) in the generator mode. What changes must be made?

(d) Calculate the power efficiency for the dc machine operating as a motor and as a generator.

4-13. The network shown below includes a voltage-controlled current source. Make a bond graph model of the network using an active 2-port as the coupling element.



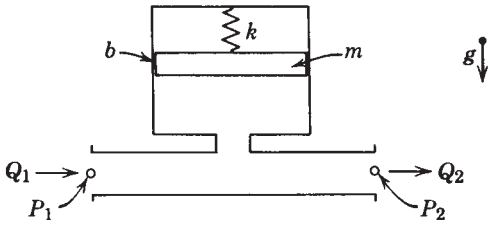
4-14. A load of mass M is suddenly dropped onto piston 1 from height L.



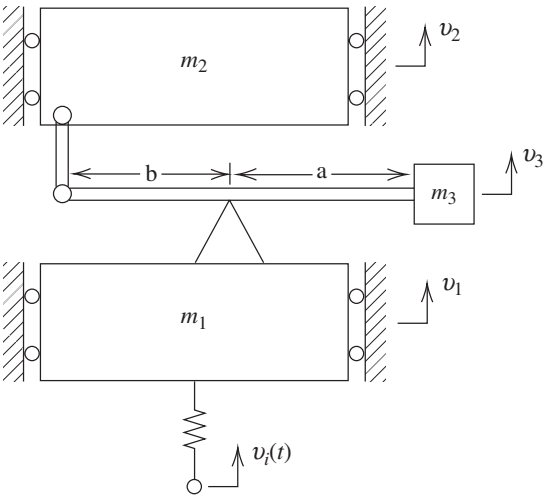
(a) Make a bond graph model of the system, assuming piston 2 is locked in place by pins. Also assume the load stays joined to piston 1.

(b) If piston 2 is not locked in place, but pistons 1 and 2 were in equilibrium before the load was dropped, modify the model in (a).

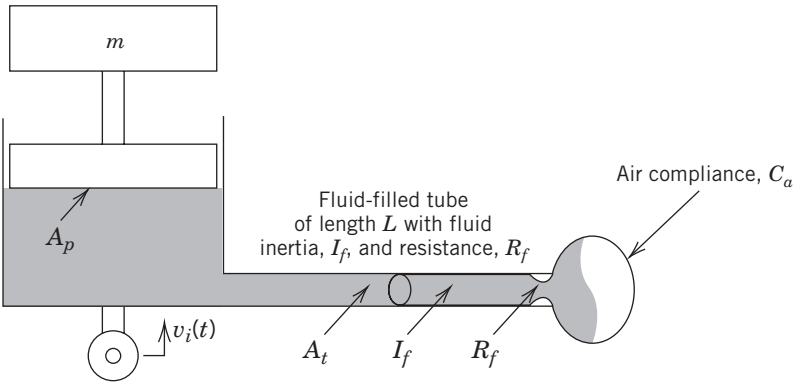
4-15. Model the spring-loaded accumulator device shown below. Include inertia and dissipation effects.



4-16. For the device shown, construct a bond graph model. For the positive directions indicated, assign a power convention.

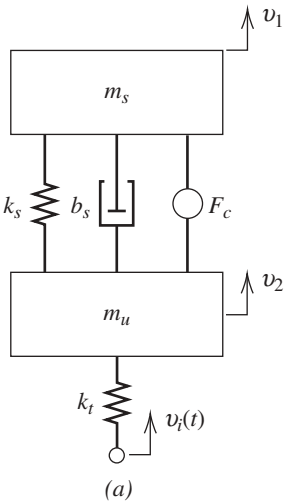


4-17. The coupled mechanical and hydraulic system is used to isolate the mass, m, from the ground motion, $v_i(t)$. Construct a bond graph model, show your assumed positive e and f directions, and assign an appropriate power convention.



4-18. A quarter-car model of a vehicle is shown with an idealized actuator capable of generating any control force, F_c.

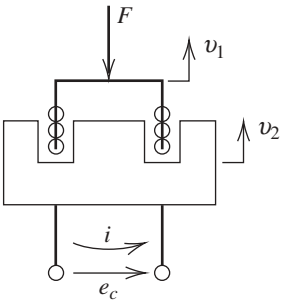
(a) Construct a bond graph model, identify positive directions, and assign a power convention.



(a)

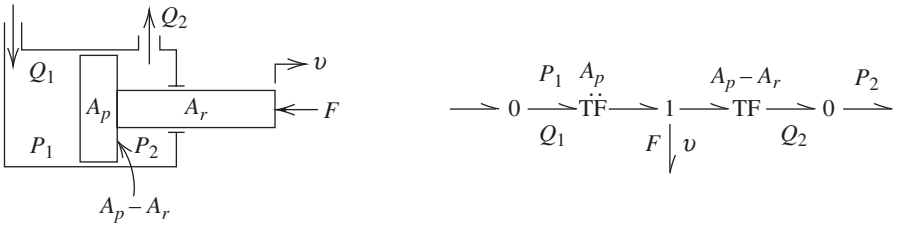
(b) A voice coil actuator with winding resistance, R_w, is to be installed as the force actuator to replace the idealized control force. The actuator will be voltage, e_c, controlled. Include the force actuator in the quarter-car model and construct the overall bond graph model.

winding resistance, R_w
coil length, l
permanent magnetic with flux density B
transduction constant $T = BL$
force, $F = Ti$

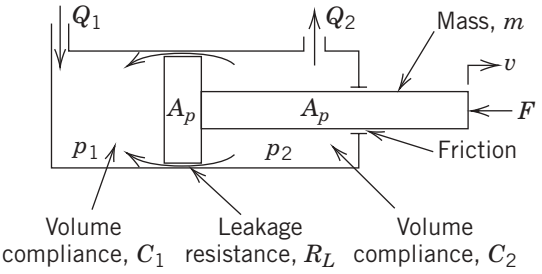


(b)

4-19. The hydraulic piston cylinder device is a fundamental component of many motion control systems. Hydraulics can generate large forces over large distances, which is necessary for big motion platforms such as flight simulators. The ideal form of the device shown is a transformer, as indicated in the bond graph.



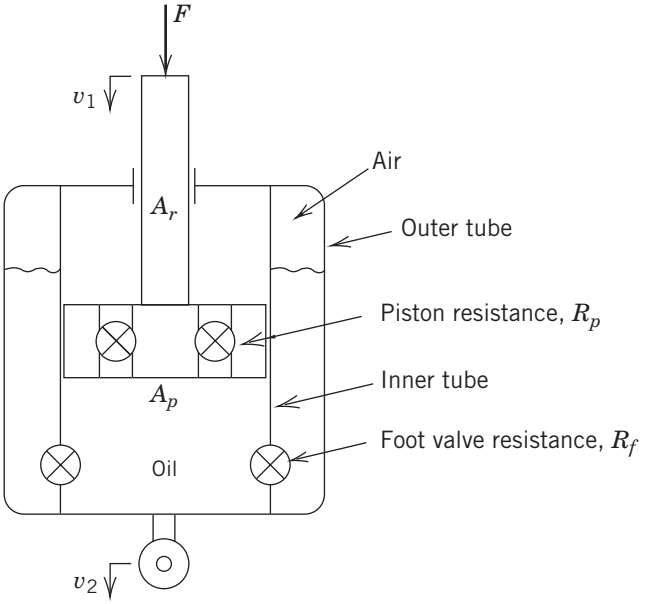
(a)



(b)

An actual device has inertia, friction, leakage past the piston, and compliance of the volumes of oil on both sides of the piston. For the realistic effects shown, construct a bond graph model and assign a power convention.

4-20. Shock absorbers for automobiles are very sophisticated devices. The piston head has several levels of check valves to generate a force–velocity characteristic quite different from what might be expected from simply forcing oil through an orifice. Also, to provide volume for the piston rod as the shock is compressed, most shock absorbers are of a twin-tube construction, as shown below. The valve that connects the inner and outer tubes is called the foot valve.



(a) Construct a bond graph model incorporating the dynamic elements shown and assign a power convention.

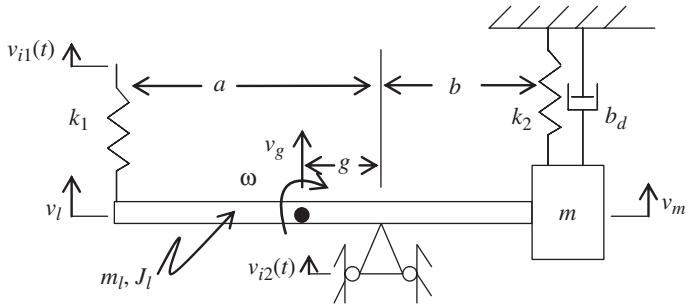
(b) State your assumptions about more realistic dynamic effects and include them in your model.

4-21. Install the shock absorber model from Problem 4-20 into the quarter-car model of Problem 4-18. Replace the damper with the shock model.

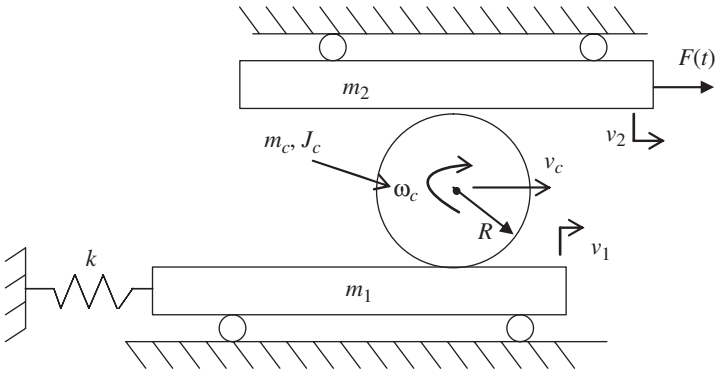
4-22. The system is similar to the example system of Figure 4.14. For this problem, the lever has mass, m, and center of mass moment of inertia, J_l. In addition, the pivot point has prescribed vertical motion set by the specified velocity, $v_{i2}(t)$. The spring at the left end has prescribed vertical velocity at the top, and there is a mass with spring and damper attached at the right end of the lever. The lever executes plane motion since it simultaneously translates and rotates.

(a) Derive the kinematic constraints for this system by relating v_l, v_{i2}, and v_m to the center of mass velocity and angular velocity, v_g and ω.

(b) State whether springs and dampers are positive in compression or tension and construct a bond graph model for this system.



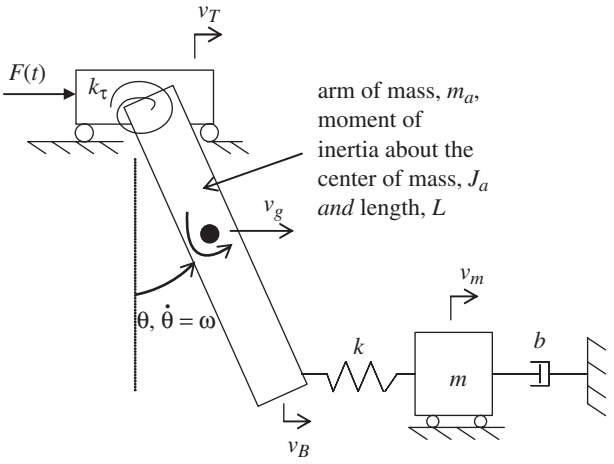
4-23. Masses m_1 and m_2 move only horizontally. Mass m_1 has an input force prescribed, and m_2 has an attached spring, k. The two masses are connected by a rolling element of mass, m_c, center of mass moment of inertia, J_c, and radius, R. The rolling element rolls without slip at the contact point with each mass.



(a) Derive the kinematic constraints relating the mass velocities v_1 and v_2 to the velocity and angular velocity of the rolling element.

(b) Construct a bond graph model of this system and show an appropriate power convention.

4-24. A gantry robot is a large manufacturing machine that consists of a very stiff table-like structure, atop which sits a platform that can be driven in the XY-plane. Suspended below the platform is the robot arm that will have a cutting tool of some kind attached at its lower end. The robot arm can be moved vertically, and the cutting tool can be moved in several directions. The result is a multiaxis cutting tool that can create large 3-D shapes. Although gantry robots are made as stiff and rigid as possible, sometimes small motions at the cutting tool end can limit the speed or precision of the machine.

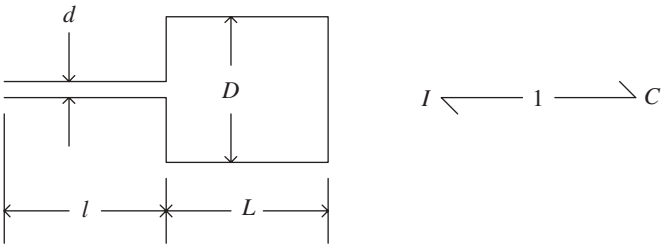


A simple model of the gantry robot is shown in the figure. An input force, $F(t)$, is used instead of the actual drive motors, and the top of the arm is constrained to move horizontally. The arm is rigid and has mass, m_a, moment of inertia about the center of mass, J_a, and length, L. The center of mass is located in the middle of the arm. The flexibility in the frame is represented by the torsional spring, k_τ, and the load being cut is represented by the spring, mass, damper, k, m, b, attached to the lower end of the arm. We desire a model that will allow the prediction of the motion–time history of the system for small angular deflections. Perhaps we can come up with some control strategy for the drive motors that will limit the vibratory motion of the arm end. For now,

(a) Derive the kinematic constraints relating v_T and v_B to the arm motion, v_g, and ω.

(b) Construct a bond graph model of this system and show an appropriate power convention.

4-25. Consider a simple model of a so-called Helmholtz oscillator formed by attaching a circular tube of diameter d and length l to a larger circular tube of diameter D and length L that is sealed at the end to form an acoustic volume.



Using the elementary bond graph shown, which neglects any resistance effects, the natural frequency of the oscillator is given by the formula

$$\omega_n = (1/IC)^{1/2}[\text{rad/s}] \text{ or } f_n = (1/IC)^{1/2}/2\pi[\text{cycles/s}] \text{ or } [\text{Hz}].$$

(a) Using the formulas for acoustic inertia and capacitance given in the text, find an expression for ω_n and f_n in terms of the dimensions of the device.

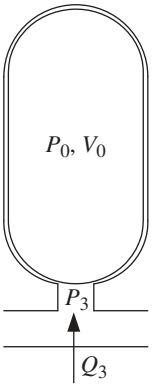
(b) What would be the natural frequency in air at standard conditions if $d = 0.01$ m and $D = l = L = 0.1$ m? Do you think you could hear this frequency if you blew across the end of the open tube as you would across an empty bottle?

4-26. Consider a length of hydraulic hose to be used to operate the ram of a high-speed actuator. There is a concern that the compliance in the system may slow down the system response. You are to decide whether the hydraulic compliance due to the compressibility of the hydraulic fluid or the flexibility of the hose walls is most important. Use these dimensions: nominal radius of the hose 15 mm, wall thickness 5 mm. Assume that the modulus of elasticity for the hose is about the same as for hard rubber and the bulk modulus for hydraulic oil has the value given in the Appendix.

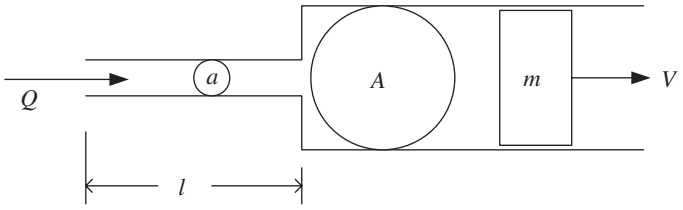
4-27. Consider a system that is a combination of Figure 4.28 and Figure 4.29. That is, a hydraulic pump is connected to a hydraulic motor by one line to a relief valve and by another line from the valve to the motor. In the new system, both lines are long and to be represented by a combination of inertia, resistance, and capacitance, as in Figure 4.29. In addition, both the pump and the hydraulic motor are positive displacement machines, represented as in Figure 4.32. The hydraulic motor is driven by a dc electrical motor as represented by the gyrator in Figure 4.34.
The load that the hydraulic motor drives can be represented by a combination of a rotary mechanical inertia and a rotary resistance. Show a complete bond graph for the system after setting the atmospheric pressure to zero (using gage pressures).

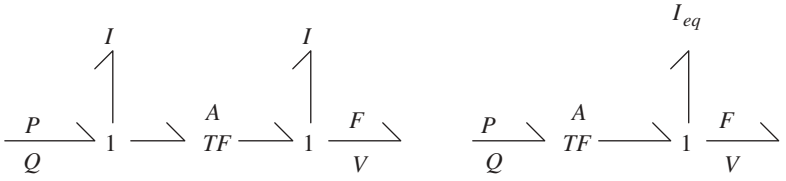
4-28. The figure shows an accumulator similar to the one in Figure 4.26. In this case, there is a distinction between the fluid pressure P_3 and the pressure of the gas in the rubber bladder under some conditions. Before the fluid system is pressurized, when P_3 is zero, the pressure in the gas is P_0 and the volume of the gas is V_0. The bladder has a rigid section at the base that prevents it from expanding into the inlet pipe. Until the fluid pressure exceeds P_0, no fluid enters the accumulator; $Q_3 = 0$. However, after some fluid has entered the accumulator, the gas is compressed and the gas pressure and the fluid pressure are essentially identical.
Consider the volume of fluid that has entered the accumulator, $V_3 = \int_0^t Q_3\,dt$, to be zero at $t = 0$ when the system is unpressurized. Make

a sketch of the nonlinear capacitance relationship between P_3 and V_3 using the isentropic assumption discussed in the text. Note that the P_3 can rise to P_0 even when $V_0 = 0$ and that $P_3 \to \infty$ when $V_3 \to V_0$.



4-29. The figure shows a pipe of area a and length l, containing hydraulic fluid, connected to a hydraulic ram of area A. This combination will eventually be used to model a hydraulic system, and the concern is whether the pipe hydraulic inertia, $\rho l/a$, is significant in comparison with the mass, m, of the ram and its connected load mass.





The bond graph on the right shows the two inertia elements separately, and the one on the left shows the two inertias combined into a single equivalent inertia associated with the velocity of the ram, V. One way to compute the equivalent inertia is to express the kinetic energy $T = mV^2/2 + IQ^2/2$ (where $I = \rho l/a$) in terms of the mass velocity by recognizing that the mass velocity and the volume flow rate are related by the transformer law $Q = AV$.

(a) Give an expression for the equivalent inertia (or equivalent mass) of the ram and pipe combination.

(b) Perform a sample calculation to find out what radius the pipe should have to make the fluid inertia of the pipe have the same effect on the effective mass of the system as the mass of the ram and load mass. Use the following parameters:

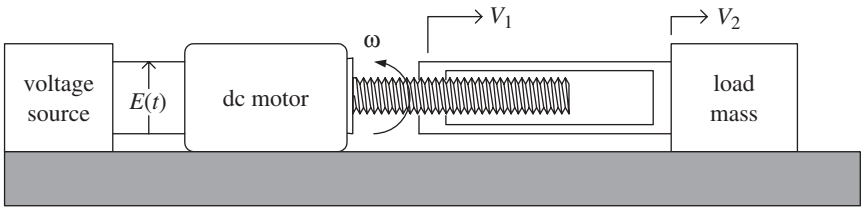
$$\rho = 900\ [\text{kg/m}^3], \quad l = 0.5\ [\text{m}], \quad m = 740\ [\text{kg}],$$
$$A = \pi R^2, \quad R = 50\ [\text{mm}], \quad a = \pi r^2$$

Answers:

(a) $I_{\text{eq}} = (m + \rho l A^2/a)$,

(b) $a = \rho l A^2/m, r = 15$ [mm].

4-30. The figure shows the physical part of a load-positioning system without the sensors and the controller electronics that control the voltage source. You are to make a bond graph model of the system, including the following effects:



(a) The voltage is a controlled effort source. Indicate this by including a signal or active bond with a double arrow impinging on an effort source.

(b) The dc motor has inductance and resistance in the armature coils.

(c) Include a rotary inertia for the motor armature and the screw drive.

(d) Use a resistor to model the friction moment associated with the motor and the screw drive.

(e) A constant S relates the angular velocity of the screw with the nut: $V_1 = S\omega$.

(f) The tube connecting the nut with the load mass is elastic with a spring constant k.

(g) The load mass has significant friction with the ground.

5

STATE-SPACE EQUATIONS AND AUTOMATED SIMULATION

In Chapter 4, bond graph models were developed for many physical engineering systems. We first treated each energy domain separately, but our goal was always to connect the energy domains into overall system models using the same symbols and structure for the entire system. This was done in Section 4.4 and will continue to be done for the rest of this text. The nine basic bond graph elements are capable of representing a very large cross section of engineering systems.

One reason for constructing bond graph models will become apparent in this chapter. It will be demonstrated just how straightforward it is to derive system equations and to obtain computer solutions to these equations. In some cases we can go from a bond graph directly to a simulation using a computer program without first obtaining equations of motion manually. This is particularly important when part of the model involves nonlinear constitutive laws since analytical results are rarely possible for such models.

In other cases, when linearized equations of motion are appropriate, we can use them to obtain analytical information about the system behavior. The real virtue of bond graph modeling becomes overwhelmingly apparent when one considers that, from a bond graph point of view, all systems models use the same symbols. Thus, only one formulation and solution procedure is needed for any of the systems we model.

To emphasize the virtue of having a formulation procedure that works for any bond graph, consider the physical system shown schematically in Figure 5.1*a* for part of an electric power steering system. A dc motor with winding resistance, R_w, rotor inertia, J_m, and rotary damping, b_τ, drives a flexible output shaft of rotary stiffness, k_τ. The shaft is connected to a rack and pinion setup where the gear has contact radius, R, and the rack has mass, m. The rack is attached to a spring and damper with parameters k and b. We desire to predict the motion of this system for a voltage input to the motor.

System Dynamics: Modeling, Simulation, and Control of Mechatronic Systems, Fifth Edition
Dean C. Karnopp, Donald L. Margolis and Ronald C. Rosenberg
Copyright © 2012 John Wiley & Sons, Inc.

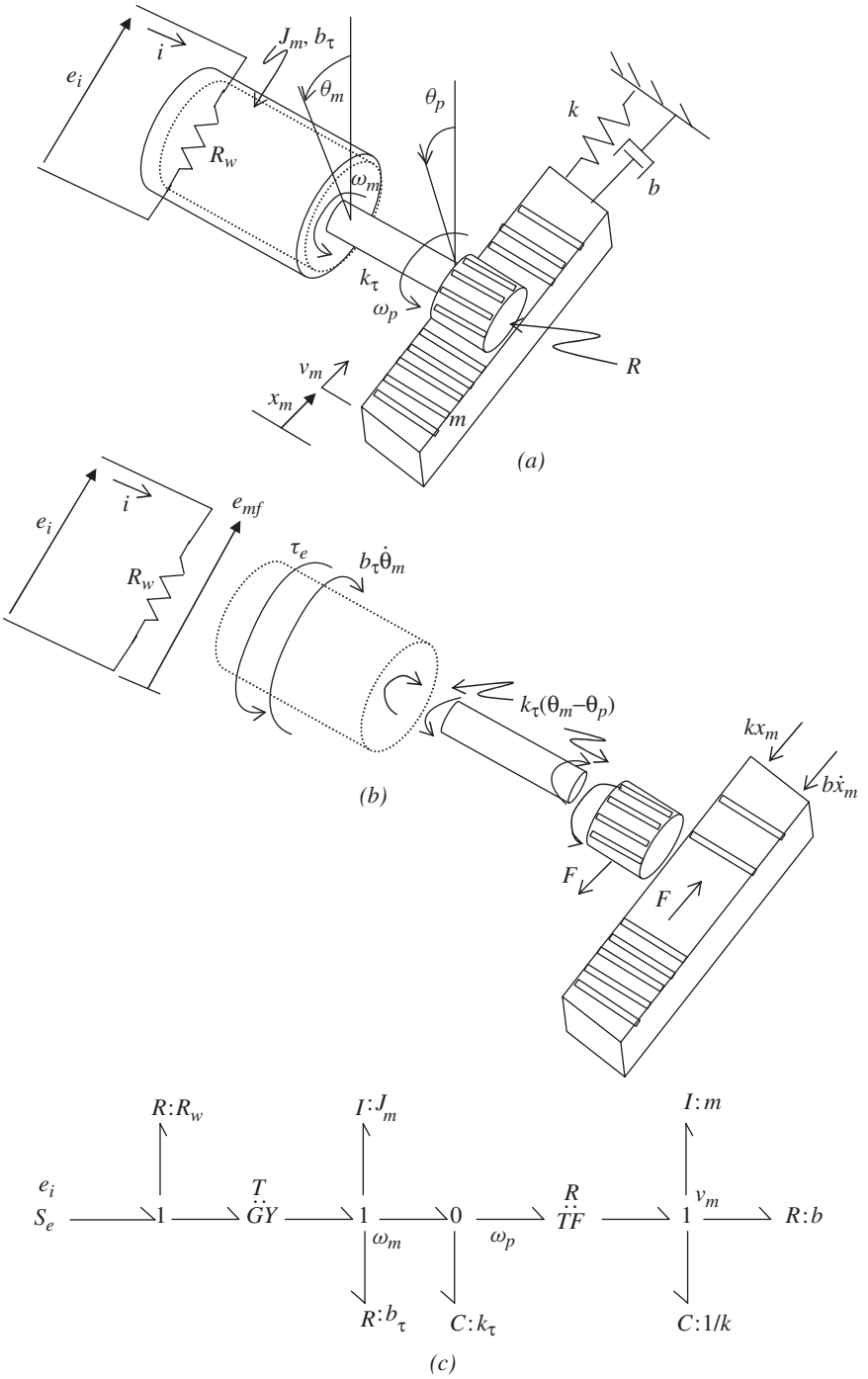


FIGURE 5.1. Part of an electric power steering system.

Without a procedure for deriving equations of motion, one might start by drawing a free-body diagram of the entire system. This is done in Figure 5.1*b*. The electric circuit voltage, e_{mf}, is the back emf of the motor. The electrical torque, τ_e, is the torque on the rotor due to the electromechanical coupling of the motor. The torque in the output shaft is shown acting on the rotor and the gear. And the internal force, F, is shown acting on the gear and the rack. With bond graphs, we always define efforts and flows, regardless of the energy domain involved. This is done in the system schematic diagram. The more conventional approach to formulation often starts with displacement variables, and these are also indicated in the schematic.

The next step is to use Kirchhoff's law for the circuit, with the result

$$e_i - R_w i - e_{mf} = 0. \tag{5.1}$$

Then use Newton's laws for translation and rotation, with the result

$$\tau_e - b_\tau \dot{\theta}_m - k_\tau(\theta_m - \theta_p) = J_m \ddot{\theta}_m, \tag{5.2}$$

$$F - kx - b\dot{x} = m\ddot{x}. \tag{5.3}$$

We must recognize the kinematic, no-slip constraint between the pinion gear and the rack,

$$\dot{x} = R\dot{\theta}_p, \tag{5.4}$$

and the moment equilibrium for the inertia-free pinion gear,

$$k_\tau(\theta_m - \theta_p) = FR. \tag{5.5}$$

At this point it might be useful to count the unknowns and see how we are doing. The unknowns are $i, e_{mf}, \tau_e, \theta_m, \theta_p, x,$ and F. Thus, we have seven unknowns and five equations. We are not done yet. The missing information is the electromechanical physics relating the electrical torque to the current and the back emf to the motor speed. These relationships are

$$\tau_e = Ti, e_{mf} = T\dot{\theta}_m, \tag{5.6}$$

where T is the coupling constant for the motor. We now have equal numbers of equations and unknowns and can theoretically proceed to a solution.

If we were to pursue this example further, we would next have to figure out substitutions to make to eliminate unwanted variables while retaining desired variables. There is no set procedure for this substitution; we just use our intuition and start substituting. If we are successful at coming up with a computable set of equations, the act of doing so will not benefit us much the next time we are confronted with a new physical system.

With bond graphs there is a definite procedure for choosing the proper variables. There is also a procedure for deriving a computable set of equations. It is the same procedure each time equations are derived. If one struggles the first time, it will be a little easier the second time, and easier each successive time. Figure 5.1*c* is a bond graph of the example system. All the dynamic elements in the schematic diagram are included in the bond graph. (Problem 5-16 revisits this system.) You will be surprised to learn how easy it is to derive the governing equations for system models using bond graph techniques compared with the unstructured conventional approach illustrated above. The formulation procedure is presented next.

5.1 STANDARD FORM FOR SYSTEM EQUATIONS

One of the most remarkable features of bond graphs is that a study of equation formulation may be carried out prior to writing any equations. To understand how this can be so, we first consider some particular forms of equations that are used to represent a system, and select one form—the state-space type—as our primary goal. All the models of dynamic systems discussed in this text involve ordinary differential equations. (Some models do represent approximations for the partial differential equations of distributed systems such as transmission lines.)

There are two limit forms for systems of differential equations, plus a wide range of possibilities within these limits. An nth-order system may be represented by

1. A single nth-order differential equation in terms of one unknown variable
2. n first-order coupled differential equations in terms of n unknown variables, or
3. Various combinations of unknowns and equations of appropriate orders (not necessarily equal)

Many important mathematical problems, methods, and results are organized in terms of the first form. At first, almost all engineering mathematics was cast in that form. The second form has certain advantages that have recommended it to mathematicians* for development of theory, but, even more important from our point of view, it is a convenient form for use by engineering system analysts, control engineers, and people conducting digital and analog computer studies. An interesting example of the third form is to be found in the sets of second-order equations generated by the Lagrangian approach to system analysis.

To illustrate each of the forms, and to give an appreciation of some of the differences, let us consider the mechanical double-oscillator example shown in

*See, for example, G. D. Birkhoff, *Dynamical Systems*, Providence, RI: American Mathematical Society, 1966.

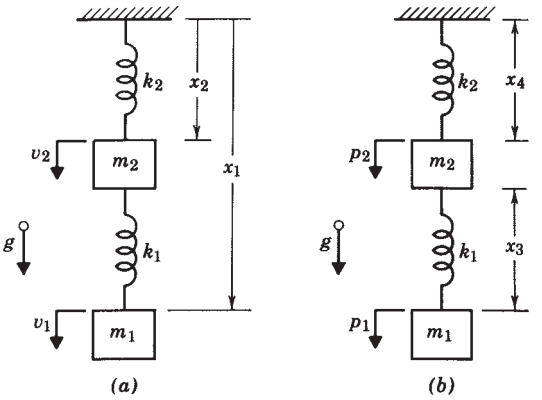


FIGURE 5.2. Mechanical double-oscillator example.

Figure 5.2. There are several choices of unknowns available. First, let us choose the position x_2 and describe the system in terms of that single displacement. The system equation in terms of x_2 turns out to be

$$\dddot{x}_2 + \left[\frac{k_2}{m_2} + k_1\left(\frac{1}{m_1} + \frac{1}{m_1}\right)\right]\ddot{x}_2 + \frac{k_1 k_2}{m_1 m_2}x_2 = k_1\left(\frac{1}{m_1} + \frac{1}{m_2}\right)g. \tag{5.7}$$

If a Lagrange equation approach is used, with x_3 and x_4 (the spring extensions) shown in Figure 5.1*b* as the generalized coordinates, the following pair of coupled second-order equations may be developed:

$$\begin{aligned}\ddot{x}_3 + k_1\left(\frac{1}{m_2} + \frac{1}{m_1}\right)x_3 - \frac{k_2}{m_2}x_4 &= 0,\\ \ddot{x}_4 - \frac{k_1}{m_2}x_3 + \frac{k_2}{m_2}x_4 &= g.\end{aligned} \tag{5.8}$$

Finally, if one concentrates closely on the energy in the system, associating a momentum or displacement variable with each distinct energy element, a set of four coupled first-order equations in terms of p_1, p_2, x_3, and x_4 may be found, as shown in Eq. (5.9). Clearly, this set could also be converted to a geometric-variable form by replacing the momentum variables by the velocities to which they are related:

$$\begin{aligned}\dot{p}_1 &= -k_1 x_3 + m_1 g,\\ \dot{p}_2 &= k_1 x_3 - k_2 x_4 + m_2 g,\\ \dot{x}_3 &= \frac{p_1}{m_1} - \frac{p_2}{m_2},\\ \dot{x}_4 &= \frac{p_2}{m_2}.\end{aligned} \tag{5.9}$$

It is possible, in principle, to transform from one of the given forms to any other, whether the system is linear or nonlinear. However, for nonlinear systems the desired transformations may be very difficult to discover. Furthermore, if one starts with the form (5.7), for example, choosing additional unknowns is a rather haphazard process unless one has considerable insight into the system being studied. Yet if a physically meaningful set of variables is already available (e.g., as displayed by (5.9)), elimination of unwanted variables can be carried out with considerable insight.

In studying engineering systems using bond graphs, there is an ideal opportunity to start the formulation in terms of significant physical variables and to generate simultaneous sets of first-order equations from the bond graphs. This is the approach we shall follow. When the system being studied is nonlinear, the form we seek is given by

$$\begin{aligned}
\dot{x}_1(t) &= \phi_1(x_1, x_2, \ldots, x_n; \quad u_1, u_2, \ldots, u_r),\\
\dot{x}_2(t) &= \phi_2(x_1, x_2, \ldots, x_n; \quad u_1, u_2, \ldots, u_r),\\
&\vdots\\
\dot{x}_n(t) &= \phi_n(x_1, x_2, \ldots, x_n; \quad u_1, u_2, \ldots, u_r),
\end{aligned} \tag{5.10}$$

where the x_i are the *state* variables, the $\dot{x}_i$ are the time derivatives of the x_i, the u_i are inputs to the system, and the ϕ_i are a set of static (or algebraic) functions.*

If the system is linear, Eqs. (5.10) take on a simpler form:

$$\begin{aligned}
\dot{x}_1(t) &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_{11}u_1 + b_{12}u_2 + \cdots + b_{1r}u_r,\\
\dot{x}_2(t) &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_{21}u_1 + b_{22}u_2 + \cdots + b_{2r}u_r,\\
&\vdots\\
\dot{x}_n(t) &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_{n1}u_1 + b_{n2}u_2 + \cdots + b_{nr}u_r,
\end{aligned} \tag{5.11}$$

where the a_{ij} and b_{ij} are constants in most cases. The example Eqs. (5.9) are of the type of Eq. (5.11).

For linear systems, equations of the form (5.11) can be put into a standard matrix formulation,

$$\dot{X} = AX + BU, \tag{5.11a}$$

where X is the vector of state variables,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \bullet \\ \bullet \\ x_n \end{bmatrix},$$

*This simply means that, given the values of the arguments on the right-hand side of Eq. (5.4), a set of values for the derivatives may be found by algebraic means.

and U is the vector of inputs,

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \bullet \\ \bullet \\ u_r \end{bmatrix},$$

The A-matrix is composed of the coefficients, a_{ij}, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \bullet & \bullet & a_{1n} \\ a_{21} & \bullet & \bullet & \bullet & a_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{n1} & a_{n2} & \bullet & \bullet & a_{nn} \end{bmatrix},$$

and the B-matrix contains the coefficients, b_{ij}, where

$$B = \begin{bmatrix} b_{11} & b_{12} & \bullet & \bullet & b_{1r} \\ b_{21} & \bullet & \bullet & \bullet & b_{2r} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ b_{n1} & b_{n2} & \bullet & \bullet & b_{nr} \end{bmatrix}.$$

We will discover how useful this formulation is when the system being studied can be modeled as a linear one.

Our task for the rest of this chapter is to learn how to select significant system variables from among the many effort, flow, momentum, and displacement variables contained in a bond graph and to organize them into relations of the form (5.10) or, for linear models, the matrix form (5.11a).

5.2 AUGMENTING THE BOND GRAPH

In the previous chapters we have developed an orderly procedure for representing all kinds of interacting physical systems using bond graphs. At this point we assume that a bond graph has been constructed and that an appropriate power convention has been assigned. Before writing any equations, it is desirable to prepare the bond graph with additional information that will make writing of equations take on a very orderly pattern. Providing the additional information consists of

1. Numbering all the bonds in consecutive order
2. Assigning to each bond a causal sense for the e and f variables

By numbering the bonds, every element and variable can be referred to unambiguously. Thus, I_4 is the inertia on bond 4, e_6 is the effort on bond 6, q_{11} is

the displacement on bond 11, and so forth. For beginners, it is suggested that after the physical system has been modeled and a bond graph exists with physically motivated labels (e.g., mass, inductance, force, volume flow, pressure or piston area symbols) a second, identical bond graph be drawn without the physical labels but with bond numbers. We call this bond graph the *computational bond graph* since it is easily processed by computers as well as humans. Since bond graphs for all physical systems are so similar, it is also suggested that the physically notated bond graphs be kept nearby so that the user can keep track of the physical effects involved in the numbered bond graph.

The second step in the augmentation process is the *assignment of causality*. Basic considerations of causality for the various elements were introduced earlier (Section 3.4). It is now appropriate to apply such information to the entire system in an orderly fashion.

In a causal sense there are two distinct types of bond graph elements. Source elements (S_e and S_f) and junction elements (0, 1, TF, and GY) must meet certain causal conditions or their basic definitions are no longer valid. As an obvious example, if a source of force does not have a causality showing that it applies a defined force on the system to which it is connected, then it has no meaning. We generalize this observation and assert that *every source element must have its appropriate causal form assigned to it*.

For each of the 2-port junction elements, TF and GY, there are two possible causal forms that preserve the basic definition of the element. If neither of these forms can be assigned, the concept of input and output associated with the element is not valid. Consequently, we assert that *every TF and GY must have one of its two allowable causal forms assigned to it*. The choice of form will generally be indicated by the adjoining system on the basis of other considerations to be discussed.

In a similar vein, we argue with respect to the ideal multiport junctions 0 and 1 that *each* 0 *and* 1 *must have one of its appropriate causal forms assigned to it* or the basic definition of the particular element will not be valid. The selection of a particular causal form will be motivated by other system considerations in general.

If a system cannot meet the causal conditions outlined above, the basic physical model on which the bond graph is based must be restudied. The indications are that an impossible situation has been created that is not capable of sensible mathematical resolution. Two such examples are shown in Figure 5.3. In part *a*, source element 2 is defined invalidly. Inspection of an electrical circuit interpretation indicates the nature of the difficulty: obviously, a modeling error has been made in putting two supposedly *independent* current sources in series. (Each source seeks to impose its own current but the two must be equal if they act in series.) In part *c* of Figure 5.3, a TF is found to have invalid causality. If this bond graph were derived from a fluid circuit, as shown in part *d*, the interpretation would be that two *independent* pressure sources had been joined by an ideal transformer of pressure (the TF), leading to a physically incompatible

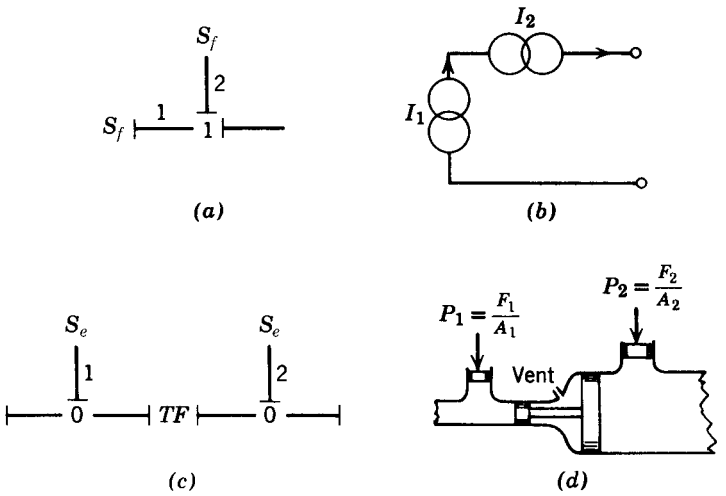


FIGURE 5.3. Two examples of invalid causality and physical interpretations.

situation. The next step is up to the system modeler, who must correct the model in an appropriate way.

Continuing our discussion of assigning causality, we come to the energy storage elements, —I and —C. Bond graphs "bookkeep" energy, and the instantaneous energy of the system is indicated by the energy variables (p's on —I's and q's on —C's) associated with the energy storage elements. The junction structure elements,

$$—\underset{|}{0}—,\quad —\underset{|}{1}—,\quad —TF—,\quad —GY—,$$

merely shuttle the power around between the source elements, the energy storage elements, and the dissipative elements. We choose to use the energy variables as the *state variables* for the systems we model. Knowledge of the energy variables dictates the energy state of the system at each instant of time. As we shall see, an energy storage element is independent if it can accept integral causality. Thus, in our causal assignments, we make every effort to assign integral causality to each energy storage element. We will see that this is not always possible. If an energy storage element is forced to accept derivative causality, then that element is not independent, and its energy variable is algebraically related to the other energy variables in the system. The energy storage element in derivative causality still stores energy, but its contribution to the system energy can be calculated algebraically from knowledge of the other energy variables, and, if the energy variable of the derivative element should happen to enter the equation formulation, it can be eliminated from the final state equations. Thus, for the formulation procedure proposed here, *the state variables are the p-variables on I-elements and the q-variables on C-elements that are in integral causality*.

At the system level the causality associated with R-elements is largely a matter of indifference. The major exception is in the case of a nonlinear constitutive law that is not bi-unique (e.g., Coulomb friction). Then the causality associated with a unique input–output relation for the element should be used. Otherwise, R-elements accept whatever causality they are assigned by other elements in the system.

The basic causality assignment procedure is summarized as follows:

Sequential Causal Assignment Procedure

1. Choose any source (S_e, S_f), and assign its required causality. Immediately extend the causal implications through the graph as far as possible, using the elements (0, 1, GY, TF) which may be connected to the source. These four elements called *constraint* elements because only certain causal patterns are allowed.
2. Repeat step 1 until all sources have been assigned their required causality.
3. Choose any storage element (C or I), and assign its preferred (integral) causality. Immediately extend the causal implications through the graph as far as possible, using the constraint elements (0, 1, GY, TF).
4. Repeat step 3 until all storage elements have been assigned a causality.

In many practical cases all bonds will be causally oriented after this stage. In some cases, however, certain bonds will not yet have been assigned. We then complete the causal assignment as follows:

5. Choose any unassigned R-element and assign a causality to it (basically arbitrary). Immediately extend the causal implications through the graph as far as possible, using the constraint elements (0, 1, GY, TF).
6. Repeat step 5 until all R-elements have been used.
7. Choose any remaining unassigned bond (joined to two constraint elements), and assign a causality to it arbitrarily. Immediately extend the causal implications through the graph as far as possible, using the constraint elements (0, 1, GY, TF).
8. Repeat step 7 until all remaining bonds have been assigned.

The procedure is straightforward and orderly. Some practice on examples will convince you of the ease and rapidity with which causality can be assigned. It is important to recognize that the constraint elements represent the physical structure in the system (e.g., Kirchhoff's voltage and current laws, Newton's law and geometric compatibility), and assigning causality to them means that they will be used correctly in a particular input–output fashion. Further discussion of the use and interpretation of causality in more complex bond graphs is given in Chapter 7.

Several situations can arise when applying causality according to the procedure given:

1. All storage elements have integral causality, and the graph is complete after step 4. This simple, common case is discussed in Section 5.3.
2. Causality is completed by using R-elements or bonds, as indicated in steps 5–8. This situation is discussed in Section 5.4.
3. Some storage elements are forced into derivative causality at step 3. This case is discussed in Section 5.5.

Now let us consider several examples of using the procedure.

Assignment of causality is carried out step by step in the example in Figure 5.4. Part *a* shows a numbered bond graph without causality or power convention. We would normally have assigned a power convention by the time we arrived at causality assignment. However, to enforce upon the reader that power convention and causality assignment are totally unrelated, we work a few examples in which the reader is not distracted by a power sign convention.

A graph without causality is sometimes said to be *acausal*. In part *b* of Figure 5.4 bond 1 is directed according to the meaning of the source element (a source of flow). Since the 0-junction connected to the source has only one flow variable determined, f_1, the other bonds cannot yet be assigned a causality. (The

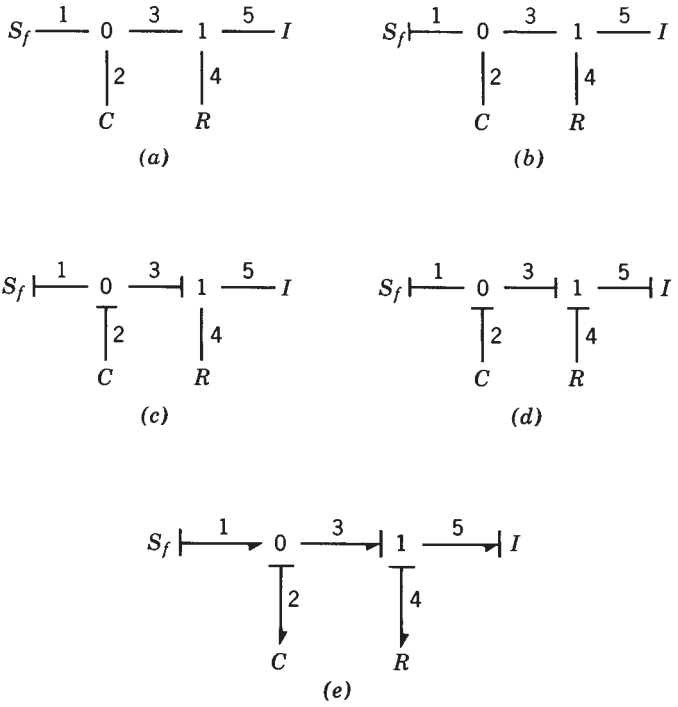


FIGURE 5.4. Causality assignment and complete augmentation of a bond graph. Example 1.

3-port 0-junction $0_{1,2,3}$ will have two flow inputs and one effort input. We now have determined only one of the two flow inputs f_1.)

There are no more sources, so we turn to step 3. In part c, C_2 is put is put into integral causality. Immediately, bond 3 must be causally directed because of the 0-junction, which can have only one effort input, e_2. However, bonds 4 and 5 are as yet undirected. (The 3-port 1-junction $1_{3,4,5}$ will have two effort inputs but only one flow input. At this point we have determined only one of the effort inputs, e_3.)

Now I_5 is assigned integral causality as shown in Figure 5.4d. Immediately, bond 4 must be causally directed as shown, due to the 1-junction, $1_{3,4,5}$, which can have only one flow input, f_5. At this point, all bonds have been causally oriented.

In part e we have added a set of power directions to the graph; the result is a completely augmented bond graph. That is, the bonds are labeled, power directions have been chosen, and causality has been assigned. Such a graph will yield its state equations to us with very little resistance, as we shall show in the next section. However, let us first augment another graph or two to gain some experience.

Figure 5.5 shows a bond graph derived from a fluid example involving pipes and reservoirs. There are three pressure sources (S_{e1}, S_{e2}, and S_{e3})

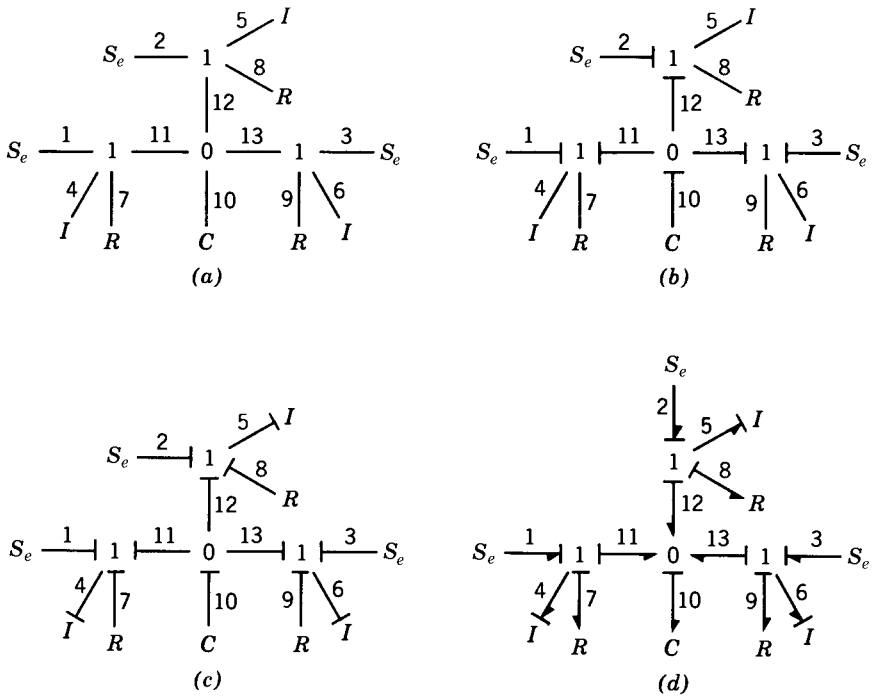


FIGURE 5.5. Augmentation of a bond graph. Example 2.

feeding through three pipes (roughly, the 1-junction complexes) into a tank (the 0-junction and C_{10}). In part a, the graph is labeled. In part b, bonds 1, 2, and 3 associated with the sources have been assigned causality. In each case no extension of causality was possible. In addition, bond 10 has been directed according to integral causality for the element C_{10} and causality has been extended to bonds 11, 12, and 13 by the 0-junction, because the common effort has been determined by the e_{10} input.

In part c, I_4 has been given integral causality which determines the causality on bond 7 due to the 1-junction. The process continues with the assignment of integral causality to I_5 and I_6 which determine the causality of bonds 8 and 9. Finally, power directions have been added to produce the fully augmented graph shown in part d. In the next section, we shall show how this graph yields four first-order equations with three inputs to the system.

As a final example in this section, consider the acausal bond graph in Figure 5.6a. The model is derived from a study of a pressure-controlled valve and includes both mechanical and fluid mechanical power. The elements S_{e1} and S_{e2} represent sources of pressure and force, respectively, and the element TF couples the two power domains. In part b, bonds 1 and 2 are causally assigned, one at a time. The causal information cannot be extended using constraint elements at this point. Bonds 3 and 4 are assigned next, and causality is extended as shown in part c. Bond 4 causality does not extend to

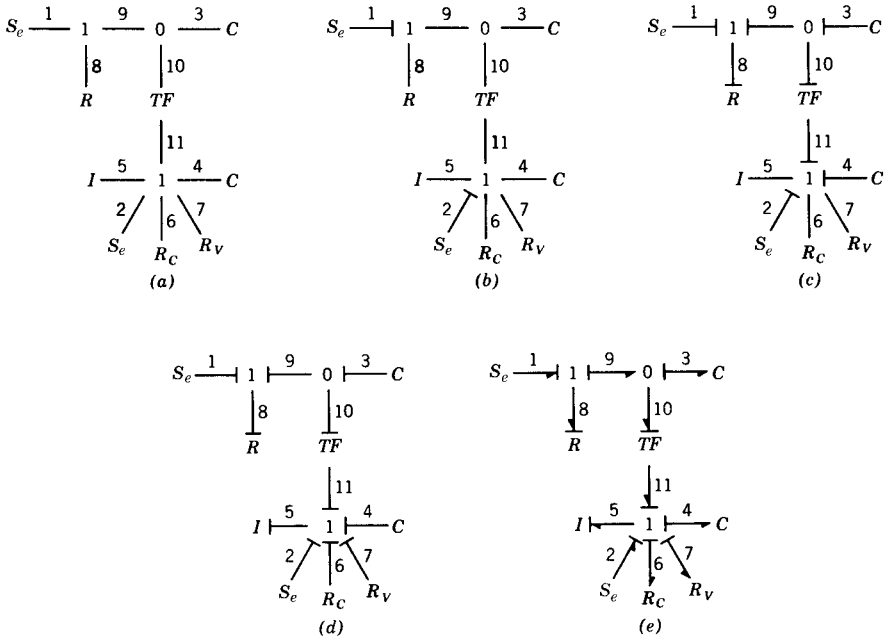


FIGURE 5.6. Augmentation of a bond graph. Example 3.

other bonds, but bond 3 has implications for bonds 8 and 9, as well as for bonds 10 and 11. Inspection of Figure 5.6*d* reveals that assigning causality to bond 5 (associated with the element I_5) determines the causality on bonds 6 and 7. With powers assigned as shown in part *e*, the bond graph is completely augmented.

Two further points may be made with respect to augmentation. The first is that, in assigning causality, the results do not depend on the order of bonds chosen except in special circumstances to be discussed later in Section 5.4. The second point is that assignment of causality and assignment of power directions are two entirely independent operations. Either one may be performed first. Typically, power directions will be first, but on occasion one may not bother to assign powers in studying aspects of system equation formulation.

5.3 BASIC FORMULATION AND REDUCTION

Once a fully augmented bond graph model is available, the equations for the system can be developed in a very orderly fashion. Frequently, when the system is small or uncomplicated in structure, state-space equations can be written down directly. However, as system size and complexity grow, the need for an organized procedure for equation generation becomes apparent.

Very general and powerful procedures are available for producing sets of system equations. In this section we concentrate on a basic pattern that is applicable in a large majority of cases encountered in engineering practice. There are three simple steps to be followed:

1. Select input and energy state variables,
2. Formulate the initial set of system equations,
3. Reduce the initial equations to state-space form.

Selection of *inputs* is straightforward. For each source element write on the graph the input variable to the system. These variables will appear in the final state-space equations if they have any effect on system behavior. The list of input variables will be called U.

As mentioned earlier, the state variables are the p-variables on the I-elements and the q-variables on the C-elements *that are in integral causality*. For many system models, we will discover that after assigning causality sequentially by the rules of the previous section, all energy storage elements are in integral causality and there are no unassigned bonds. It is this case that is discussed in this section on basic formulation procedures. When this causal pattern emerges, then explicit first-order differential equations of the type indicated in Eqs. (5.10) for nonlinear systems or Eq. (5.11a) for linear systems result, and equation derivation is straightforward.

The state variables can be placed in a vector called the state vector, X. The derivative of the state variables will always be either an effort or a flow variable, since $\dot{p} = e$ and $\dot{q} = f$. On the computational bond graph it is convenient to

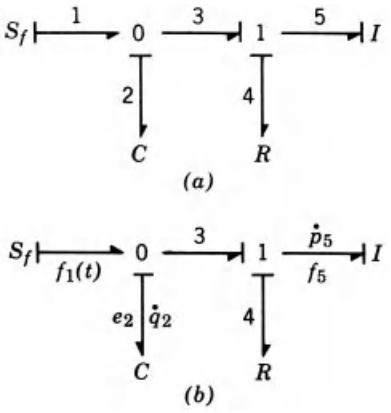


FIGURE 5.7. Equation formulation. Example 1.

show $\dot{p}$'s as the efforts on the I-elements in integral causality and $\dot{q}$'s as the flows on C-elements in integral causality. In addition, it will prove useful to indicate the *co-energy variables* on the energy storage elements. These are the flow variables on the I-elements and the effort variables on the C-elements. We will discover that the co-energy variables will appear in the initial formulation and then will be eliminated during the reduction process. As a first example of basic formulation, consider the augmented bond graph in Figure 5.7*a*, which is repeated from Figure 5.4*e*. The causality has been assigned, and all energy storage elements are in integral causality, all bonds have a causal assignment, and there are no causal conflicts on the bonds of the constraint elements. The state variables are q_2 and p_5. The formulation always starts by writing the rate of change of a state variable equals an effort or a flow. We then use causality to track the efforts and flow through the bond graph. Causality will show us the substitutions to make to eliminate unwanted variables while retaining appropriate ones, as the procedure leads us to explicit first-order equations. It makes no difference which state variable we start with. We will later learn that computer programs exist that can derive the equations following the same rules we use. These programs number the bonds consecutively and derive equations starting with the state variable with the lowest bond number. We will do the same here. Thus, starting with q_2, we write,

$$\dot{q}_2 = f_2, \tag{5.12}$$

and follow the causality of the model to direct us through the bond graph and lead us directly to the state equations. From Figure 5.7 we see that f_2 is an *input to* the C-element and an *output from* the 0-junction. This causal output is caused by the inputs, $f_1(t)$ and f_3, and according to the power convention,

$$\dot{q}_2 = f_1(t) - f_3. \tag{5.13}$$

The flow $f_1(t)$ is the input flow from the flow source, and this variable is known and remains in the formulation. (The notation $f_1(t)$ may be used to remind us that f_1 is a *known* input function of time.) The flow, f_3, is not wanted in the formulation, so we must look further.

We just used f_3 as an input to the 0-junction, but it is also an output from the 1-junction. It is true that f_3 equals both f_4 and f_5, but it is *caused by* f_5 (look at the causality on the 1-junction). Thus, we write

$$\dot{q}_2 = f_1(t) - f_5. \tag{5.14}$$

The flow, f_5, is a co-energy variable and directly related to the state variable, p_5, through the constitutive relationship for the element. If I_5 is a nonlinear element, then

$$f_5 = \Phi_I^{-1}(p_5), \tag{5.15}$$

and, if I_5 is a linear element, as it is for this example, then

$$f_5 = \frac{p_5}{I_5}. \tag{5.16}$$

With this substitution we obtain the first state equation as

$$\dot{q}_2 = f_1(t) - \frac{p_5}{I_5}. \tag{5.17}$$

The second state equation starts with

$$\dot{p}_5 = e_5, \tag{5.18}$$

and we need to follow the causal path and track down e_5. From Figure 5.7 we see that e_5 is an *input to* the I-element and an *output from* the 1-junction. The 1-junction indicates that e_5 is *caused by* e_3 and e_4, and according to the power convention,

$$e_5 = e_3 - e_4; \tag{5.19}$$

thus,

$$\dot{p}_5 = e_3 - e_4. \tag{5.20}$$

Neither of these efforts is wanted in the final formulation, so we must continue to use causality to determine what substitutions to make. The effort e_3 is an output from the 0-junction and is *caused by* e_2. But e_2 is a co-energy variable and is directly related to the state variable, q_2, by the constitutive relationship for the C-element,

$$e_2 = \Phi_C^{-1}(q_2), \tag{5.21}$$

or, for the linear element used in this example,

$$e_2 = \frac{q_2}{C_2}. \tag{5.22}$$

This substitution will be made in (5.20), but first let's track down e_4. The effort, e_4, is an output from a resistance element where, in this causality,

$$e_4 = \Phi_R(f_4), \tag{5.23}$$

or, for the linear assumption being used for this example,

$$e_4 = R_4 f_4. \tag{5.24}$$

Eq. (5.24) is used in (5.20) to eliminate e_4, but we introduce, f_4. This flow is a *causal output* from the 1-junction where the causal flow input is the co-energy variable, f_5, which is directly related to the state variable p_5, as indicated in (5.16). Using the substitutions from (5.22), (5.24), and (5.16) in (5.20), we obtain the second state equation as

$$\dot{p}_5 = (1/C_2)q_2 - (R_4/I_5)p_5. \tag{5.25}$$

Equations (5.17) and (5.25) are the two state equations for this system. Since this example is linear, the final step is to place these equations into the standard matrix form of (5.11),

$$\frac{d}{dt}\begin{bmatrix} q_2 \\ p_5 \end{bmatrix} = \begin{bmatrix} 0 & -1/I_5 \\ 1/C_2 & -R_4/I_5 \end{bmatrix}\begin{bmatrix} q_2 \\ p_5 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} f_1(t). \tag{5.26}$$

This is the perfect starting point for analysis if the system is linear. One need only follow the causality, and the substitutions needed to eliminate unwanted variables while retaining state and source variables automatically occur. The reader may not believe this yet, so we present a second example.

Figure 5.8 shows an augmented bond graph that originally came from an electrical circuit. When causality was assigned, all energy storage elements ended up in integral causality, and all bonds had a causal assignment. This is always an indication that formulation of equations is straightforward. The state variables for this example are the momentum variable on bond 2, p_2, and the displacement variable on bond 5, q_5 and input variable is $e_i(t)$. If the physical bond graph were nearby we would realize that the momentum variable is the flux linkage for the inductor of L_2 henrys and the displacement variable is the charge on the capacitor of C_5 farads. The derivatives of the state variables are indicated on the energy storage elements, as are the associated co-energy variables, f_2 and e_5.

The formulation starts by writing

$$\dot{p}_2 = e_2 \tag{5.27}$$

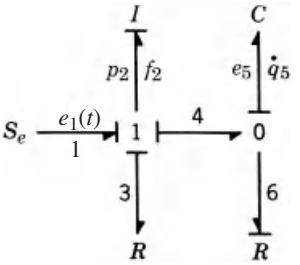


FIGURE 5.8. Equation formulation. Example 2.

and then pursuing e_2 by using the causality of the bond graph. The effort, e_2, is an output from the 1-junction and is caused by $e_1, e_3,$ and e_4, according to the power convention; thus,

$$\dot{p}_2 = e_1(t) - e_3 - e_4. \tag{5.28}$$

The effort, $e_1(t)$, is a source variable and remains in the final equations. The effort e_4 is an output from the 0-junction and is caused by the causal input, e_5. But e_5 is a co-energy variable directly related to a state variable by

$$e_5 = \frac{q_5}{C_5}. \tag{5.29}$$

Before substituting into (5.28), notice that e_3 is an output from a resistance with causality dictating

$$e_3 = R_3 f_3, \tag{5.30}$$

and f_3 is caused by the co-energy variable, f_2, and f_2 is related directly to a state variable by

$$f_2 = (1/I_2)p_2. \tag{5.31}$$

Using (5.29), (5.30), and (5.31) in (5.28) yields the state equation

$$\dot{p}_2 = e_1(t) - (R_3/I_2)p_2 - (1/C_5)q_5. \tag{5.32}$$

The second state equation starts with

$$\dot{q}_5 = f_5, \tag{5.33}$$

followed by using causality to track the path for f_5. The flow, f_5, is the input to C_5 and the output from the 0-junction. This output is caused by the causal inputs f_4 and f_6, according to the power convention; thus,

$$\dot{q}_5 = f_4 - f_6. \tag{5.34}$$

The flow f_4 is the output from the 1-junction, caused by the co-energy variable, f_2, where

$$f_2 = (1/I_2)p_2. \tag{5.35}$$

The flow f_6 is the causal output from a resistance, where causality indicates that

$$f_6 = \Phi_R^{-1}(e_6) = (1/R_6)e_6 \tag{5.36}$$

for this linear example. But e_6 is caused by the co-energy variable, e_5, which is directly related to a state variable according to (5.29). Using (5.35) and (5.36) in (5.34) yields the final state equation as

$$\dot{q}_5 = (1/I_2)p_2 - (1/R_6C_5)q_5. \tag{5.37}$$

The final step for linear systems is to put state equations (5.32) and (5.37) into the standard matrix format,

$$\frac{d}{dt}\begin{bmatrix} p_2 \\ q_5 \end{bmatrix} = \begin{bmatrix} -R_3/I_2 & -1/C_5 \\ 1/I_2 & -1/R_6C_5 \end{bmatrix}\begin{bmatrix} p_2 \\ q_5 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e_1(t). \tag{5.38}$$

Perhaps the reader is starting to see the power of using causality and the equation formulation procedure presented in this section. Any approach to equation derivation always requires elimination of some variables in favor of others in order to end up with a computable set of equations. Bond graphs and causality simply make the substitutions obvious so that appropriate variables are eliminated while others are retained, without having to guess the appropriate substitutions to make. One more example should be sufficient to demonstrate the consistency of our approach to the formulation of state equations.

Figure 5.9 is the augmented bond graph originally shown in Figure 5.6 for development of causality. The bond graph represents a physical system with hydraulic and mechanical parts. To keep track of what parts are hydraulic and what parts are mechanical, we really need to have the physical bond graph. For equation formulation, we need only have the augmented bond graph since equation formulation follows the same procedure regardless of the physical system. This example has a transformer that couples the hydraulic and mechanical sides of the model. The modulus of the transformer is physically an area, A. In Figure 5.9, the modulus, m, is defined by noting the constitutive relationships for the transformer. We will derive the equations in terms of the generic definition of the transformer shown in the figure, but before we can actually use the equations to predict system response, we will have to determine how the modulus, m, relates to the area, A. (A physical variable bond graph would tell us whether m as defined is actually A or $1/A$. How the transformer modulus is actually used depends on the causality that was determined for the transformer bonds.)

The causal assignment once again yielded all energy storage elements in integral causality and no bonds were left unassigned. The state variables are q_3, the

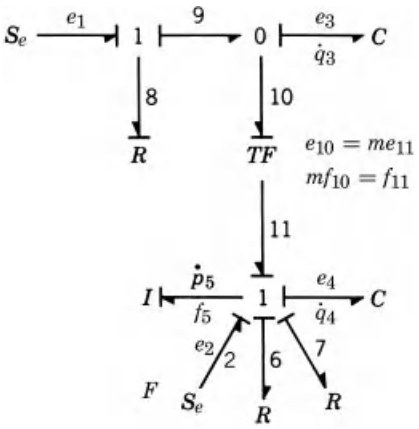


FIGURE 5.9. Equation formulation. Example 3.

displacement variable on C_3; q_4, the displacement variable on C_4; and p_5, the momentum variable on I_5. The input variables from the sources are e_1 and e_2. Also noted on the bond graph are the associated co-energy variables. Equation derivation starts with

$$\dot{q}_3 = f_3, \tag{5.39}$$

but

$$f_3 = f_9 - f_{10},$$

and

$$\begin{aligned} f_9 = f_8 &= (1/R_8)e_8 = (1/R_8)(e_1 - e_9) \\ &= (1/R_8)(e_1 - e_3) = (1/R_8)\left(e_1 - (1/C_3)q_3\right). \end{aligned} \tag{5.40}$$

The flow f_{10} is the output flow from the transformer, caused by the input flow, f_{11}, according to

$$f_{10} = (1/m)f_{11} = (1/m)f_5 = (1/mI_5)p_5. \tag{5.41}$$

Substituting (5.40) and (5.41) into (5.39) yields the first state equation as

$$\dot{q}_3 = (1/R_8)\left(e_1 - (1/C_3)q_3\right) - (1/mI_5)p_5. \tag{5.42}$$

The second state equation is particularly simple,

$$\dot{q}_4 = f_4 = f_5 = (1/I_5)p_5. \tag{5.43}$$

And the final state equation is

$$\dot{p}_5 = e_5 = e_2 - e_4 - e_6 - e_7 + e_{11}. \tag{5.44}$$

The effort e_2 is a source variable and remains in the final formulation. The effort e_4 is directly related to a state variable,

$$e_4 = (1/C_4)q_4. \tag{5.45}$$

The efforts e_6 and e_7 are outputs from resistances, where

$$\begin{aligned} e_6 &= R_6 f_6 = R_6 f_5 = (R_6/I_5)p_5 \\ e_7 &= R_7 f_7 = R_7 f_5 = (R_7/I_5)p_5 \end{aligned} \tag{5.46}$$

and e_{11} is the output effort from the transformer,

$$e_{11} = (1/m)e_{10} = (1/m)e_3 = (1/m)q_3/C_3. \tag{5.47}$$

Substituting (5.45), (5.46), and (5.47) into (5.44) yields the final state equation as

$$\dot{p}_5 = e_2 - (1/C_4)q_4 - [(R_6 + R_7)/I_5]p_5 + (1/mC_3)q_3. \tag{5.48}$$

Equations (5.42), (5.43), and (5.48) are a complete set of linear state equations and can be cast into the standard matrix format,

$$\frac{d}{dt}\begin{bmatrix} q_3 \\ q_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} -(1/R_8C_3) & 0 & -(1/mI_5) \\ 0 & 0 & 1/I_5 \\ 1/mC_3 & -1/C_4 & -(R_6 + R_7)/I_5 \end{bmatrix}\begin{bmatrix} q_3 \\ q_4 \\ p_5 \end{bmatrix} + \begin{bmatrix} 1/R_8 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}. \tag{5.49}$$

We hope that by now the reader is convinced that by choosing the state variables to be the p's on I's and the q's on C's in integral causality, and by following the input/output causal paths, the state equations are straightforwardly derived. If the system is linear, then the resulting equations can be put into a standard matrix format that is the perfect starting point for linear analysis, which is covered in the next chapter. If the system is nonlinear, equations are still derived easily by using bond graphs and causality, but there is no nice matrix representation from which to launch an analysis. Instead, computer simulation is required, which is covered later in this chapter and at an advanced level in Chapter 13.

Unfortunately, models do not always lend themselves to straightforward formulation. Sometimes modeling assumptions lead to perfectly good models that have mathematical formulation difficulties. Such problems are the topic of the next two sections.

5.4 EXTENDED FORMULATION METHODS—ALGEBRAIC LOOPS

Constructing models of physical systems always requires modeling assumptions. Any physical system can be infinitely complicated. It is we, as engineers, who must construct models that are simple enough to work with and yet retain sufficient dynamics to answer the questions being asked. Along the way we decide whether the mass of some component is important, whether a part is sufficiently stiff to be modeled as rigid, whether components can be represented as linear over some operating range, and so forth. The result of these assumptions can sometimes lead to mathematical formulation problems that have nothing to do with the quality of the model, yet they must be addressed in order to make use of the model. One of these formulation problems is caused by *algebraic loops.* This topic is discussed at an advanced level in Section 13.2 of Chapter 13. Here, the basic problem is introduced.

Consider the physical system in Figure 5.10*a*. It consists of a mass, m, acted upon by the input force, $F(t)$. The mass is attached to some dampers and a spring, as shown, and the spring has an input velocity at its right end, $v_i(t)$. The mass velocity is labeled v_m, and the attachment point of the dampers and spring is indicated as v'. The dampers and spring are assumed positive in tension, and mass at the attachment point of these elements, m', *is assumed to be negligible*. This model is perfectly reasonable, and yet it has a serious formulation problem that is vividly exposed with the use of bond graph causality.

Figure 5.10*b* shows the bond graph for this system constructed using the procedure for mechanical translation. Some liberty has been taken in that physical labels are on the bond graph along with numbering the bonds. Normally, we would have a physical bond graph and a separate computational bond graph. Also in Figure 5.10*b*, causality has been assigned up through step 4 of the sequential assignment procedure presented in Section 5.2.

From the sequential assignment of causality, the inertial element, I_2, has set the flow input to the 1-junction, constraining bond 3 to have flow output from the 1-junction. Since f_3 is a causal flow input to a 0-junction, there is no constraint to assign further causality. The compliance element, C_8, is in integral causality, and e_8 sets the effort on the 0-junction. This constrains the effort, e_7, to be an output from the 0-junction and an input to the 1-junction with bonds 5 and 6 unassigned. At this point there are no further causal assignments that we are constrained to make. Both of the energy storage elements are in integral causality, but there remain causally unassigned bonds.

Whenever causality has been assigned according to the procedure of Section 5.2 and one must proceed to step 5, because there are causally unassigned bonds remaining, state equations of the form (5.10) or (5.11) cannot be derived in the straightforward manner of the previous section. When this situation arises, it means that there is an *algebraic loop* where some effort or flow variables are algebraically related to themselves as well as to other source variables and state variables. We will demonstrate what happens if this condition is ignored.

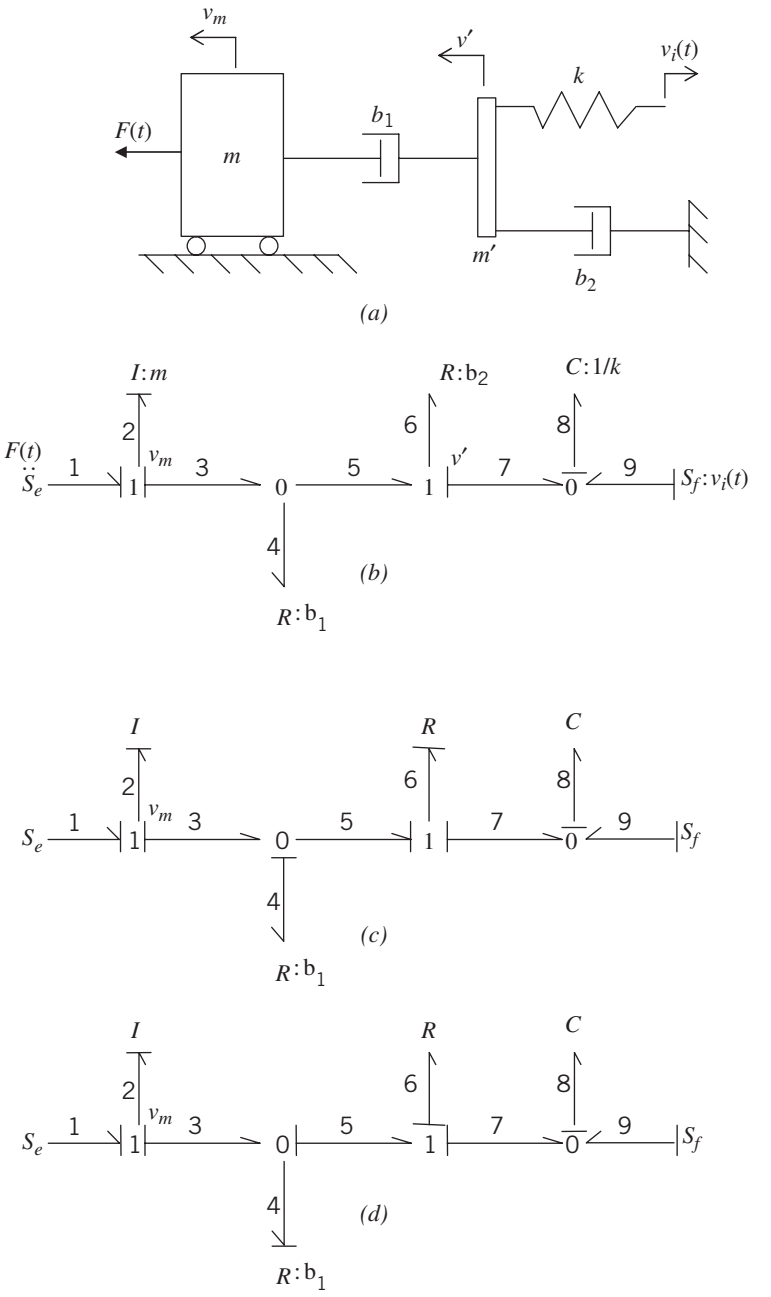


FIGURE 5.10. Extended formulation—algebraic loop. Example 1.

The procedure for handling algebraic loops is to first make an arbitrary causal assignment on one of the causally unassigned bonds and then propagate the causal information as far as possible. If the causality completes after one arbitrary assignment, then there is one algebraic loop. If a second arbitrary assignment is required, then a second algebraic loop is present, and so on. The more loops involved, the more difficult the final formulation becomes. In Figure 5.10*c* an arbitrary assignment has been made on R_4, which remained causally unassigned after the sources and energy storage elements had been assigned according to the procedure. The effort, e_4, has been assigned as an output from the R-element. This requires e_5 to be an output from the 0-junction, and now with e_5 and e_7 as inputs to the 1-junction, we are constrained to have R_6 in the causality indicated. Thus, making one arbitrary assignment has yielded a causally complete bond graph, and thus there is only one algebraic loop for this example.

From the appearance of Figure 5.10*c*, one would not realize that there is a problem. The bond graph causality is complete, all energy storage elements are in integral causality, and there are no causal conflicts. Let's see what happens if we attempt to derive the state equations using the procedure of the previous section. The state variables are p_2 and q_8, and the co-energy variables are f_2 and e_8. The input variables from sources are e_1 and f_a. Then,

$$\dot{p}_2 = e_1 - e_3 = e_1 - e_4 = e_1 - R_4 f_4 = e_1 - R_4 \left(p_2/I_2 - f_6\right). \tag{5.50}$$

So far, all looks fine. But,

$$f_6 = (1/R_6)e_6 = (1/R_6)(e_5 - e_7) = (1/R_6)\left(e_4 - (1/C_8)q_8\right), \tag{5.51}$$

and

$$e_4 = R_4 f_4 = R_4(f_3 - f_5) = R_4\left(p_2/I_2 - f_6\right) = \cdots \tag{5.52}$$

The flow, f_6, in Eq. (5.52) comes from Eq. (5.51), which leads back to e_4 again and we find ourselves in an endless loop of substitutions. This loop is not the fault of bond graphs. It is a direct result of the modeling assumptions that went into the schematic of Figure 5.10*a*. However, bond graphs and causality let you know there is a problem before any equations are derived.

The procedure for handling algebraic loops is to first recognize the problem from the resulting incomplete causality. This is followed by making an arbitrary causal assignment and extending the information according to the sequential assignment rules. This continues until all bonds are assigned. While it makes no difference which bond is selected for the arbitrary assignment, it is recommended from step 5 of the *sequential causality assignment procedure* that when a resistance element is involved in the algebraic loop (i.e., it has a causally unassigned bond), the resistance element be chosen for the arbitrary assignment. This was done in Figure 5.10*c* where the element R_4 was chosen for the arbitrary assignment.

As mentioned earlier, an algebraic loop is an indicator that some effort or flow variable involved in the loop is algebraically related to itself and other source and state variables. The next step in the procedure is to derive this algebraic relationship. When a resistance element is involved in the loop, it is customary to choose the variable that was arbitrarily selected to be the output from the R-element and the input to the rest of the system. For the example here, since R_4 was assigned arbitrarily, we select e_4 as the variable to be related to itself. Thus, we write e_4 on the left-hand side of an equation and then use causality in the normal way to derive the desired relationship, thus,

$$\begin{aligned} e_4 &= R_4 f_4 = R_4(f_3 - f_5) = R_4\left((1/I_2)p_2 - f_6\right) = R_4\left((1/I_2)p_2 - (1/R_6)e_6\right) \\ &= R_4\left((1/I_2)p_2 - (1/R_6)\left(e_4 - (1/C_8)q_8\right)\right). \end{aligned} \tag{5.53}$$

Equation (5.53) is an algebraic relationship that relates e_4 to source variables, state variables, and to e_4 itself.

The next step is to solve the resulting relationship for the algebraic variable, e_4 in this case,

$$e_4 = \frac{(R_4/I_2)}{(1 + R_4/R_6)}p_2 + \frac{(R_4/R_6C_8)}{(1 + R_4/R_6)}q_8. \tag{5.54}$$

The final step is to derive the state equations following the standard procedure, only stop the formulation when the algebraic variable enters the formulation. For our example,

$$\begin{aligned} \dot{p}_2 &= e_1 - e_3 = e_1 - e_4, \\ \dot{q}_8 &= f_9 + f_7 = f_9 + f_6 = f_9 + (1/R_6)\left(e_4 - (1/C_8)q_8\right). \end{aligned} \tag{5.55}$$

We now substitute (5.54) into (5.55) and end up with the state equations in standard explicit form,

$$\begin{aligned} \dot{p}_2 &= e_1 - \frac{(R_4/I_2)}{(1 + R_4/R_6)}p_2 - \frac{(R_4/R_6C_8)}{(1 + R_4/R_6)}q_8, \\ \dot{q}_8 &= f_9 - (1/R_6C_8)q_8 + \frac{1}{R_6}\left[\frac{R_4/I_2}{(1 + R_4/R_6)}p_2 + \frac{R_4/R_6C_8}{(1 + R_4/R_6)}q_8\right]. \end{aligned} \tag{5.56}$$

To complete this discussion of algebraic loops, Figure 5.10*d* shows the same example system with a different arbitrary causal assignment. This time f_4 is assigned as the output from R_4, and the causality completes as shown. It makes no difference what bond is assigned arbitrarily. This *system* has an algebraic

problem consisting of one algebraic loop. Choosing f_4 as the variable to relate to itself, we write,

$$\begin{aligned} f_4 &= (1/R_4)e_4 = (1/R_4)e_5 = (1/R_4)(e_6 + e_7) = (1/R_4)\,(R_6 f_6 + (1/C_8)q_8) \\ &= (1/R_4)\,(R_6(f_3 - f_4) + (1/C_8)q_8) \\ &= (1/R_4)\,(R_6\,((1/I_2)p_2 - f_4) + (1/C_8)q_8)\,. \end{aligned} \tag{5.57}$$

Solving for f_4 yields

$$f_4 = \frac{R_6}{(1 + R_6/R_4)R_4 I_2}\,p_2 + \frac{1}{(1 + R_6/R_4)R_4 C_8}\,q_8. \tag{5.58}$$

Equation derivation would follow the standard procedure except that we would stop when f_4 enters the formulation,

$$\begin{aligned} \dot{p}_2 &= e_1 - R_4 f_4, \\ \dot{q}_8 &= f_9 + f_5 = f_9 + (1/I_2)p_2 - f_4. \end{aligned} \tag{5.59}$$

The formulation is completed when (5.58) is substituted into (5.59). It is left to any non-believing reader to show that the resulting state equations are identical to (5.56).

So far in this section we have determined that some modeling assumptions can lead to algebraic formulation problems, and that bond graphs identify such a problem by the presence of incomplete causality after using the sequential causal assignment procedure up to step 4. A procedure has been presented through an example of a linear system that shows how to solve the algebraic problem and end up with state equations in standard form.

For linear systems, this procedure will work for multiple loops involving several algebraic variables, although solving by hand becomes formidable rather quickly. For nonlinear systems, it may not be possible to invert some of the algebraic relationships. For example, friction is a very common resistance effect in mechanical systems and is represented by an R-element in a bond graph. For the most common friction representation (see Figure 8.2 in Section 8.1 of Chapter 8), the velocity (or flow) must be an input and the friction force (or effort) must be an output from the element. If the friction element was involved in an algebraic loop that required its inversion (calculation of velocity from knowledge of the force), in many cases it could not be done because a given force may have no corresponding velocity or several.

We probably should look into the modeling assumptions that created the problem in the first place. In Figure 5.10*a*, the connection of the spring and damper was *assumed* to be massless. The connection is not actually massless, it was just assumed to be and the assumption was probably justifiable during the modeling process. Consider Figure 5.11*a*, where a small mass, m', has been associated with the connection point. Figure 5.11*b* shows the bond graph for the system with causality assigned. No arbitrary assignment is required and there is no algebraic

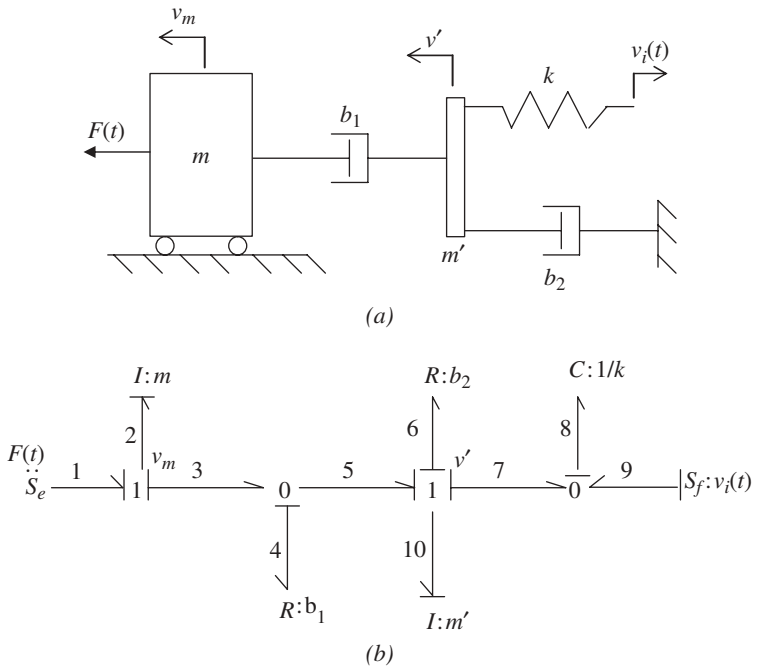


FIGURE 5.11. Eliminating an algebraic loop by revisiting modeling assumptions in Example 1.

loop. Equation derivation is straightforward for this system. If this had been the first model constructed, we would never have known that a potential problem exists. In Chapter 13, the concept of algebraic loops is dealt with more formally, and the idea of revisiting modeling assumptions when formulation issues arise is developed in much more detail. Reference [1] describes a formal vector-field method for generating state space equations.

5.4.1 Extended Formulation Methods—Derivative Causality

In this section we study systems where, after sequential causal assignment, one or more energy storage elements are forced into derivative causality. As with algebraic loops, the occurrence of derivative causality is due to the modeling assumptions made during the model construction. Storage elements in derivative causality are not dynamically independent of the remaining storage elements, and they do not contribute a state variable to the final equations.

Mathematically, Eqs. (5.10) or (5.11), are called *explicit* differential equations because the state variable derivatives are given as explicit functions of the state variables and the input variables. It is this form of the equations of motion that is best suited to analysis and simulation. When derivative causality arises, the state equations become *implicit,* meaning that some of the state variable derivatives

become functions of themselves as well as of the state and input variables. In principle, implicit state equations can be algebraically transformed to explicit ones but, as we will see, this is not necessarily a simple task and many computer simulation programs are not equipped to do it.

The independent state variables remain the p's on I's and the q's on C's in *integral* causality. The derivative elements do still store energy, and this energy is accounted for in the bond graph. By being dynamically dependent, their associated energy variable, p or q, is algebraically related to the independent state variables. So, knowledge of the independent state variables allows algebraic determination of the energy variables on the derivative elements, and thus calculation of their contribution to the system energy. The algebraic relationships, relating the dependent and the independent energy variables must be derived before it is possible to have a state representation in explicit form. This is demonstrated through example.

Consider the physical system shown in Figure 5.12*a*. It consists of a permanent magnet dc motor driving a rotational load, J, through a flexible shaft of torsional stiffness, k_τ. There is also some rotational damping with constant, b_τ. On the electrical side of the motor, the winding resistance, R_w, and the winding inductance, L, are important dynamic effects. It has been decided, for this application, that the rotational inertia of the motor is negligible and it is not included. Figure 5.12*b* shows a bond graph for this system where the motor has been characterized as a gyrator with coupling constant, T[N · m/A] (see Section 4.4.2 to remind you how to model dc motors).

In the figure, the bonds have been numbered and causality has been assigned using the sequential assignment procedure. The inductance, I_3, has been put into integral causality. As a result, the compliance, C_6, was constrained to be in derivative causality. The state variables for this system are p_3 and p_8, and only two state equations need to be derived. The single input variable is e_c. The energy variable, q_6, is not an independent state variable, but is algebraically dependent on the independent state variables.

We cannot simply ignore the presence of the derivative element. If we try to derive the equations without its consideration, then we would write,

$$\begin{aligned}\dot{p}_3 &= e_1 - e_2 - e_4 = e_1 - (R_2/I_3)p_3 - Tf_5 \\ &= e_1 - (R_2/I_3)p_3 - T\,(f_6 + (1/I_8)p_8)\,. \end{aligned} \tag{5.60}$$

The appearance of f_6 is where the derivative element enters the formulation, and f_6 must be dealt with before final formulation can be done.

We next recognize that

$$f_6 = \dot{q}_6,$$

but there is no state equation for q_6, and more work must be done.

The next step is to derive the algebraic relationship relating q_6 to the state variables and source variables. When we do this, we start by writing the constitutive

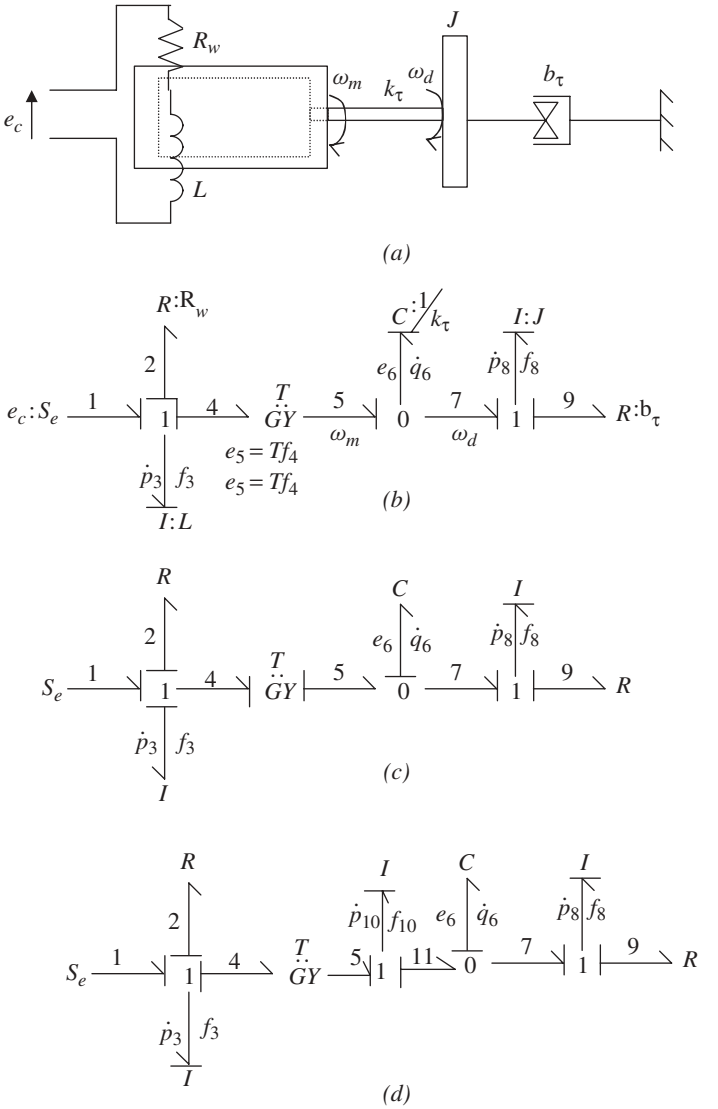


FIGURE 5.12. Derivative causality in a bond graph. Example 1.

law "backwards," that is,

$$q_6 = C_6 e_6. \tag{5.61}$$

(For an I-element in derivative causality, we would write $p_i = I_i f_i$. Also, for nonlinear elements, we would start with $q_6 = \Phi_C(e_6)$ for compliance elements in derivative causality, and $p_i = \Phi_I(f_i)$ for inertial elements in derivative causality.)

Starting with Eq. (5.61), we follow causality in the standard fashion to find out how e_6 is determined:

$$e_6 = e_5 = Tf_4 = Tf_3 = (T/I_3)p_3. \tag{5.62}$$

Thus,

$$q_6 = (C_6T/I_3)p_3. \tag{5.63}$$

This is the algebraic relationship that allows the energy variable, q_6, to be algebraically determined from the state variable, p_3. Taking the derivative of (5.63) yields

$$\dot{q}_6 = (C_6T/I_3)\dot{p}_3. \tag{5.64}$$

We can now substitute (5.64) into (5.60), with the result

$$\dot{p}_3 = e_1 - (R_2/I_3)p_3 - T\left((C_6T/I_3)\dot{p}_3 + (1/I_8)p_8\right). \tag{5.65}$$

When dealing with derivative causality, we end up with an implicit equation with derivatives of state variables on both sides of the equation. Equation (5.65) is particularly simple to reduce to the explicit form because q_6 depends on only one state variable. We can simply bring $\dot{p}_3$ from the right side to the left, with the result

$$\dot{p}_3 = \frac{e_1}{(1 + C_6T^2/I_3)} - \frac{(R_2/I_3)p_3}{(1 + C_6T^2/I_3)} - \frac{(T/I_8)p_8}{(1 + C_6T^2/I_3)}. \tag{5.66}$$

The second and last state equation starts with

$$\dot{p}_8 = e_7 - (R_9/I_8)p_8.$$

But e_7 is not *caused* by the compliance element; instead,

$$e_7 = e_5 = Tf_4 = (T/I_3)p_3, \tag{5.67}$$

with the final result

$$\dot{p}_8 = (T/I_3)p_3 - (R_9/I_8)p_8. \tag{5.68}$$

Equations (5.66) and (5.68) can be put into the standard matrix format for linear systems,

$$\frac{d}{dt}\begin{bmatrix} p_3 \\ p_8 \end{bmatrix} = \begin{bmatrix} -R_2/I_3Q & -T/I_8Q \\ T/I_3 & -R_9/I_8 \end{bmatrix}\begin{bmatrix} p_3 \\ p_8 \end{bmatrix} + \begin{bmatrix} 1/Q \\ 0 \end{bmatrix} e_1, \tag{5.69}$$

where

$$Q = 1 + C_6T^2/I_3.$$

When this example was presented, the I-element on bond 3 in Figure 5.12*b* was put into integral causality and this resulted in the C_6 element being in derivative causality. In Figure 5.12*c* the bond graph is repeated, but this time the C_6 element is put into integral causality, with the result that the I_3 element ends up in derivative causality. We have some flexibility in choosing which elements are in integral causality, but, once chosen, some elements will be constrained to be in derivative causality. The derivative causality resulted from the original modeling assumptions, and without changing these assumptions, the model will have derivative causality regardless of causal assignment choices defining which of the energy variables are considered to be independent.

When deriving equations for Figure 5.12*c*, the effort $e_3 = \dot{p}_3$ will enter the formulation. Thus, we will need to write the constitutive law "backwards" for the I_3 element and determine how the energy variable, p_3, is algebraically related to the state variables. Thus,

$$p_3 = I_3 f_3 = I_3 f_4 = (I_3/T)e_5 = (I_3/TC_6)q_6 \tag{5.70}$$

and

$$e_3 = \dot{p}_3 = (I_3/TC_6)\dot{q}_6. \tag{5.71}$$

Now we can derive the equations using standard procedures, starting with

$$\dot{q}_6 = f_5 - f_7 = (1/T)e_4 - (1/I_8)p_8 = (1/T)(e_1 - e_2 - e_3) - (1/I_8)p_8. \tag{5.72}$$

We see that e_3 entered the formulation, and we will eliminate it using (5.71); we should note that $e_2 = R_2 f_2$ and f_2 is *not caused* by I_3, but must be tracked through the causal information. Continuing,

$$\dot{q}_6 = (1/T)\,(e_1 - (R_2/TC_6)q_6 - (I_3/TC_6)\dot{q}_6) - (1/I_8)p_8. \tag{5.73}$$

Again we see that a derivative of a state variable is on both sides of the equation. And again, this equation is very simple to get into the standard form by simply bringing the $\dot{q}_6$ terms from the right-hand side to the left, eventually resulting in

$$\dot{q}_6 = \frac{1/T}{(1 + I_3/T^2C_6)}e_1 - \frac{R_2/T^2C_6}{(1 + I_3/T^2C_6)}q_6 - \frac{1/I_8}{(1 + I_3/T^2C_6)}p_8. \tag{5.74}$$

The second state equation is

$$\dot{p}_8 = (1/C_6)q_6 - (R_9/I_8)p_8. \tag{5.75}$$

Comparing (5.66) and (5.68) with (5.74) and (5.75), we see that the two state representations of the same system have different appearances. This is not

surprising since we are using different state variables for the two representations. Rest assured that if both sets of equations were solved, the predicted system dynamics would be identical from both state representations.

It has been emphasized that modeling assumptions are the cause of algebraic problems in formulations. For the example system we have just shown it was assumed that the rotational inertia of the electric motor was negligible and was not included in the model. Figure 5.12*d* shows the bond graph for the system when the rotational inertia of the motor, J_m, is included. The angular velocity of the motor, ω_m, is on bond 5 in the bond graphs in parts *b* and *c*. In part *d*, a 1-junction has been inserted, exposing ω_m, and the I-element representing the motor inertia has been attached (it is bond 10 in Figure 5.12*d*). Causality has been assigned, and there is no derivative causality.

The new modeling assumption has eliminated any formulation problems. If these modeling assumptions had been made originally, we would have never known that a potential formulation problem exists. It should be noted that the model in Figure 5.12*d* has four state variables and four state equations, whereas the model in part *b* has only two state variables and two state equations. Including the motor inertia has increased the number of state variables by two, one being the new I-element in integral causality and the other coming from the storage element formerly in derivative causal element that is now in integral causality. If we were going to do pencil and paper analysis, then two additional equations is a significant further complication. If our intention is computer solution, then the change is almost insignificant. The concept of revisiting modeling assumptions to alleviate formulation problems is discussed in detail in Chapter 13.

One more example involving derivative causality may serve to sufficiently illustrate the pattern of formulation and reduction. Figure 5.13*a* shows masses m_1 and m_2 on the left and right ends of a massless lever. A force, $F(t)$, acts on m_1, and a spring, k, is attached to m_2. Positive force and velocity directions are indicated. Part *b* of the figure shows a bond graph for this system, with bonds numbered and the transformer modulus defined by the relationships shown for the transformer and representing the input force by e_4. In part *c* causality has been assigned, putting I_2 into integral causality, with the result that I_1 is in derivative causality. The independent state variables are p_2 and q_3, and the dependent variable p_1 is not a state variable but will impact the equation formulation. Thus, we write the constitutive law backwards for this element and derive the algebraic relationship relating p_1 to the other state variables. Thus,

$$p_1 = I_1 f_1 = I_1(a/b) f_6 = (I_1/I_2)(a/b) p_2 \tag{5.76}$$

and, taking the derivative of Eq. (5.76),

$$e_1 = \dot{p}_1 = (I_1/I_2)(a/b)\dot{p}_2. \tag{5.77}$$

Equation derivation starts with

$$\dot{p}_2 = e_6 - e_3 = (a/b)e_5 - (1/C_3)q_3 = (a/b)(e_4 - \dot{p}_1) - (1/C_3)q_3. \tag{5.78}$$

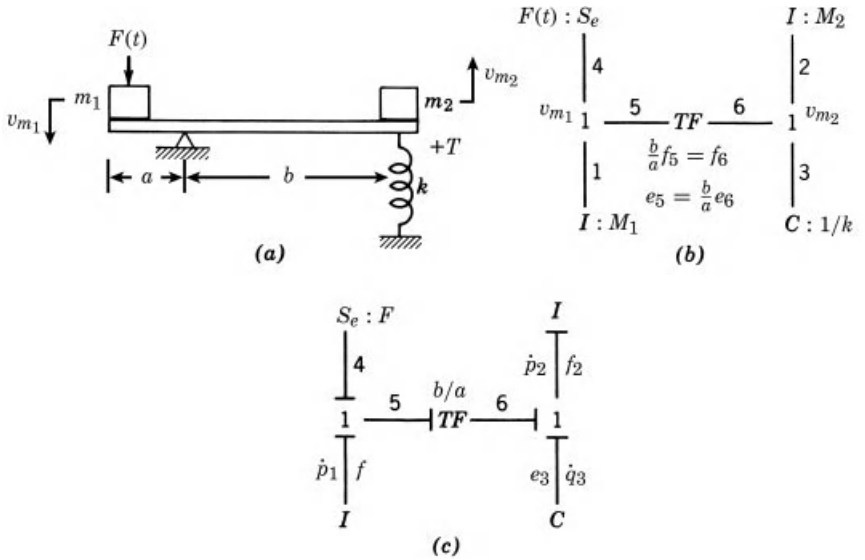


FIGURE 5.13. Derivative causality in a bond graph. Example 2.

We substitute from (5.77), with the result

$$\dot{p}_2 = (a/b)\,(e_4 - (I_1/I_2)(a/b)\,\dot{p}_2) - (1/C_3)q_3, \tag{5.79}$$

where the state variable derivative $\dot{p}_2$ appears on both sides of the equation. The final result is

$$\dot{p}_2 = \frac{(a/b)}{1 + (I_1/I_2)(a/b)^2} e_4 - \frac{1/C_3}{1 + (I_1/I_2)(a/b)^2} q_3, \tag{5.80}$$

and the second state equation is

$$\dot{q}_3 = (1/I_2)p_2. \tag{5.81}$$

The reader must appreciate that the two examples shown for derivative causality are quite simple in that (1) only one energy storage element was in derivative causality and (2) the dependent energy variable was algebraically related to only one independent state variable. For more complex systems the algebra required to reduce a set of equations to a standard equation format can be formidable at best, and, for nonlinear systems, practically impossible at worst.

It is always possible to revisit the modeling assumptions that resulted in the formulation problem and change these assumptions to yield a model that is virtually the same as the one originally intended but without the formulation problems. There are some computational penalties that accompany the addition of very

small added inertias or very stiff added compliances in a model. But for some systems these penalties are a very small price to pay for eliminating formulation problems. As mentioned, this topic is discussed in detail in Chapter 13.

For the system in Figure 5.13, it is interesting to see how derivative causality could have been avoided. Figure 5.14*a* shows a physical system where the lever has been divided into two massless pieces, with the right end of one and the left end of the other pivoted at the same point. The two pieces are coupled by a torsional spring of stiffness, k_τ, which represents the fact that the lever is not actually rigid, but has some flexibility. The system is otherwise identical to the system of Figure 5.13*a*. The angular velocity of the left side lever is ω_1 and that of the right side lever is ω_2. The torsional spring "sees" the difference

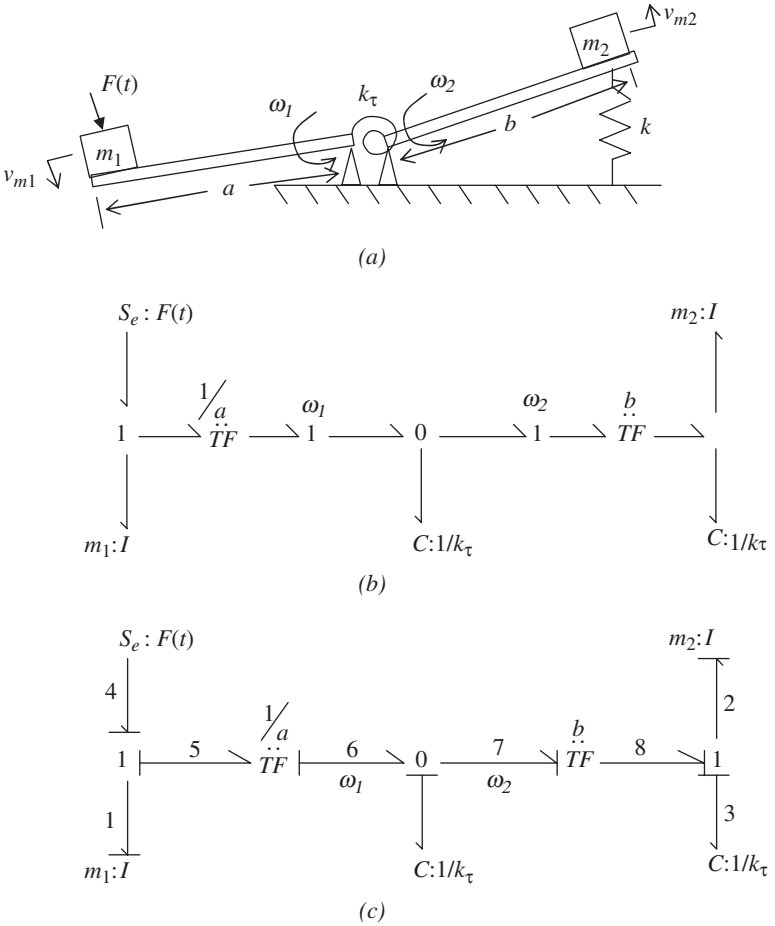


FIGURE 5.14. Avoiding derivative causality by changing the modeling assumptions for Example 2.

between these angular velocities, and the bond graph for the system is shown in Figure 5.14*b*.

Figure 5.14*c* shows the bond graph reduced and with causality assigned. There is no derivative causality, and equation formulation would be straightforward. If the torsional spring is made stiff enough, then there will be very little difference between ω_1 and ω_2, and a close approximation to the original system will result.

5.5 OUTPUT VARIABLE FORMULATION

We have seen that it is very convenient to derive state equations for all types of systems using the energy variables, p's and q's, as the state variables. More often than not, however, the system outputs that we want to observe and plot are not the p and q variables. Can you imagine interpreting the motion of a mechanical system from a plot of a mass momentum, or the behavior of a circuit from a plot of the flux linkage in a coil? You might want to look at these responses, but you also might want to look at mass velocities and displacements, circuit voltages and currents, hydraulic pressures, and so on. It turns out that any effort or flow on any bond in a bond graph can be related to the state variables. Thus, we solve systems of equations in terms of the state variables, and then derive output equations that relate the desired outputs to the state variables. If the system solution is done computationally, then it is a simple matter to append output equations at the end of the program so that desired outputs can be viewed.

Consider the system in Figure 5.9, which was used earlier to demonstrate equation formulation. If we desire the flow on bond 1 as an output, then we simply follow the causal path and relate f_1 to the state and source variables. Thus,

$$f_1 = f_8 = (1/R_8)e_8 = (1/R_8)\,(e_1 - e_9) = (1/R_8)\,(e_1 - (1/C_3)q_3)\,. \qquad (5.82)$$

By following causality we have related the desired output to the known source variable, e_1, and the state variable, q_3. If we wanted the effort on bond 5, then

$$\begin{aligned} e_5 &= e_2 - e_6 - e_7 - e_4 + e_{11} \\ &= e_2 - (R_6 + R_7)\,(1/I_5)p_5 - (1/C_4)q_4 + (1/mC_3)q_3. \end{aligned} \qquad (5.83)$$

In this case, q_3, q_4, and p_5 are state variables and e_2 is an input variable from a source.

These two output equations were derived assuming that the system was linear, which was assumed when this system was used for equation derivation. When the system is linear, the output equations can be put into a matrix format,

$$Y = CX + DU, \qquad (5.84)$$

where Y is a vector of desired outputs,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \bullet \\ \bullet \\ y_k \end{bmatrix}, \tag{5.85}$$

U is the vector of inputs,

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \bullet \\ \bullet \\ u_r \end{bmatrix}, \tag{5.86}$$

and C is a $k \times n$ matrix of coefficients and D is a $k \times r$ matrix of coefficients where the coefficients are made up of system parameters. For the output equation (5.82) and (5.83), the matrix form would be

$$\begin{bmatrix} f_1 \\ e_5 \end{bmatrix} = \begin{bmatrix} -1/R_8C_3 & 0 & 0 \\ 1/mC_3 & -1/C_4 & -(R_6 + R_7)/I_5 \end{bmatrix} \begin{bmatrix} q_3 \\ q_4 \\ p_5 \end{bmatrix} + \begin{bmatrix} 1/R_8 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, \tag{5.87}$$

If the system is nonlinear, the desired outputs can still be expressed in terms of the state variables and source variables, but the nice matrix notation of Eq. (5.84) is not available. For example, in Figure 5.9, if the resistance element R_8 and the compliance element C_3 are nonlinear such that f_8 is a function of e_8,

$$f_8 = \Phi_R^{-1}(e_8), \tag{5.88}$$

and e_3 is a function of q_3,

$$e_3 = \Phi_C^{-1}(q_3), \tag{5.89}$$

then the output, f_1, would be derived as

$$f_1 = f_8 = \Phi_R^{-1}(e_8) = \Phi_R^{-1}(e_1 - e_3) = \Phi_R^{-1}\left(e_1 - \Phi_C^{-1}(q_3)\right). \tag{5.90}$$

Causality still shows us the path from the output of interest to the known variables, but the relationships are not linear. From a computational point of view it makes theoretical difference whether the equations are linear or nonlinear. If the output we desire is the power on some bond, then the output equation will be nonlinear whether or not the system is linear. Power is always composed of the

product of an effort and a flow. In Figure 5.9, if we desire the power on bond 1, P_1, as an output variable, then

$$P_1 = e_1 f_1 = (e_1/R_8)\,(e_1 - (1/C_3)q_3)\,, \tag{5.91}$$

where f_1 is the linear version from Eq. (5.82). Since we are most often dealing with computational solutions of system equations, it really does not matter as to the complexity or linearity of the output equations, other than getting them typed and formatted correctly for the program being used.

5.6 NONLINEAR AND AUTOMATED SIMULATION

A primary reason for constructing a physical engineering model of a dynamic system is to predict the response of the system prior to actually building it. A good model can go a long way toward avoiding mistakes in a first prototype of a complex system. Sometimes it is justified to assume that a system behaves linearly, thus allowing the model to be constructed from linear inertance, compliance, and resistance elements and restricted to small motions with no nonlinear kinematics. When this is justified, linear analysis can be performed, sometimes analytically but most often computationally. Some of the more important linear concepts for engineers are developed in the next chapter.

More often than not, at some point in the system development process, a more realistic nonlinear model must be developed in order to really understand the system, or to test a control strategy that was developed using linear thinking, or to see what happens when displacements hit hard boundaries or power amplifiers saturate, for example. For nonlinear systems there is really only one choice for predicting the system response, and that is computer simulation. This topic is discussed thoroughly in Chapter 13. Here we discuss how to organize system equations for computer simulation and demonstrate with an example. We also discuss the concept of automated simulation where bond graph models are communicated graphically to the computer and causality and the equivalent of state equations are automatically generated in symbolic form even for nonlinear systems.

5.6.1 Nonlinear Simulation

Figure 5.15*a* is the quarter-car model repeated from Figure 4.10, where it was used to demonstrate bond graph construction procedures. This time, the suspension spring and damper and tire spring are all treated as nonlinear elements. A sketch of the constitutive relationships for each element is in the schematic. Also shown next to $v_i(t)$ is a roadway bump the vehicle is going to hit. This bump is shown as a height versus distance and has the appearance of a speed bump such as found in parking lots. We will characterize the bump as $\frac{1}{2}$ of a sine wave, and interpret the bump as an equivalent velocity as required for bond graph flow

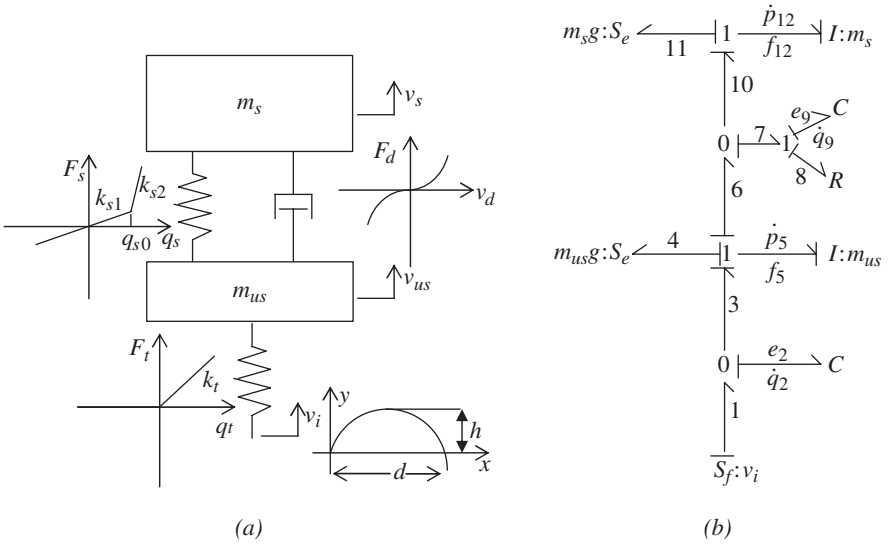


FIGURE 5.15. Quarter-car simulation example with nonlinear elements.

inputs. For the spatially distributed input,

$$y = h \sin \pi \frac{x}{d} \quad \text{for } 0 < \frac{x}{d} \leq 1$$

and

$$y = 0 \quad \text{for } (x/d) > 1. \tag{5.92}$$

If we imagine the vehicle traveling at constant horizontal velocity, U, toward the right, then $x = Ut$ and (5.92) becomes

$$y = h \sin \pi (U/d)t \quad \text{for } 0 < (U/d)t \leq 1. \tag{5.93}$$

A reasonable characterization of the vertical input velocity is to multiply the forward velocity times the slope of the roadway, dy/dx, thus,

$$\begin{aligned} v_i(t) &= (h/d)\pi U \cos \pi (U/d)t \quad \text{for } 0 < (U/d)t \leq 1, \\ v_i(t) &= 0 \text{ otherwise.} \end{aligned} \tag{5.94}$$

We will characterize the suspension spring as a linear spring with slope, k_{s1}, up until a compressive displacement of q_{s0}, at which point the slope changes to a stiffer spring, k_{s2}. Thus,

$$\begin{aligned} F_s &= k_{s1} q_s \quad \text{for } q_s \leq q_{s0}, \\ F_s &= k_{s1} q_{s0} + k_{s2} (q_s - q_{s0}) \quad \text{for } q_s > q_{s0}. \end{aligned} \tag{5.95}$$

The damper will be characterized as

$$F_d = Bv_d^3, \tag{5.96}$$

and, finally, the tire spring will be characterized as linear in compression while generating no force in extension, that is, no force when the tire is off the ground. Thus,

$$\begin{aligned} F_t &= k_t q_t \quad \text{for } q_t \geq 0, \\ F_t &= 0 \quad \text{for } q_t < 0. \end{aligned} \tag{5.97}$$

The bond graph for the example system is shown in Figure 5.15*b*. Causality has been assigned and bonds have been numbered. Notice that there are no parameters listed with the suspension spring and damper or with the tire spring. These are nonlinear elements and this is indicated on the bond graph by the absence of a single identifying parameter. The state variables are q_2, p_5, q_9, and p_{12}.

Equations (5.10) are the functional form of standard nonlinear state-space equations. These are called explicit first-order differential equations and can be expressed in shorthand form as

$$\dot{X} = F\,(X, U) \tag{5.98}$$

where x is the vector of state variables,

$$\dot{X} = \begin{bmatrix} x_1 \\ x_2 \\ \bullet \\ \bullet \\ x_n \end{bmatrix},$$

U is the vector of inputs,

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \bullet \\ \bullet \\ u_r \end{bmatrix},$$

and f is a vector of functions of states and inputs,

$$F(\bullet) = \begin{bmatrix} f_1(X, U) \\ f_2(X, U) \\ \bullet \\ \bullet \\ f_n(X, U) \end{bmatrix}.$$

It turns out that equations that can be derived in this condensed form are particularly easy to solve numerically (see Chapter 13). However, it is enough to

know that equations could be put into this form, but it is not necessary to do so in order to carry out numerical simulation. In fact, it is rare to derive equations in the explicit form of (5.98) when the intention is numerical simulation. This is demonstrated with our example.

Equation formulation for the bond graph of Figure 5.15*b* begins with

$$\dot{q}_2 = f_1 - (1/I_5)p_5, \tag{5.99}$$

where f_1 is the input velocity, $v_i(t)$, from (5.94). This state equation turns out to be linear even though the system is nonlinear. The second equation is

$$\dot{p}_5 = -e_4 + e_2 - e_7. \tag{5.100}$$

The effort $e_4 = m_{\text{us}}g$, the effort, e_2 comes from the nonlinear tire spring characterized by (5.97), and

$$e_7 = e_8 + e_9. \tag{5.101}$$

The effort e_8 is the damper force from (5.96), and the effort e_9 is from the nonlinear suspension spring from Eq. (5.95).

The next state equation is

$$\dot{q}_9 = (1/I_5)p_5 - (1/I_{12})p_{12}, \tag{5.102}$$

which is another linear state equation.

The final state equation is

$$\dot{p}_{12} = -e_{11} + e_7, \tag{5.103}$$

where $e_{11} = m_s g$ and e_7 has already been discussed.

The equation formulation is now complete. Bond graph causality told us that explicit equations would result, but there is no need to reduce the equations to this final form. It is perfectly all right to leave the equations in uncondensed form as long as all the variables are accounted for. Equations (5.94)–(5.97) with (5.99)–(5.103) are a complete, computable set of equations ready to simulate the quarter-car system. We need only specify initial conditions for the state variables, and the simulation can proceed.

In addition, many commercial equation solvers have a *sort* feature. This means that the user need not pay any attention to the order in which the equations appear in the program. The program determines the order in which the equations must be used and properly sorts the equations for the user. For example, Eq. (5.100) requires e_7 to evaluate $\dot{p}_5$, but e_7 appears in order below Eq (5.100) in Eq. (5.101). If a program were proceeding in order to evaluate the state variable derivatives, an error would occur, because the current value of e_7 would not be known at the time when the program attempts to compute $\dot{p}_5$. A sort program assures that all equations needed to compute state variable derivatives have been

evaluated before the derivatives are computed. (This is possible in the absence of algebraic loops and derivative causality.)

Some simulation results for this example are presented below, but first we discuss automated simulation.

5.6.2 Automated Simulation

It is hoped that the reader recalls how straightforward it is to assign causality to a bond graph and thus know, before deriving any equations, if any formulation problems such as algebraic loops or derivative causality exist. With causality assigned, state variables and input variables are determined. Computers can follow the same rules as humans and can assign sequential causality to a bond graph, choose state variables, and derive the first-order state equations. There are several commercial programs that start from a graphical description of a bond graph and automatically deliver state equations ready for simulation. The user need only input the parameters, define any nonlinearities, specify the input functions, and specify the initial conditions. The program will then simulate the system and plot any desired outputs, including any state variable, any effort or flow variable, or any output defined by the user as a combination of state variables and input variables. These programs are so straightforward to use that a user unfamiliar with bond graph modeling could carry out a simulation if a bond graph model including sign half-arrows were provided. In Chapter 1 some references are given that describe some of the better known commercial bond graph processors.

For our example system, a bond graph processor program would follow the causality of Figure 5.15*b* and deliver the following uncondensed equations:

$$
\begin{aligned}
e_1 &= e_2 & e_7 &= e_8 + e_9 \\
e_3 &= e_2 & \dot{q}_9 &= f_9 \\
f_2 &= f_1 - f_3 & e_8 &= R_8 f_8 ?? \\
\dot{q}_2 &= f_2 & f_{10} &= f_{12} \\
f_3 &= f_5 & f_{11} &= f_{12} \\
f_4 &= f_5 & e_{12} &= e_{10} - e_{11} \\
f_6 &= f_5 & \dot{p}_{12} &= e_{12} \\
e_5 &= e_3 - e_4 - e_6 & e_2 &= (1/C_2) q_2 ?? \\
\dot{p}_5 &= e_5 & f_5 &= (1/I_5) p_5 ?? \\
e_6 &= e_7 & e_9 &= (1/C_9) q_9 ?? \\
e_{10} &= e_7 & f_{12} &= (1/I_{12}) p_{12} ?? \\
f_7 &= f_6 - f_{10} & & \\
f_8 &= f_7 & & \\
f_9 &= f_7 & & \qquad (5.104)
\end{aligned}
$$

If the reader follows the causality of Figure 5.15*b* it will be discovered that every input/output causal implication is reflected in Eqs. (5.104). Of particular interest are the equations with question marks next to them. When an automated program derives equations, it has no idea which components are linear or nonlinear nor how to characterize any specific nonlinearities. The program does recognize which elements might be nonlinear, and it highlights these with a question mark or some other indicator. For this example, the resistance element, R_8, is nonlinear and the user would have to modify the linear equation in (5.104) with

$$e_8 = Bf_8^3 \tag{5.105}$$

from (5.96) above. Neither of the inertial elements is nonlinear, so the user would only need to provide the inertia parameter for each. Both compliance elements are nonlinear, so the equation for e_2 in Eq. (5.104) would have to be modified using Eq. (5.97), and the equation for e_8 would have to be modified according to Eq. (5.95). Finally, the input, f_1, would be defined according to Eq. (5.94), and e_4 and e_{11} must be defined as the respective weights of the unsprung and sprung vehicle masses. Along with initial conditions (initial values of the state variables), the simulation is ready to be run as long as a sorting equation solver is used.

For the example system, the parameters used in the simulation are shown in Table 5.1.

The system was simulated using a commercial equation solver for two different forward speeds, U, listed in Table 5.1. Figure 5.16 shows the velocity input from (5.94) for the lower vehicle forward speed of 2 mph. Figure 5.17 shows the suspension spring and damper forces along with the tire spring force. The vehicle starts from static conditions where the masses are not moving vertically

TABLE 5.1. Parameters for the Example Simulation

$m_s = 320$ kg
$\dfrac{m_{\text{us}}}{m_s} = \dfrac{1}{6}$
f_s, suspension frequency, $= 1.0$ Hz
$\omega_s = 2\pi f_s$, rad/s
For the suspension spring, $k_{s1} = m_s\omega_s^2$, N/m
$k_{s2} = 10k_{s1}$, N/m
For the tire spring, $k_t = 10k_{s1}$, N/m
Initial conditions for springs, $q_{s-\text{ini}} = m_s g/k_{s1}$ $q_{\text{us}-\text{ini}} = (m_s + m_{\text{us}})g/k_t$
Breakpoint for suspension spring, $q_{s0} = 1.3q_{s-\text{ini}}$
B, damper parameter, $= 1500.0$ N/(m/s)^3
For the input bump, $h = 0.25$ m
$d = 1.0$ m
Forward velocity, $U = 2$ and 30 mph $= 0.9$ and 13.5 m/s

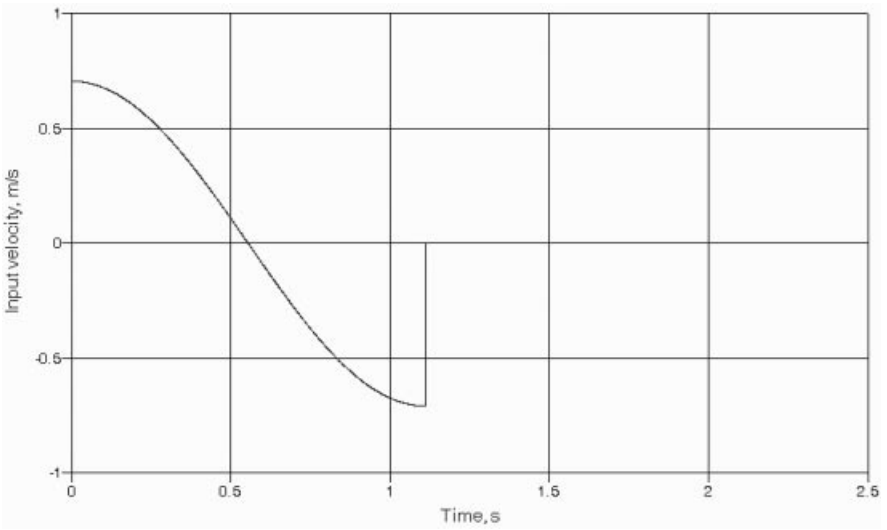


FIGURE 5.16. Input velocity for the simulation example for a vehicle speed of 2 mph.

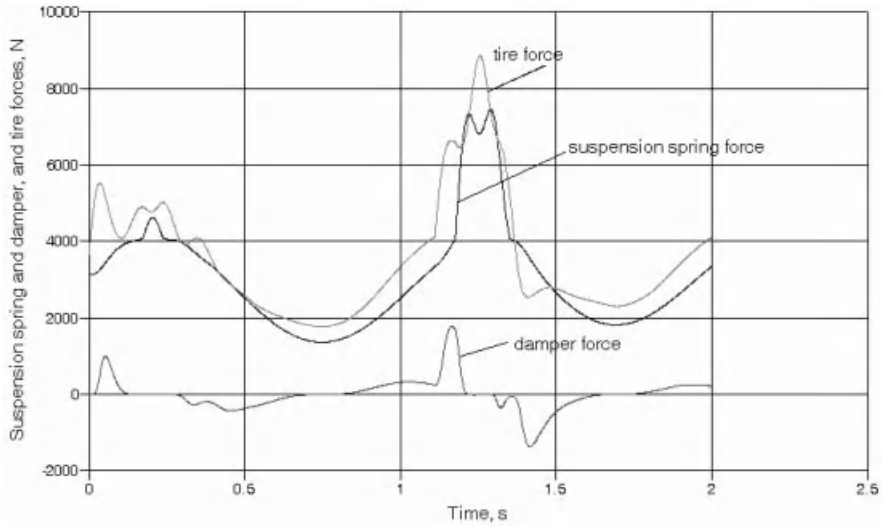


FIGURE 5.17. Suspension spring and damper forces and tire forces for a vehicle speed of 2 mph.

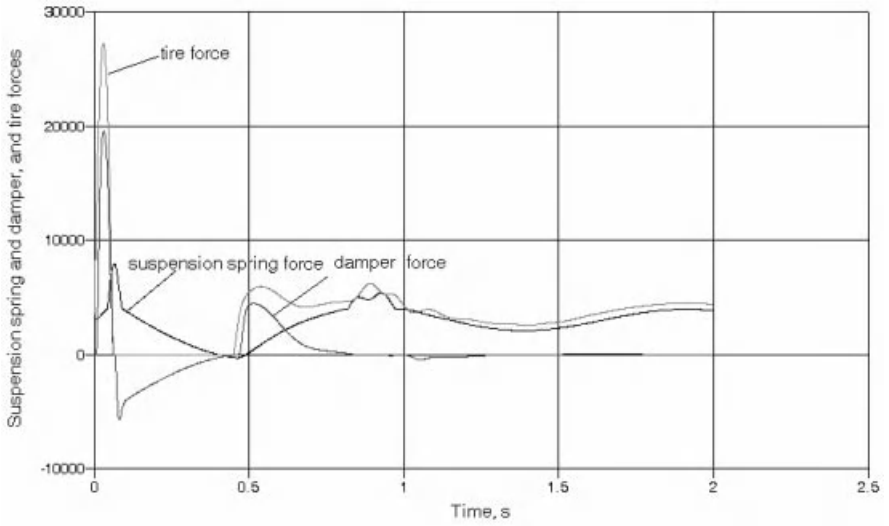


FIGURE 5.18. Suspension spring and damper forces and tire forces for a vehicle speed of 30 mph.

and the springs are compressed to their initial values. As the bump is traversed, the springs become more compressed and the damper force becomes nonzero as a relative velocity is established across the suspension. The sharp change in slope of the suspension spring force at about 1.2 s is due to the spring displacement becoming more compressed than the breakpoint, q_{s0}, and thus the spring becomes much stiffer. The tire displacement never becomes negative in this simulation, and thus the tire force follows a linear law as indicated in (5.97).

When the vehicle forward speed is increased to 30 mph, the force results are as shown in Figure 5.18. This time the tire force becomes zero just after the vehicle strikes the bump. The tire spring first compresses and then uncompresses until the tire deflection becomes negative, the wheel is off the ground, and the tire force becomes zero. When the wheel returns to the ground, the tire spring becomes compressed once again. It is interesting to plot the damper force versus the relative velocity across the suspension as shown in Figure 5.19. The cubic constitutive law is very obvious in this view. Figure 5.20 shows the suspension spring force versus the suspension deflection, and again the nonlinear constitutive behavior is very apparent. Finally, Figure 5.21 shows the tire force plotted against the tire deflection, and once again the nonlinear behavior is exposed.

It is hoped that the reader has gained some appreciation for the virtues of bond graph modeling for obtaining simulation results for nonlinear systems. Chapter 13 covers this topic more thoroughly. In the previous chapters, we felt it necessary to introduce the modeling concepts using linear systems, since there are a lot of new concepts to learn, including system construction rules, power convention, causality, and equation derivation. While linear systems play a very large part

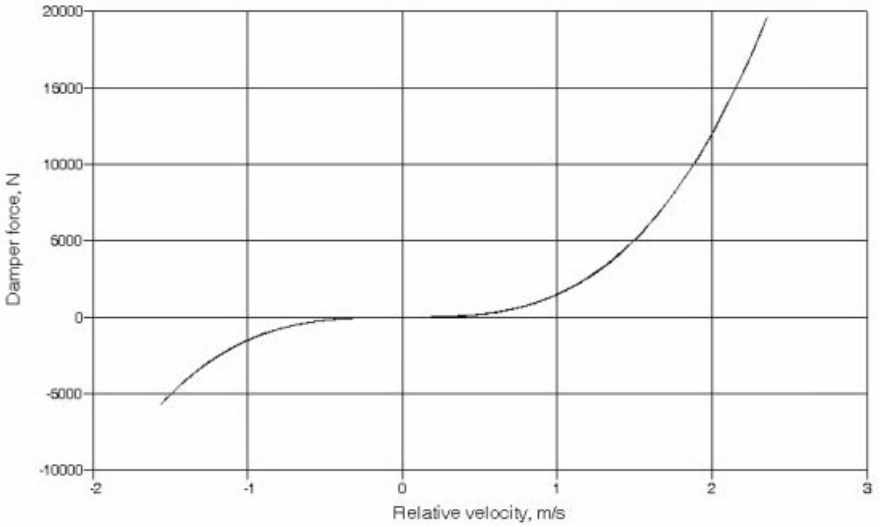


FIGURE 5.19. Damper force versus relative velocity across the suspension.

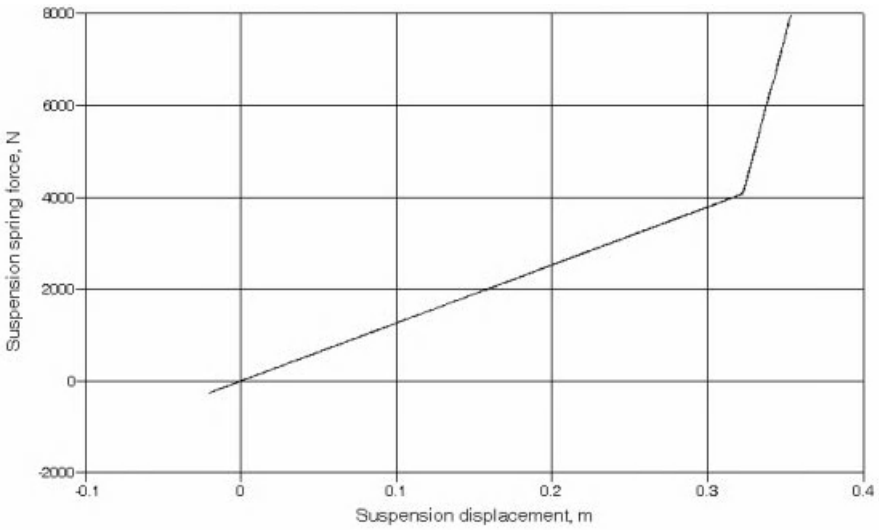


FIGURE 5.20. Suspension spring force versus suspension deflection.

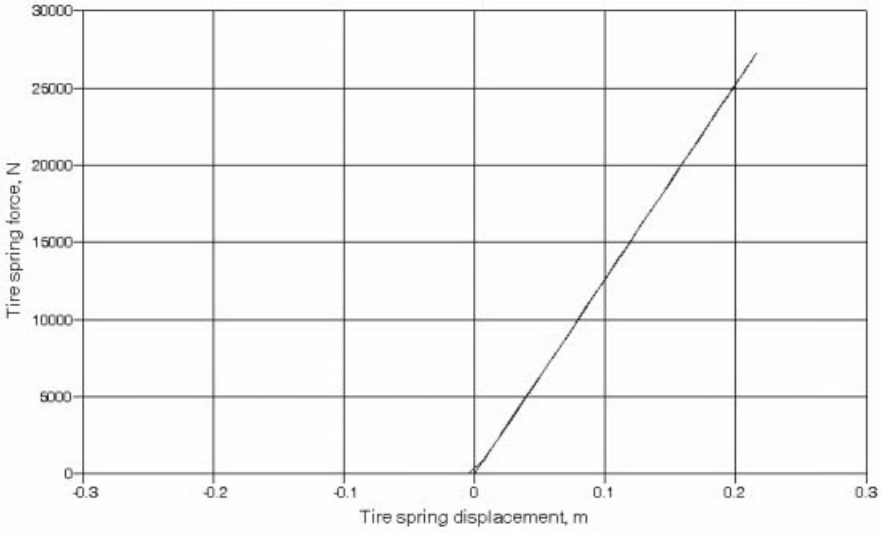


FIGURE 5.21. Tire spring force versus tire deflection.

in understanding the basic behavior of systems, the fact is that most systems are nonlinear, and, at some point in the development of an engineering system, the nonlinear behavior will have to be included. The nonlinear model of the system will probably be computer simulated since there is almost no chance that analytical methods can be useful.

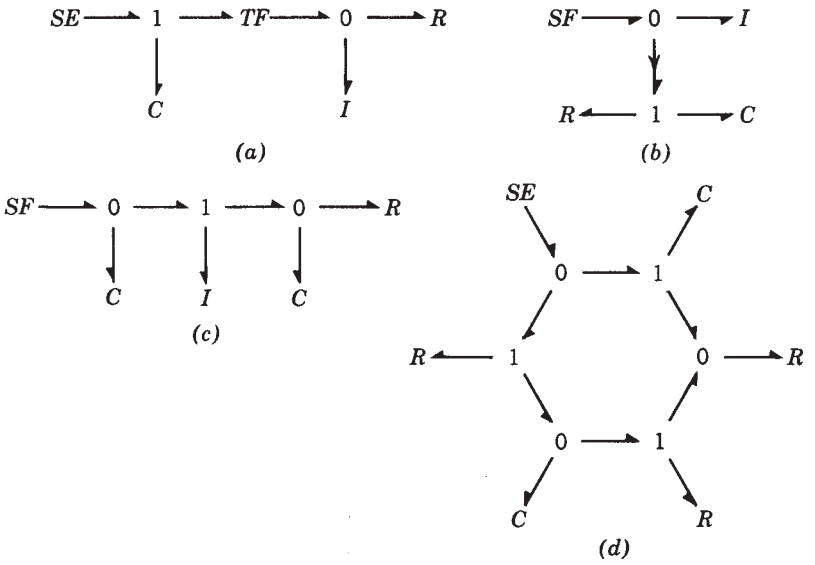
In Chapter 6, some of the more important linear system analysis tools are presented. These linear techniques are important in understanding system dynamics and designing control systems and they are widely used even when they are not strictly appropriate. Later chapters are devoted to advanced bond graph methods for constructing realistic overall system models that include nonlinear constitutive laws when appropriate.

REFERENCE

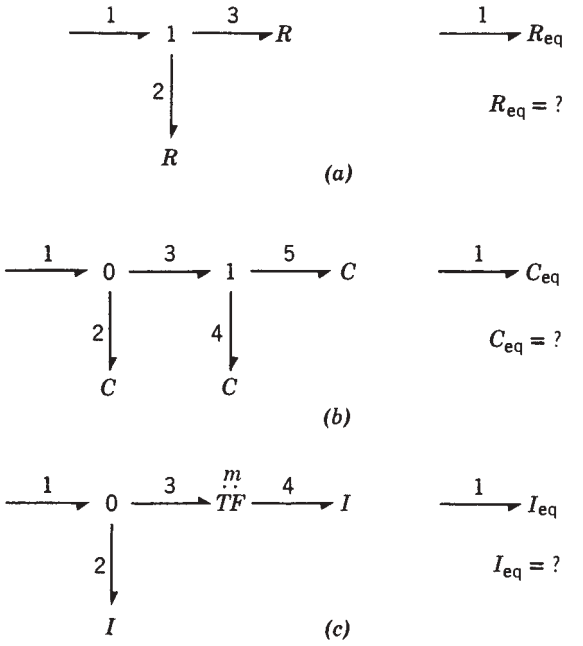
[1] R. C. Rosenberg, "State space formulation for bond graph models of multiport systems," *Trans. ASME J . Dyn. Syst. Meas. Control*, **93**, Ser. G, No. 1, 35–40 (Mar. 1971).

PROBLEMS

5-1. For each of the following bond graphs, assign causality, predict the number of state variables, and write a set of state equations. All elements may be assumed to be linear with constant coefficients.



5-2. For each of the problems below, you are asked to find an equivalent expression for the system given. Try it first for constant-coefficient elements and then for nonlinear characteristics using the first case as a guide.



5-3. For the problems listed below, in each case make a bond graph model of the circuit, schematic, or network diagram. Then augment the graph and write state equations. Interpret the equations physically. You may assume the elements are linear.

(a) Problem 4-1(e), electrical circuit

(b) Problem 4-3(h), mechanical translation

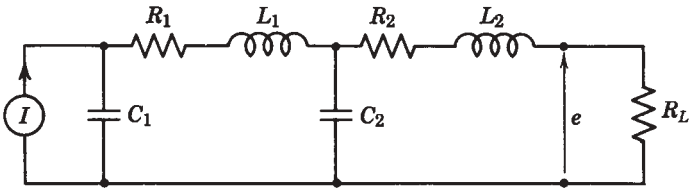
(c) Problem 4-5(a), mechanical rotation

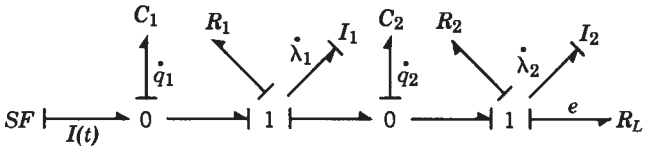
(d) Problem 4-6(a) and (b), hydraulic

(e) Problem 4-5(c), mechanical transduction

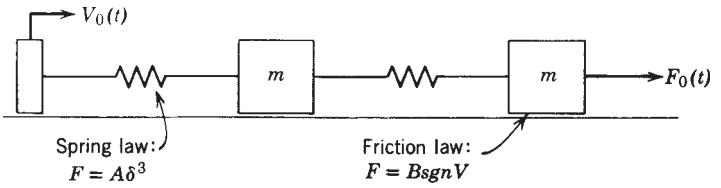
(f) Problem 4-5(e), mechanical transduction

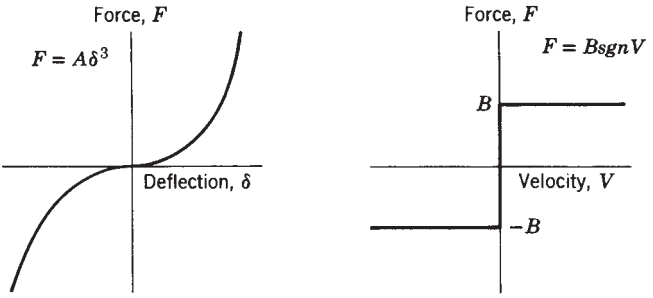
5-4. For the electrical circuit shown below, verify the bond graph, write state equations, and develop an output equation for the voltage across the load resistor (R_L) in terms of state variables and input variables. Assume all elements have constant coefficients.





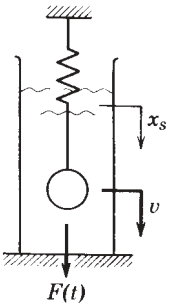
5-5. The mechanical system shown below has two nonlinear springs with constitutive laws, as shown. Note that δ is a deflection. The friction relations are also nonlinear, with a signum (sign) characteristic as shown. Make a bond graph and write state equations.





5-6. Make a bond graph model of the pulley system shown in part b of Problem 4-4. Augment the graph and predict the number of state variables. Write a suitable set of state equations. Modify the graph to include pulley inertias and pin-joint friction. Write state equations for the resulting graph.

5-7. A ball is suspended at the end of a spring in a container of fluid, as shown below. The extension of the spring measured from free length is x_s, and the velocity of the mass is v, measured downward. The spring is nonlinear, characterized by the relation



$$F_s = \phi_s(x_s).$$

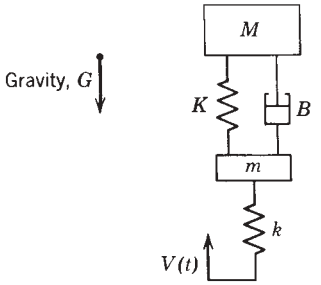
The damping effect of the fluid is proportional to the square of the velocity, corrected for sign:

$$F_D = b\,|v|\,v.$$

(a) Formulate state equations for the system in terms of x_s and p (mass momentum).

(b) Transform the equations from x_s, p variables to x_s, v variables.

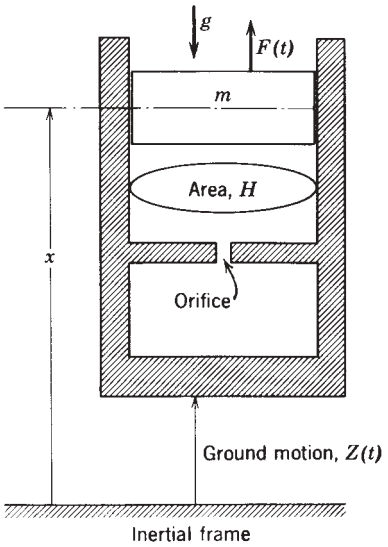
5-8. Consider the simple model of a vehicle shown below, where the suspended mass is M; the main suspension stiffness is K; the damper coefficient is B; the tire mass is m; the tire spring stiffness is k; and the velocity input from the roadway is $V(t)$.



(a) Make a bond graph model, augment the graph, and write state equations.

(b) The tire mass is assumed to be negligible: let $m \to 0$, but retain all other values. Make a bond graph and derive state equations.

5-9. A commercial type of air spring isolator is shown. The mass is supported by air pressure contained by means of a virtually frictionless bellows arrangement (not shown). Damping is provided by adjusting the area of the orifice between the two chambers. The orifice resistance law is generally nonlinear, as is the capacitance relation, since the pressure–volume relation for the gas is approximately $PV^n = \text{const}$.



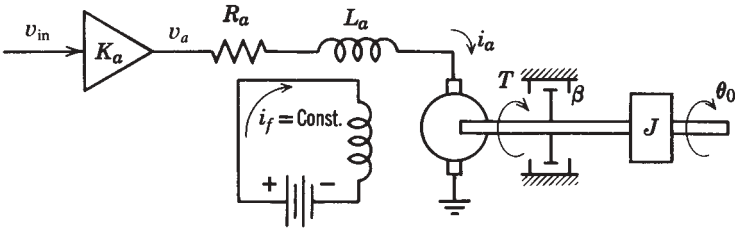
(a) Make a bond graph for the system that allows both ground motion $Z(t)$ and force $F(t)$ as inputs.

(b) Write state equations, using general functions for resistance and capacitance relations.

(c) Using the bond graph, show an analogous all-mechanical system.

5-10. For the water storage system of Problem 4-7, including the inertia effects introduced in part b, augment the graph, predict the number of state variables, and write state equations. (If you encounter formulation difficulties, indicate clearly the steps to be taken to achieve state equations in standard form.)

5-11. On the diagram, physical-system variables and parameters are identified, where



θ_0 = output position angle;

v_{in} = input voltage;

v_a = output voltage of linear amplifier;

i_a = motor armature current;

i_f = motor field current, assumed constant;

K_a = gain of linear amplifier, assumed to have no significant time constants;

R_a = resistance of armature winding;

L_a = inductance of armature winding;

J = inertial load;

b = viscous-damping constant;

K_T = torque constant of motor;

K = back-emf constant of motor.

The differential equations that govern the dynamics of the system are

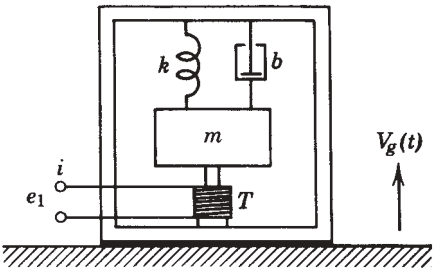
$$J\ddot{\theta}_0 + \beta\dot{\theta}_0 = K_T i_a, \quad L_a\dot{i}_a + R_a i_a = V_a - K_v\dot{\theta}_0.$$

(a) Construct a bond graph for the system.

(b) Write state-space equations, and verify that they are equivalent to those listed above.

(c) Compare the two methods for analyzing this system. For example, is the system third or second order? Are K_T and K_v related in any way?

5-12. Consider the seismometer sketched below:



The input is ground motion $V_g(t)$; an electrical transducer using a permanent magnet moving in a coil reacts to the relative motion between the case and the seismic mass m.

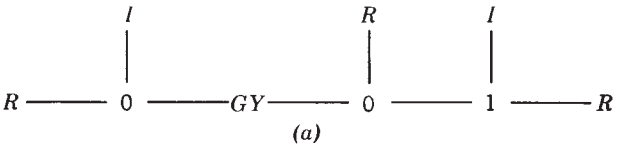
(a) Construct a bond graph model of the device, leaving the electrical port as a free bond.

(b) Assume the device is connected to a voltage amplifier, so that $i = 0$. Find an expression relating $V_g(t)$ and e_1.

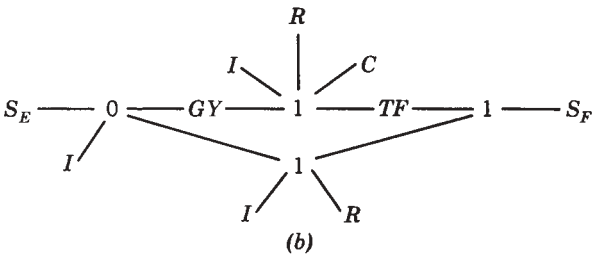
(c) Sometimes it is preferable to use current rather than voltage as output, due to noise considerations. Suppose a current amplifier is used, so that $e_1 = 0$. Relate output current i to input signal $V_g(t)$ for the case when coil resistance and inductance are neglected.

(d) Reconsider the problem of (c) with coil resistance included.

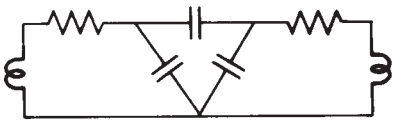
5-13. The three systems shown below each possess a feature that makes the writing of state-space equations interesting. In each case assign causality sequentially and predict any difficulties. Where possible, write state-space equations assuming linear element characteristics.



(a)

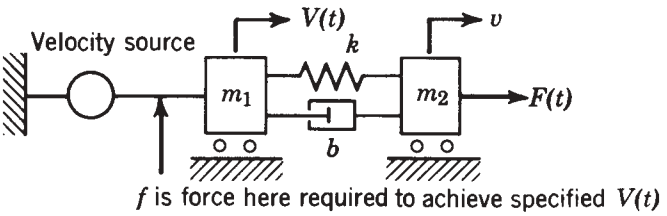


(b)



(c)

5-14. Consider the mechanical system shown below.

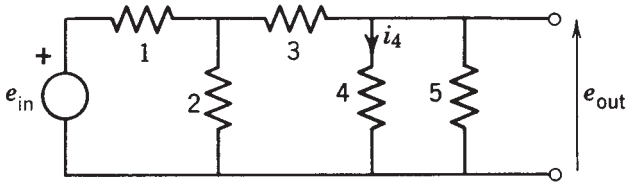


The mass, m, plays a somewhat unusual role in the system.

(a) Show that a state space for the system can be found that does not even depend on m_1.

(b) Show that most system output variables depend statically on the state variables and the inputs, F and V, but because of m_1 this is *not* true of f (defined above).

5-15. The resistive circuit shown below offers an opportunity to exploit the use of causality and helping (or auxiliary) variables.



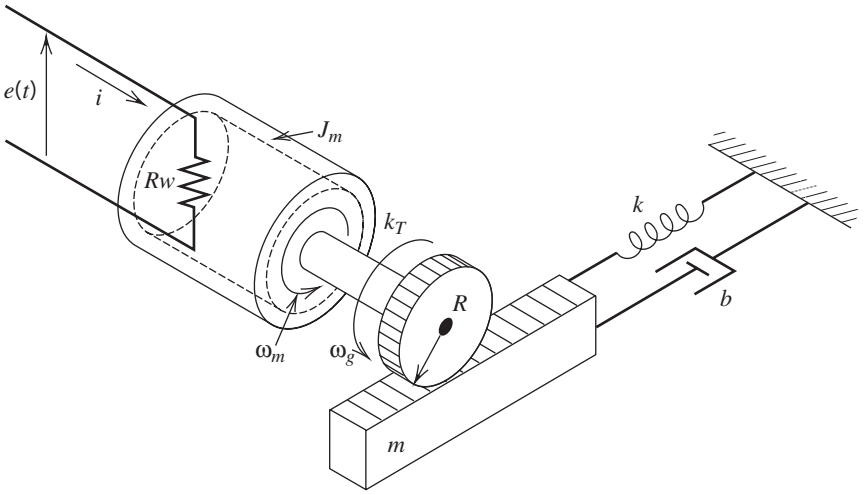
(a) Make a bond graph model of the circuit, and assign causality.

(b) Find e_{out} in terms of e_{in}.

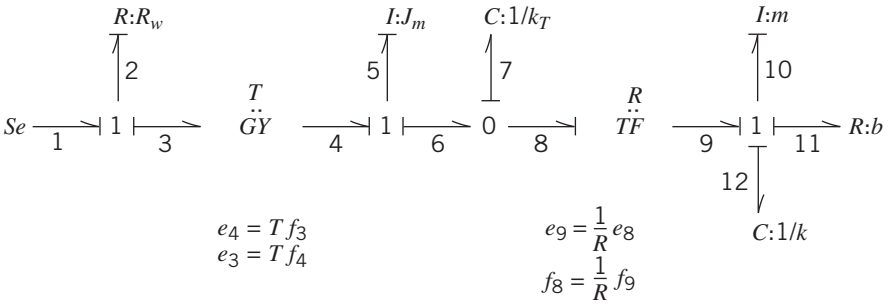
(c) Find i_4 in terms of e_{in}.

(d) If a load resistance R_L is put on at the e_{out} port, modify the graph and your solution to part (b) above.

5-16. The dc motor shown has winding resistance R_w, rotary inertia J_m, and output shaft compliance k_T. The pinion gear of radius R drives a rack of mass m with attached spring k and damper b. The bond graph for this system is shown below, where T is the motor constant such that the motor torque τ is related to the current i by $\tau = Ti$. Derive a state-space representation of this system and put into matrix form, $\dot{x} = Ax + be$:



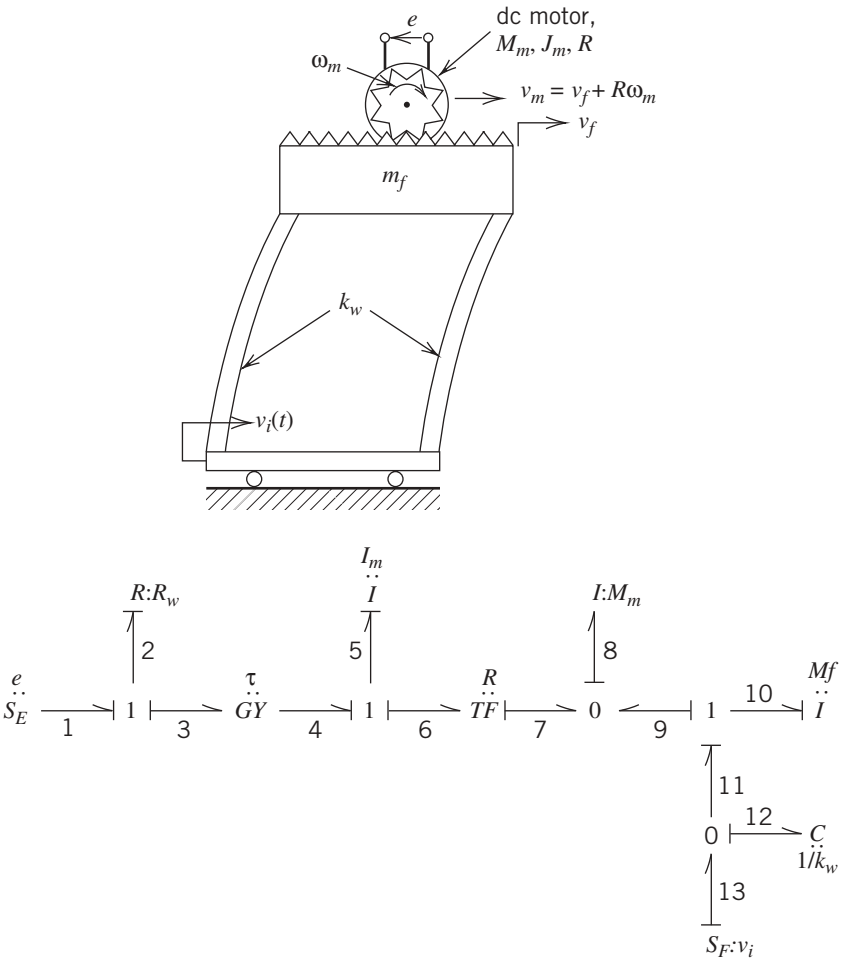
(a)



(b)

Write output equations for e_4, f_7, and e_2.

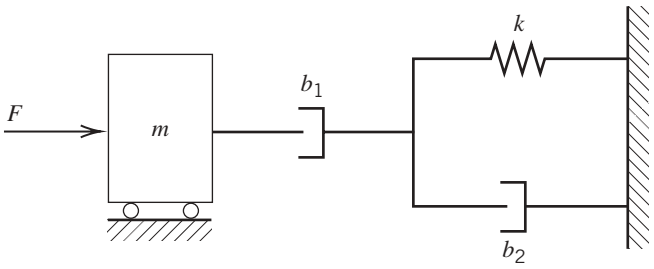
5-17. A one-story building model is shown with a dc motor on the "roof." The motor can be accelerated such that a reaction force is generated. The idea would be to control the motor voltage so that the reaction force helps calm the building in the event of an earthquake, represented by input velocity $v_i(t)$.



The bond graph for this system has derivative causality. Using procedures from this chapter, derive the state equations for this system. The algebra may become tedious, but describe with equations and words how you would come up with final state equations.

5-18. For Problem 5-17, propose the addition of a physical energy storage element that eliminates derivative causality. Redraw the bond graph and describe what the added element represents. Derive the state equations.

5-19. The system shown has an algebraic loop:



(1) Construct a bond graph model, assign causality, and expose the algebraic loop.

(2) Make an arbitrary causal assignment and perform the algebra necessary to allow derivation of state equations. Derive the state equations.

(3) Add a physical energy storage element that will eliminate the algebraic loop, and redraw the bond graph. Derive the system equations.

6

ANALYSIS AND CONTROL OF LINEAR SYSTEMS

6.1 INTRODUCTION

In the past, a major reason for studying linear models of systems was convenience, both in the model formulation and in the solution of the resulting differential equations. But in this book, we have taken pains to point out that with the use of bond graph methods, nonlinear models are not much more difficult to formulate than linear ones. This is particularly true when only the 1-port elements have nonlinear characteristics. Chapters 7, 9, and 12 demonstrate that some types of physical systems have even more complicated types of nonlinearities, but even in these cases bond graph methods result in equation sets that can be handled by computer simulation, as discussed in Chapter 13.

An important reason for the study of linear system models is that complete analytical solutions to linear differential equations exist, and this allows one to develop an intuitive feeling for system dynamics in general that is hard to acquire if only computer solutions are available. Everyone dealing with system dynamics is expected to be familiar with concepts relating to eigenvalues, free response, transfer functions and frequency response, even though they do not strictly apply to real (nonlinear) systems.

In addition, there are many fields of engineering, such as acoustics, electric circuit design, structural vibration, and the design of actuators of various types, in which linear models have proved to be particularly effective. Moreover, the theory and practice of automatic control have traditionally been based on linear system theory. So there are good reasons for taking time in this chapter to review linear system theory and its application to automatic control.

Since linear ordinary differential equations are treated in virtually all engineering and scientific curricula, we will not attempt to present a mathematically

Dean C. Karnopp, Donald L. Margolis and Ronald C. Rosenberg
Copyright © 2012 John Wiley & Sons, Inc.

rigorous and complete account here. Our purpose is rather to point out some features of the theory of linear ordinary differential equations that are significant for the understanding and design of dynamic systems and methods of controlling such systems so as to perform in a desirable manner.

Although a system model can, in principle, be described by a single high-order differential equation as mentioned in Chapter 5, in this chapter we concentrate on equivalent sets of first-order equations. As shown in Chapter 5 this *state equation* form is produced naturally by bond graph models, and it is particularly suitable for machine computations. When the model is linear, it is convenient to use a vector matrix form as a way of representing the equations. In fact, matrices will be used mainly as a convenient notation to describe system models of any order. Only minimal knowledge of matrix operations will be assumed.

It is still quite common for books on system dynamics or automatic control to introduce Laplace transforms for solving constant coefficient linear differential equations (see Reference [1]). This technique changes a differential equation problem to an algebraic problem involving the Laplace variable s instead of time, t. At one time, the use of Laplace transform techniques was one of the convenient ways to solve linear differential equations analytically. But with the advent of computer simulation it has become easy to solve linear equations numerically, so now this aspect of Laplace transform methods is of diminished utility.

In the present exposition an alternative approach will be used based on the general assumption that solutions are of the form Ae^{st} for special values of the variable s. The results derived below are the same as if Laplace transforms had been used, but the approach seems more fundamental and avoids the necessity of discussing the transform techniques in detail.

6.2 SOLUTION TECHNIQUES FOR ORDINARY DIFFERENTIAL EQUATIONS

Before discussing the characteristics of linear state equations, it is useful to note that it is not difficult in principle to find one or more solutions to general state equations, using a computer if necessary. In most of the examples encountered so far, the physical system model yielded state equations of the form

$$\begin{aligned}
\dot{x}_1 &= f_1(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r),\\
\dot{x}_2 &= f_2(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r),\\
&\vdots\\
\dot{x}_n &= f_n(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r),
\end{aligned} \tag{6.1}$$

where $x_1, \ldots, x_n$ are *state variables* (typically p's or q's in bond graph terms) and $u_1, \ldots, u_r$ are input variables (typically e's or f's from sources).

In addition, there may be *output variables* $y_1, \ldots, y_s$ statically related to the state and input variables:

$$\begin{aligned} y_1 &= g_1(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r), \\ y_2 &= g_2(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r), \\ &\vdots \\ y_s &= g_s(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r). \end{aligned} \tag{6.2}$$

In general, the f and g functions are nonlinear, but we will mainly consider linear functions. The linear versions of Eqs. (6.1) and (6.2) are

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_{11}u_1 + b_{12}u_2 + \cdots + b_{1r}u_r, \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_{21}u_1 + b_{22}u_2 + \cdots + b_{2r}u_r, \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_{n1}u_1 + b_{n2}u_2 + \cdots + b_{nr}u_r, \end{aligned} \tag{6.1a}$$

and

$$\begin{aligned} y_1 &= c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n + d_{11}u_1 + d_{12}u_2 + \cdots + d_{1r}u_r, \\ y_2 &= c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n + d_{21}u_1 + d_{22}u_2 + \cdots + d_{2r}u_r, \\ &\vdots \\ y_s &= c_{s1}x_1 + c_{s2}x_2 + \cdots + c_{sn}x_n + d_{s1}u_1 + d_{s2}u_2 + \cdots + d_{sr}u_r. \end{aligned} \tag{6.2a}$$

Instead of writing out all the a, b, c, and d coefficients for linear systems, as in Eqs. (6.1a) and (6.2a), one can write the equations in a condensed vector matrix form as follows:

$$[\dot{x}] = [A][x] + [B][u], \tag{6.1b}$$

$$[y] = [C][x] + [D][u]. \tag{6.2b}$$

A comparison of this symbolic form with the explicit forms, Eqs. (6.1a) and (6.2a), should make it clear that $[x]$ is an n-dimensional column vector of the state variables, $[\dot{x}]$ is a vector of the derivatives of the state variables, $[u]$ is an r-dimensional column vector of the input variables, $[A]$ is an $n \times n$ matrix of the a_{ij} coefficients, $[B]$ is an $n \times r$ matrix of the b_{ij} coefficients, $[C]$ is an $s \times r$ matrix of the c_{ij} coefficients, and $[D]$ is an $s \times r$ matrix of the d_{ij} coefficients. All that one needs to know at this point is that *row by column* multiplication of matrices and column vectors in Eqs. (6.1b) and (6.2b) yields the explicit forms of Eqs. (6.1a) and (6.2a).

In this chapter we will be concerned with the special case of system equations that not only are linear but also have constant coefficients. Occasionally, the

situation arises in which linear equations have coefficients that vary in time. In matrix form this could be represented as

$$[\dot{x}] = [A(t)][x] + [B(t)][u]. \tag{6.3}$$

The theory of systems in which some of the coefficients vary in time is quite complicated, and since such equations do not arise often, we will not deal with them here.

Equations (6.1), (6.1a), (6.1b), (6.2), (6.2a), and (6.2b) describe *state-determined systems*. To completely specify a problem to be solved, more information than just the equations is needed. Typically, the input variables are specified as functions of time:

$$\begin{aligned} u_1 &= u_1(t), \\ u_2 &= u_2(t), \\ &\vdots \\ u_r &= u_r(t). \end{aligned} \tag{6.4}$$

In addition, a particular trajectory in state space must be singled out for attention. One main concern will be with *initial-condition problems* in which the state at some initial time, t_0, is known:

$$\begin{aligned} x_1(t_0) &= x_{10}, \\ x_2(t_0) &= x_{20}, \\ &\vdots \\ x_n(t_0) &= x_{n0}, \end{aligned} \tag{6.5}$$

where x_{10} to x_{n0} are the initial states. Note that given the information in Eqs. (6.4) and (6.5), the initial output variables are determined; for example,

$$y_1(t_0) = g_1[x_{10}, x_{20}, \ldots, x_{n0}; u_1(t_0), u_2(t_0), \ldots, u_r(t_0)].$$

One may also determine $\dot{x}_1(t_0)$ to $\dot{x}_n(t_0)$ using Eq. (6.1); for example,

$$\dot{x}_1(t_0) = f_1[x_{10}, x_{20}, \ldots, x_{n0}; u_1(t_0), u_2(t_0), \ldots, u_r(t_0)].$$

In principle, it is easy to compute how the system changes in a short interval of time, Δt, using the concept of a derivative directly. For example,

$$x_1(t_0 + \Delta t) \cong x_1(t_0) + \dot{x}_1(t_0)\Delta t = x_{10} + \dot{x}_1(t_0)\Delta t. \tag{6.6}$$

This equation is really just a rearrangement of the definition of a derivative,

$$\dot{x}_1 \equiv \frac{dx_1}{dt} = \lim_{\Delta t \to 0} \left[\frac{x_1(t_0 + \Delta t) - x_1(t_0)}{\Delta t} \right],$$

in which Δt is small but finite. Since $\dot{x}_1(t_0)$ depends on the initial state and the u's at t_0 that are known, one can use Eq. (6.6) to find the state at $t_0 + \Delta t$.

This is sometimes called *Euler's formula* for integrating equations, and it may be applied to all the state equations at $t = t_0$. It can be reapplied at $t = t_0 + \Delta t$, since the u's are known at any time, and after applying Eq. (6.6) at t_0, the state at $t_0 + \Delta t$ is then known. The equation may then be applied recursively to march the solution along in time as far as desired. The process described above is readily adapted to automatic digital computation, but there is a snag. Equation (6.6) is not exactly correct except in the limit as $\Delta t \to 0$. As Δt is made very small, the approximation generally becomes increasingly accurate, but it takes more steps and more evaluations of the functions to cover any given time span of interest. The problem of choosing suitable time increments Δt and recursion formulas analogous to Eq. (6.6) to optimize digital integration accuracy falls in the domain of *numerical methods*, and there is a wealth of literature on this subject. Here it suffices to note that methods to integrate state equations exist in profusion and that, despite its obviousness, Euler's formula often turns out to be quite inefficient compared to other techniques.

The integration of state equations is straightforward, but when the equations are linear, one may say a great deal about the totality of possible systems behavior without the necessity of finding many solutions for various initial conditions and forcing functions. As will be demonstrated, linear systems obey the *principle of superposition*, which states that any scaled sum of solutions to linear state equations is also a solution. The superposition property allows one to consider a particular solution to the state equations to be the sum of a *response due to initial conditions*, or *free response* (in which all input variables vanish), and a *response due to inputs* or *forced response*. Additionally, if we know the forced response of a linear system to several inputs, each acting one at a time, then the solution when all inputs are acting together is just the sum of each individual response. Also, since the free-response component of the total system response has to do only with the $[A]$ matrix in the state equations, this response may be used to characterize the system in a useful manner. We may discover natural frequencies, time constants, and stability properties that, in contrast to the general case, have nothing to do with initial conditions or forcing terms and help in understanding how the system would behave in a variety of situations.

There are three extremely important linear system concepts that contribute the most to our understanding the dynamics of physical systems and our ability to develop controllers that can change a systems natural behavior. These concepts are associated with the words *eigenvalue*, *transfer function*, and *frequency response* and they are presented in the next sections.

6.3 FREE RESPONSE AND EIGENVALUES

Consider now the linear case represented in the form of Eqs. (6.1a) and (6.2a) or in the vector matrix form of Eqs. (6.1b) and (6.2b) and assume that the input variables are all zero. When a solution for this case is found, it will be called the *free response* and it will have to do only with initial conditions, not input

variable forcing. We consider a first-order system at the onset and then move on to an example with multiple state variables.

6.3.1 A First-Order Example

Figure 6.1 shows a simple model of an elastomeric mount commonly used to provide isolation between a load (not shown) and the ground. The mount exhibits some spring-like effect and some damping-like effect and these are represented by the linear spring with spring constant k and the damper with constant b. The applied force at the top F_{in} completes this simple model. The bond graph for this system is also shown in Figure 6.1 and causality dictates that there is one state variable q, the displacement across the spring.

The state equation for this first-order system is

$$\dot{q} = -(k/b)q + (1/b)F_{\text{in}}. \tag{6.7}$$

The *free response* of the system does not depend on the external forcing so we let $F_{\text{in}} = 0$ in order to pursue the free response. Thus, the governing equation simplifies to

$$\dot{q} = -(k/b)q. \tag{6.8}$$

For all free-response equations such as Eq. (6.8), there is a *trivial solution* in which the state variables as well as their derivatives are all zero. (In this case, the trivial solution to Eq. (6.8) is obviously $\dot{q} = q = 0$.) A more general assumption that proves useful for all linear systems is that the solution is exponential, for example,

$$q(t) = Qe^{st}, \tag{6.9}$$

where the constants Q and s are to be adjusted, if possible, to make sure that $q(t)$ fits the differential equation and the initial conditions.

Substituting Eq. (6.9) into Eq. (6.8) (noting that $\dot{q} = Qse^{st}$), we find

$$Qse^{st} = -(k/b)Qe^{st} \tag{6.10}$$

or

$$(s + k/b)\, Qe^{st} = 0. \tag{6.11}$$

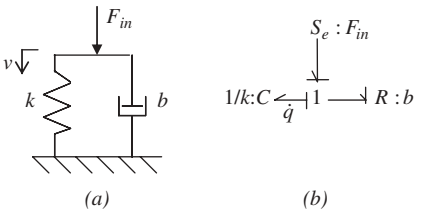


FIGURE 6.1. First-order example system: (*a*) schematic diagram; (*b*) bond graph.

For a solution to exist, either

$$Qe^{st} = 0,$$

in which case we return to the trivial solution, or else

$$s + k/b = 0 \quad \text{or} \quad s = -k/b. \tag{6.12}$$

Equation (6.12) is called the *characteristic equation* and the value of s that is a solution to the characteristic equation is called the *eigenvalue*.

The solution of Eq. (6.9) using the eigenvalue of Eq. (6.12) is

$$q(t) = Qe^{-(k/b)t} = Qe^{-t/\tau}, \tag{6.13}$$

where the *time constant*, τ, has been defined as $\tau = b/k$ seconds. (The time constant must have the same units as t so that the exponent $-t/\tau$ is dimensionless, as it must be.)

To complete the solution, the constant Q can be found from the initial value of the displacement $q(0) = q_0$:

$$q(0) = Qe^{-0} = Q, \tag{6.14}$$

where the constant $q(0)$ is simply the value of the spring displacement at $t = 0$. The solution is then

$$q(t) = q(0)e^{-t/\tau}, \qquad t \geq 0. \tag{6.15}$$

This solution is shown in Figure 6.2. For this simple exponential solution, it is readily shown that, at any point on the curve, if one extends a straight line along the local slope of the curve to the steady-state value of q (zero in this case), it will intersect after τ time units have elapsed. Thus, τ sets the basic timescale of the response. If a system response is measured experimentally, one can easily estimate τ as shown in Figure 6.2 and one can check to see how well a linear, first-order constant-parameter system would reproduce the measured response

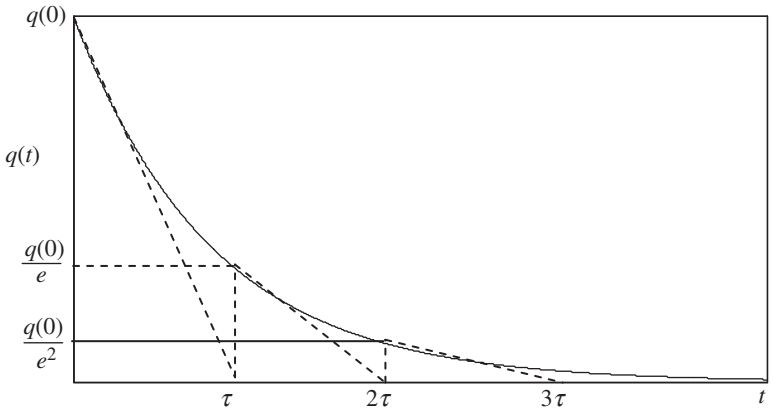


FIGURE 6.2. Free response of the system in Figure 6.1.

by computing τ at several points on the experimental curve. In the parlance of system dynamics we say that the response is 63% complete after 1 time constant ($e^{-1} = 0.37$) and virtually complete after about 4 time constants ($e^{-4} \approx 0.02$). Theoretically, of course, $q(t)$ would continue to approach zero exponentially for a truly linear system but for a real physical system, nonlinear effects will become important for very small values of q.

6.3.2 Second-Order Systems

Although knowledge of first-order system response will prove very useful, it must be obvious that more complex systems respond in ways much more complicated than the exponential of Figure 6.2. Surprisingly, the pattern of analysis, somewhat generalized, does carry over for arbitrarily high-order systems. In fact, understanding the response of a first-order system and then the response of a second-order system, the pattern will become evident, and the extension to nth-order systems is straightforward.

Consider, then, the simple second-order example shown in Figure 6.3. The spring and damper connecting the cart to the load mass could very well be the mount from Figure 6.1. To make the correspondence between the system from Figure 6.3 and the general equations (6.1a) and (6.2a), we write the state equations in a standard form using causality and the formulation procedure from Chapter 5. Thus,

$$\begin{aligned} \dot{x} &= 0x - (1/m)p + 1V + 0F, \\ \dot{p} &= kx - (b/m)p + bV + 1F. \end{aligned} \tag{6.16}$$

These equations are shown in Figure 6.3*c*.

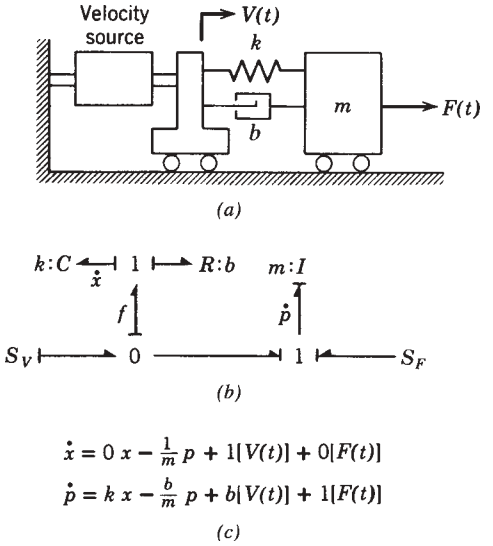


FIGURE 6.3. Example system: (*a*) schematic diagram; (*b*) bond graph; (*c*) state equations.

The force f, which is the sum of the forces in the spring and damper, is just one of several possible output variables. Following the causal information in the bond graph yields

$$f = kx - (b/m)p + bV + 0F. \tag{6.17}$$

The system has two state variables, $x_1 = x$, $x_2 = p$; two input variables, $u_1 = V$, $u_2 = F$, and one output variable, $y_1 = f$. The matrices for the forms of Eqs. (6.1b) and (6.2b) are then related to the system parameters as follows:

$$[A] = \begin{bmatrix} 0 & -1/m \\ k & -b/m \end{bmatrix}, \quad [B] = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}, \quad [C] = [k \quad -b/m], \quad [D] = [b \quad 0]. \tag{6.18}$$

Note that in Figure 6.3*c* the notations $V(t)$ and $F(t)$ are used as a reminder that these input quantities must be specified functions of time before the state equations can be solved.

For the free response and to derive the eigenvalues there is no dependence on the external inputs so we let $V(t) = F(t) = 0$, such that

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} 0 & -1/m \\ k & -b/m \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}. \tag{6.19}$$

In general terms, this equation could be written simply as $[\dot{x}] = [A][x]$, which is Eq. (6.1b) with $[u] = [0]$.

The *trivial solution* to Eq. (6.19) is evident. It is

$$x \equiv 0, \quad p \equiv 0. \tag{6.20}$$

This solution corresponds to the case in which the spring is unstretched and the mass has no momentum (or velocity). In other words, a possible solution for the free response of a linear system is that it does absolutely nothing. As in the first-order system example, this solution is, again, not very useful and is of no help in fitting given initial conditions. Suppose a solution satisfying

$$x(t_0) = x_0, \quad p(t_0) = p_0 \tag{6.21}$$

is desired, where x_0 is the initial stretch in the spring and $p_0 = mv_0$ is the initial momentum of the mass.

To find the solution, assume that both energy-state variables may be represented by an amplitude times an exponential time function. In the example, we assume

$$x(t) = Xe^{st}, \quad p(t) = Pe^{st}, \tag{6.22}$$

where X and P are constant amplitudes and s is a number with the dimension of inverse time. (A complication not evident in the first-order example is that X, P, and s all may turn out to be complex numbers.) If the assumption of Eq. (6.22)

is substituted into Eq. (6.19), the result is

$$\begin{bmatrix} sX \\ sP \end{bmatrix} e^{st} = \begin{bmatrix} 0 & -1/m \\ k & -b/m \end{bmatrix} \begin{bmatrix} X \\ P \end{bmatrix}. \tag{6.23}$$

The left side of Eq. (6.23) can be written in a form similar to the right side by using the 2×2 *unit matrix*

$$[I] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6.24}$$

such that

$$\begin{bmatrix} sX \\ sP \end{bmatrix} e^{st} = s \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ P \end{bmatrix} e^{st} = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} X \\ P \end{bmatrix} e^{st}. \tag{6.25}$$

With this idea, the two sides of Eq. (6.23) can be combined. A single matrix is then involved.

$$\begin{bmatrix} s & 1/m \\ -k & s + b/m \end{bmatrix} \begin{bmatrix} X \\ P \end{bmatrix} e^{st} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{6.26}$$

We can see now that the single matrix on the left side of Eq. (6.26) is just $[sI - A]$. This is a very important matrix that is composed simply of the negative of the $[A]$-matrix with $[sI]$ adding s terms along the main diagonal. This matrix will be important for systems of any order when analyzing the free response where $[I]$ is the general $n \times n$ unit matrix with 1's on the main diagonal and 0's elsewhere, and $[A]$ is the system matrix from Eq. (6.1b).

Now Eq. (6.26) shows that the assumption of Eq. (6.22) has reduced the differential equation problem to an algebraic problem. (The Laplace transformation also achieves this reduction.) One way to solve Eq. (6.26) would be to have e^{st} vanish. Not only is this not really possible for any finite value of s, but the result leads only to the trivial solution of Eq. (6.20). Rejecting this possibility, one can divide out the e^{st} term and concentrate on solving for X and P that satisfy

$$\begin{bmatrix} s & 1/m \\ -k & s + b/m \end{bmatrix} \begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{6.27}$$

Once again, the *trivial solution* appears, since clearly $X = 0$, $P = 0$ is a solution to Eq. (6.27). The elementary theory of linear algebra contains the result that if the determinant of the matrix $[sI - A]$ is not zero, there is a single unique solution. Since we know that there always is a trivial solution, we look for the condition for which *nonzero* values of X and P can result. This requires that the determinant of the matrix $[sI - A]$ *should be zero*. To continue with our pursuit of a nontrivial solution, then, we require

$$\text{Det} \begin{vmatrix} s & 1/m \\ -k & s + b/m \end{vmatrix} = 0 \tag{6.28}$$

or

$$s^2 + (b/m)s + (k/m) = 0. \tag{6.29}$$

The requirement that this determinant vanish leads to the requirement that a polynomial in s must vanish. Equation (6.29) is the *characteristic equation* for this system and leads to a set of *eigenvalues or characteristic values* for s. Equation (6.6) was the characteristic equation for our first example. (In German, a characteristic value is called an *eigenwert*, and in English the mongrel term *eigenvalue* is commonly used. Why "characteristic wert" never caught on remains a mystery.) In what follows, we will generally use the term *eigenvalue* in place of the words *characteristic value*.

Solving Eq. (6.29), we find two eigenvalues for this second-order system:

$$\begin{aligned} s_1 &= -(b/2m) + [(b/2m)^2 - (k/m)]^{1/2}, \\ s_2 &= -(b/2m) - [(b/2m)^2 - (k/m)]^{1/2}. \end{aligned} \tag{6.30}$$

Note that s_1 and s_2 will be complex for small values of the damping constant b since then Eq. (6.30) will involve the square root of a negative number. For each value of s that is determined from an equation such as Eq. (6.29), a nonzero solution satisfying Eq. (6.27) may be found. For an nth-order system, there will be n eigenvalues. In our second-order example, there are two. For high-order systems it will generally not be possible to solve for the characteristic values without resorting to numerical methods, but the existence of n roots of an nth-order characteristic equation is ensured by the theory of linear algebra.

Much useful information is contained in the eigenvalues of Eq. (6.30) themselves, but before discussing what insights can be attained by simply solving the characteristic equation, let us outline the procedure for finding a complete free response satisfying given initial conditions.

Having pursued the solution to this point, the next question might be: Since we used the condition for there to exist nonzero values of X and P, what are these nonzero values? For each value of s, s_1, and s_2, let us attempt to determine X and P values from Eq. (6.27). For s_1 we attempt to find X_1 and P_1 by substituting the eigenvalue back into Eq. (6.27) with the result for s_1:

$$s_1 X_1 + (1/m) P_1 = 0, \qquad -k X_1 + (s_1 + b/m)\, P_1 = 0. \tag{6.31}$$

By all appearances it looks like we have two equations for the two unknowns X_1, P_1 and we should be able to solve for each individually. However, the condition that $\mathrm{Det}[sI - A] = 0$ has additionally guaranteed that the two equations in (6.31) are not independent. In fact, when s_1 is substituted into Eqs. (6.31), they each will be found to contain the same information. Thus, the best we can do is determine one of the constants in terms of the other, thus:

$$X_1 = -(1/ms_1) P_1 \quad \text{or} \quad X_1 = (1/k)[s_1 + (b/m)] P_1. \tag{6.32}$$

If the value of s_1 from Eq. (6.30) is substituted into either of the equations in Eq. (6.32) it will be found that they both give the same relation between X_1 and P_1. Whenever an eigenvalue is substituted into equations for amplitudes, such as Eq. (6.27), the equations will be found *not* to be independent. In the present

example this means that the ratio of X_1 to P_1 in (6.32) is determined, but not the magnitude of either one. In general for higher-order systems, the most one can expect to solve for is the ratios of $n-1$ amplitudes in terms of one arbitrary amplitude.

If we return to Eq. (6.27) and use s_2 from Eq. (6.30) to solve for amplitudes X_2 and P_2, we again find that both equations are equivalent so only the ratio of X_2 to P_2 is determined:

$$X_2 = -(1/ms_2)P_2 \quad \text{or} \quad X_2 = (1/k)[s_2 + (b/m)]P_2. \tag{6.33}$$

Now whenever X_1 and P_1 are related according to Eq. (6.32), Eq. (6.27) will be satisfied, and the original differential equation Eq. (6.26) will be satisfied by

$$x(t) = X_1 e^{s_1 t}, \qquad P(t) = P_1 e^{s_1 t}.$$

Similarly, another solution is

$$x(t) = X_2 e^{s_2 t}, \qquad p(t) = P_2 e^{s_2 t}$$

whenever X_2 and P_2 satisfy Eq. (6.33). It is also easy to verify that another valid free-response solution is obtained by adding the two solutions together:

$$x(t) = X_1 e^{s_1 t} + X_2 e^{s_2 t}, \qquad p(t) = P_1 e^{s_1 t} + P_2 e^{s_2 t}. \tag{6.34}$$

This is an example of the principle of *superposition* for linear systems and is readily verified by simply substituting Eq. (6.34) into Eq. (6.27) and noting that as long as the components of the sum satisfy the differential equations, so does the sum.

Although s_1 and s_2 are completely determined by Eq. (6.30), the four quantities X_1, X_2, P_1, and P_2 are so far related only by the two relations (6.32) and (6.33). Two more conditions are necessary to specify the free response completely, and these are provided by the two initial conditions of Eq. (6.21) and solving:

$$x_0 = x(t_0) = X_1 e^{s_1 t_0} + X_2 e^{s_2 t_0}, \tag{6.35}$$

$$p_0 = p(t_0) = P_1 e^{s_1 t_0} + P_2 e^{s_2 t_0}. \tag{6.36}$$

Thus, the four equations (6.32), (6.33), (6.35), and (6.36) completely specify the free response of the system, satisfying the given initial conditions. Although in this simple case one could actually solve these equations completely, this is rarely attemped now that computer simulation is so easy.

What is important at this point in the development is that the eigenvalues appear as arguments of exponentials and appear in the system response as shown in Eq. (6.34). It is nice to know that we could solve for the complete free response but we are rarely interested in the free response alone. We are much more concerned with the forced response or the total response including both the free response and the response due to forcing. It is understanding the nature of the free response that is important and we pursue this goal in the next section.

6.3.3 Example: The Undamped Oscillator

To gain understanding of the physical system response from Figure 6.3 we first set the dashpot parameter b to zero, leaving a system composed of only a spring and a mass. The characteristic equation (6.29) becomes

$$s^2 + (k/m) = 0, \tag{6.37}$$

which yields the eigenvalues

$$s_1 = +j\,(k/m)^{1/2} = j\omega_n, \qquad s_2 = -j\,(k/m)^{1/2} = -j\omega_n, \tag{6.38}$$

where $j \equiv (-1)^{1/2}$ and the *undamped natural frequency* ω_n is defined as $(k/m)^{1/2}$. In distinction to the first-order example that had a *real* eigenvalue, we now have encountered purely *complex* eigenvalues.

The solution from Eq. (6.34) then becomes

$$x(t) = X_1 e^{j\omega_n t} + X_2 e^{-j\omega_n t}, \tag{6.39}$$

$$p(t) = P_1 e^{j\omega_n t} + P_2 e^{-j\omega_n t} \tag{6.40}$$

where the purely complex eigenvalues end up in complex exponentials in the solution.

Recall, if you can, that the sine and cosine functions can be expressed in terms of complex exponentials as

$$\cos \omega_n t = \frac{e^{j\omega_n t} + e^{-j\omega_n t}}{2} \tag{6.41}$$

$$\sin \omega_n t = \frac{e^{j\omega_n t} - e^{-j\omega_n t}}{2j}, \tag{6.42}$$

yielding the Euler relationships:

$$e^{+j\omega_n t} = \cos \omega_n t + j \sin \omega_n t \tag{6.43}$$

$$e^{-j\omega_n t} = \cos \omega_n t - j \sin \omega_n t. \tag{6.44}$$

Thus, the complex eigenvalues indicate an oscillatory time response at the frequency ω_n rad/s.

If Eqs. (6.43) and (6.44) are substituted into the solution from Eqs. (6.39) and (6.40), it appears that the real response $x(t)$ and $p(t)$ have the complex number $j = \sqrt{-1}$ as part of the solution. It is not difficult to show that while X_1, X_2 and P_1, P_2 are constants, they are complex constants and, in fact, are complex conjugate constants. That is, if

$$X_1 = a_x + jb_x \tag{6.45}$$

then

$$X_2 = a_x - jb_x; \tag{6.46}$$

and similary, if

$$P_1 = a_p + jb_p, \tag{6.47}$$

then

$$P_2 = a_p - jb_p. \tag{6.48}$$

There are still only two constants associated with $x(t)$ and two constants associated with $p(t)$, so nothing has changed with respect to obtaining a solution if one chooses to pursue it.

Using (6.45) through (6.48) with (6.43) and (6.44) in the solution expressions (6.39) and (6.40) will yield the time response as

$$\begin{aligned} x(t) &= A_x \cos \omega_n t + B_x \sin \omega_n t, \\ p(t) &= A_p \cos \omega_n t + B_p \sin \omega_n t, \end{aligned} \tag{6.49}$$

where, if all the algebra was carried out, one would discover:

$$A_x = 2a_x,\ B_x = -2b_x,\ \ A_p = 2a_p,\ \ B_p = -2b_p. \tag{6.50}$$

The real constants A_x, B_x, A_p, B_p would be determined from initial conditions.

The really important result here is that the complex eigenvalues $s_{1,2} = \pm j\omega_n$ indicate an oscillatory response at the frequency ω_n rad/s or a cyclic frequency $f_n = \omega_n/2\pi$ Hz.

Another way to represent Eq. (6.49) is as a single harmonic with a phase angle φ and an amplitude A, thus:

$$x(t) = A \cos(\omega_n t + \varphi) \tag{6.51}$$

and

$$A = \left[A_x^2 + B_x^2\right]^{1/2}, \quad \phi = \tan^{-1}(B_x/A_x). \tag{6.52}$$

There are still just two unknown constants, A and ϕ. A similar expression can be written for $p(t)$.

Another commonly used notation involves the "real part of" operator $\mathrm{Re}(\cdot)$ and the "imaginary part of," operator $\mathrm{Im}(\cdot)$. Sine and cosine functions are readily expressed in terms of these functions as follows:

$$\cos \omega_n t = \mathrm{Re}\left(e^{j\omega_n t}\right), \tag{6.53}$$

$$\sin \omega_n t = \mathrm{Im}\left(e^{j\omega_n t}\right). \tag{6.54}$$

Then $x(t)$ may be expressed in yet another way as

$$x(t) = A\,\mathrm{Re}\left(e^{j(\omega_n t+\phi)}\right) = \mathrm{Re}\left(Ae^{j\phi}e^{j\omega_n t}\right). \tag{6.55}$$

Equation (6.55) is sometimes simpler to use than some other equivalent representations, since it deals with only an amplitude A, the angle ϕ and one of the pair $e^{+j\omega_n t}$, $e^{-j\omega_n t}$.

6.3.4 Example: The Damped Oscillator

When the dashpot coefficient b is positive and not too large, the free response of the system will be oscillatory but will eventually damp out. The eigenvalues of Eq. (6.30) clearly change character, depending on whether the expression under the radical is positive or negative. If

$$(b/m)^2 < 4k/m, \tag{6.56}$$

the system is said to be *underdamped*; if

$$(b/m)^2 > 4k/m, \tag{6.57}$$

the system is said to be *overdamped*; and if

$$(b/m)^2 = 4k/m, \tag{6.58}$$

the system is said to be *critically damped*.

Often the characteristic equation (6.29) is written as follows:

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = 0, \tag{6.59}$$

in which the undamped natural frequency ω_n is as defined by Eq. (6.38), and the *damping ratio* ζ is given as

$$\zeta \equiv b/2m\omega_n = b/2(mk)^{1/2}. \tag{6.60}$$

Then the system is underdamped, overdamped, or critically damped, depending on whether $\zeta < 1$, $\zeta > 1$, or $\zeta = 1$. With these definitions, the eigenvalues for the underdamped case of Eq. (6.30) may be written

$$\begin{aligned} s_1 &= -\zeta\omega_n + j\omega_n(1-\zeta^2)^{1/2}, \\ s_2 &= -\zeta\omega_n - j\omega_n(1-\zeta^2)^{1/2}. \end{aligned} \tag{6.61}$$

We would still interpret the imaginary part of the eigenvalue as a frequency and it is customary to define the *damped natural frequency* as

$$\omega_d \equiv \omega_n\sqrt{1-\varsigma^2}, \tag{6.62}$$

allowing the eigenvalues to be expressed as

$$s_1 = -\zeta\omega_n + j\omega_d, \;\; s_2 = -\zeta\omega_n - j\omega_d. \tag{6.63}$$

The free-response solution, Eq. (6.34), becomes a bit more complicated than in the undamped case because the eigenvalues are now complex numbers rather than pure imaginary numbers. For $x(t)$ the free response looks like

$$x(t) = X_1 e^{(-\varsigma\omega_n + j\omega_d)t} + X_2 e^{(-\varsigma\omega_n - j\omega_d)t}. \tag{6.64}$$

From the rules for multiplication of exponentials, Eq. (6.64) can be written as

$$x(t) = e^{-\varsigma\omega_n t}(X_1 e^{j\omega_d t} + X_2 e^{-j\omega_d t}). \tag{6.65}$$

The term $e^{-\zeta\omega_n t}$ is real and decreases exponentially with time. (In the previous section, ζ was equal to zero, and this factor remained unity.) The second factor in Eq. (6.65) looks very much like the response from Eq. (6.39) and can be reduced to a harmonic oscillation at the damped natural frequency ω_d with the general form

$$x(t) = e^{-\varsigma\omega_n t}(A_x \cos\omega_d t + B_x \sin\omega_d t). \tag{6.66}$$

The end result is a decaying exponential multiplying an oscillatory term. Figure 6.4 shows the general appearance of a decaying oscillatory response.

It appears that interpreting the meaning of eigenvalues is not that difficult. They can be complex numbers, meaning they have a real part and an imaginary part. It appears that if one eigenvalue is a complex number, then another eigenvalue is the complex conjugate of the first. It also appears that the real parts of both eigenvalues, $-\varsigma\omega_n$ in this example, indicates decay of the response. (What if the real part were positive?) The imaginary part of the eigenvalue, ω_d in this example, indicates that the system can oscillate.

From Eq. (6.62) it is apparent that when damping is present, the oscillation frequency is different than the undamped natural frequency. From (6.62) we can derive

$$\left(\frac{\omega_d}{\omega_n}\right)^2 + \zeta^2 = 1. \tag{6.67}$$

This relationship is plotted in Figure 6.5. From this figure it is apparent that the damping ratio must be quite high to reduce the oscillation frequency significantly.

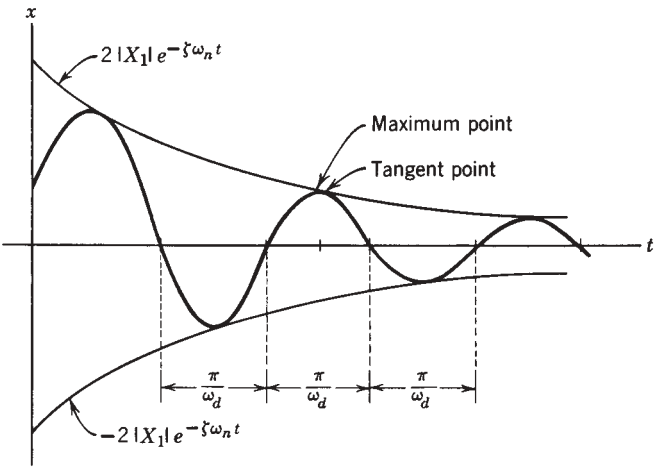


FIGURE 6.4. Time response for underdamped system.

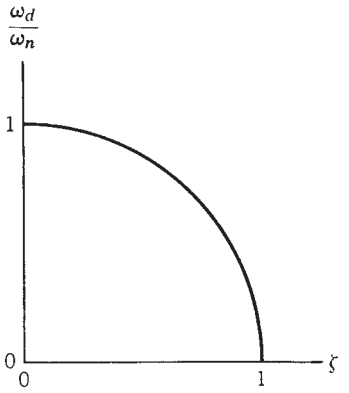


FIGURE 6.5. Variation of damped natural frequency with damping ratio.

Since so much understanding of the general nature of the free response is contained in the roots of the characteristic equation (i.e., the eigenvalues), it is useful to show how the damping ratio and the damped and undamped natural frequencies are correlated with the locations in the "complex" plane of the eigenvalues s_1 and s_2. This is shown in Figure 6.6, where the real part of the eigenvalue is shown along the horizonatal $\mathscr{R}$-axis and the imaginary part is along the vertical $\mathscr{I}$-axis. For the underdamped case with $\varsigma < 1$ the eigenvalues are complex conjugates and this is shown as s_1 and s_2 in the figure.

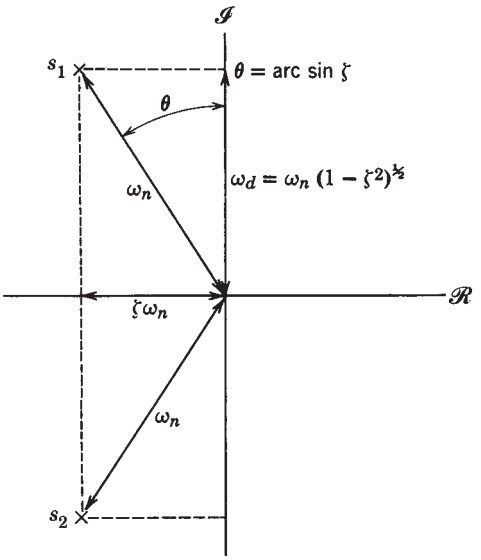


FIGURE 6.6. Correlation of natural frequencies and damping ratio with location of eigenvalues in the complex plane for underdamped case.

Stable systems have all their eigenvalues in the *left* half of the complex plane; that is, for stable systems the real components of all the $e^{s_i t}$ factors in the free response represent *decreasing* exponentials. Starting from any initial conditions, all stable systems have only components that decay toward the zero state. *Unstable* systems have at least one eigenvalue in the *right* half plane so that some components *increase* exponentially with time. The intermediate case of the undamped oscillator corresponds to purely imaginary values of s on the $j\omega$-axis.

For the overdamped oscillator, $\zeta > 1$, the character of the eigenvalues of Eq. (6.30) changes, and it is convenient to rewrite Eqs. (6.61) as

$$s_1 = -\zeta\omega_n + \omega_n(\zeta^2 - 1)^{1/2}, \quad s_2 = -\zeta\omega_n - \omega_n(\zeta^2 - 1)^{1/2}. \tag{6.68}$$

These roots are plotted in Figure 6.7. Note that for $1 < \zeta < \infty$ the roots s_1 and s_2 are negative and real. In such a case, terms of the form e^{st} are conveniently written in terms of a time constant τ, as was done for the first-order case. Thus, in Figure 6.7, s-values far out on the negative real axis correspond to quickly decreasing real exponentials.

Also, it is clear that the real part of s_1 and s_2 in Figure 6.6, $\zeta\omega_n$, corresponds to a time constant of $1/\zeta\omega_n$ (in fact, time constants are always $\tau_i = 1/\mathrm{Re}(s_i)$). Eigenvalues on the positive real axis also possess time constants but represent exponential solutions that increase in time. A value of $s = 0$ implies a time-response component of $e^{0t} = 1$, or a constant component: an infinite time constant.

Before leaving this section it is useful to summarize our observations about eigenvalues.

1. There is always the same number of eigenvalues as there are system state variables when using a bond graph to model a system.
2. Eigenvalues can be real numbers, imaginary numbers, and/or complex numbers with real and imaginary parts.
3. Complex eigenvalues will always appear as complex conjugate pairs.

Eigenvalues always appear in the system response as arguments of exponentials. This allows the interpretation that real parts of eigenvalues always indicate

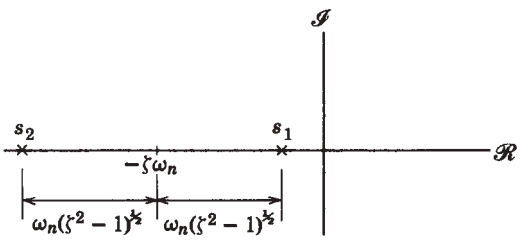


FIGURE 6.7. Eigenvalues for overdamped case.

exponential growth or decay and imaginary parts of eigenvalues are frequencies at which the system response oscillates.

System stability is indicated by the signs of the real eigenvalues and/or the real parts of the complex eigenvalues. A system with all its eigenvalues in the left half of the complex plane is stable. A system with one or more of its eigenvalues in the right half of the complex plane will have at least part of its free response increasing exponentially with time and is thus unstable.

A study of the coefficients of the characteristic equation can often determine the nature of the eigenvalues without actually solving for them numerically. See the discussion of the Routh stability criterion in Reference [1], for example.

6.3.5 The General Case

Although only first- and second-order systems have been studied above, the characteristics of the free response of all linear systems are not much more complicated. The procedure for finding the free response is summarized below:

1. Neglect all forcing terms. In the general equations of (6.1a) or (6.1b), let $u_1, u_2, \ldots, u_r$ all vanish.
2. Assume each state variable x_i will possess solutions of the form $X_i e^{st}$. After canceling out the e^{st} terms, Eq. (6.1a) becomes a purely algebraic equation:

$$\begin{aligned} (s - a_{11})X_1 - a_{12}X_2 - \cdots - a_{1n}X_n &= 0, \\ -a_{21}X_1 + (s - a_{22})X_2 - \cdots - a_{2n}X_n &= 0, \\ &\vdots \\ -a_{n1}X_1 - a_{n2}X_2 - \cdots + (s - a_{nn})X_n &= 0. \end{aligned} \tag{6.69}$$

   Note that this equation is the general version of Eq. (6.36b) and, in condensed form, can be written $[sI - A][X] = [0]$. This set of equations always possesses a trivial solution, $X_1 = X_2 = \cdots = X_n = 0$, which is the *only* solution unless the determinant of $[sI - A]$ is zero:

$$\text{Det} \begin{vmatrix} s - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & s - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & s - a_{nn} \end{vmatrix} = 0. \tag{6.70}$$

   This is the general version of the example determinant shown in Eq. (6.28).

   When the determinant is expanded out into an nth-order polynomial in s, the characteristic equation results.
3. Solve the characteristic equation for the n eigenvalues $s_1, s_2, \ldots, s_n$. The s-values will either be real or, if complex, will occur in complex conjugate pairs. The location of the eigenvalues in the complex plane indicates the

type of component of the free response that is associated with the eigenvalues. Figure 6.8 indicates how root location may be correlated with time response. Stable systems possess eigenvalues with negative real parts.

4. For each value of s corresponding to an eigenvalue, Eq. (6.69) yields a partial solution. When the eigenvalues are distinct, that is, when no two of the n eigenvalues are exactly equal, then $n-1$ of the n equations in Eq. (6.69) will be found to be independent when an eigenvalue is substituted for s.* One could then solve for $n-1$ of the X's in terms of the last X, for example. The total solution for the free response is

$$\begin{aligned} x_1(t) &= X_{11}e^{s_1 t} + X_{12}e^{s_2 t} + \cdots + X_{1n}e^{s_n t}, \\ x_2(t) &= X_{21}e^{s_1 t} + X_{22}e^{s_2 t} + \cdots + X_{2n}e^{s_n t}, \\ &\vdots \\ x_n(t) &= X_{n1}e^{s_1 t} + X_{n2}e^{s_2 t} + \cdots + X_{nn}e^{s_n t}, \end{aligned} \tag{6.71}$$

where $X_{11}, X_{21}, X_{31}, \ldots, X_{n1}$ appear in Eq. (6.69) when $s = s_1$; $X_{12}, X_{22}, X_{32}, \ldots, X_{n2}$ appear in Eq. (6.69) when $s = s_2$; and so on. The n^2 values of X_{ji} in Eq. (6.71) cannot be completely determined by writing Eq. (6.69) for each eigenvalue s_i, since in this manner only $n(n-1)$ independent equations will be obtained. The remaining n conditions are represented by the n arbitrary initial conditions

$$\begin{aligned} x_1(t_0) = x_{10} &= X_{11}e^{s_1 t_0} + X_{12}e^{s_2 t_0} + \cdots + X_{1n}e^{s_n t_0}, \\ x_2(t_0) = x_{20} &= X_{21}e^{s_1 t_0} + X_{22}e^{s_2 t_0} + \cdots + X_{2n}e^{s_n t_0}, \\ &\vdots \\ x_n(t_0) = x_{n0} &= X_{n1}e^{s_1 t_0} + X_{n2}e^{s_2 t_0} + \cdots + X_{nn}e^{s_n t_0}, \end{aligned} \tag{6.72}$$

where $x_{10}, x_{20}, \ldots, x_{n0}$ are the given initial conditions.

Only for very low-order systems is it possible to carry out the operations outlined in Eqs. (6.71) and (6.72) that lead to analytical expressions for the free response. Yet any number of computer simulation programs can numerically solve the unforced equations, given the matrix and the initial conditions. Thus, it is easy to obtain plots of the free response of all the state and output variables if desired.

Furthermore, there are many programs that will calculate the eigenvalues, given only the $[A]$ matrix. Thus, one can understand the qualitative nature of all the components of a system's free response using the ideas behind Figure 6.8 without computing any specific free responses. For example, it is easy to see if

*The special case when roots of the characteristic equation are repeated is treated in all books on differential equations. It is rarely important in practice, so we will not discuss it here.

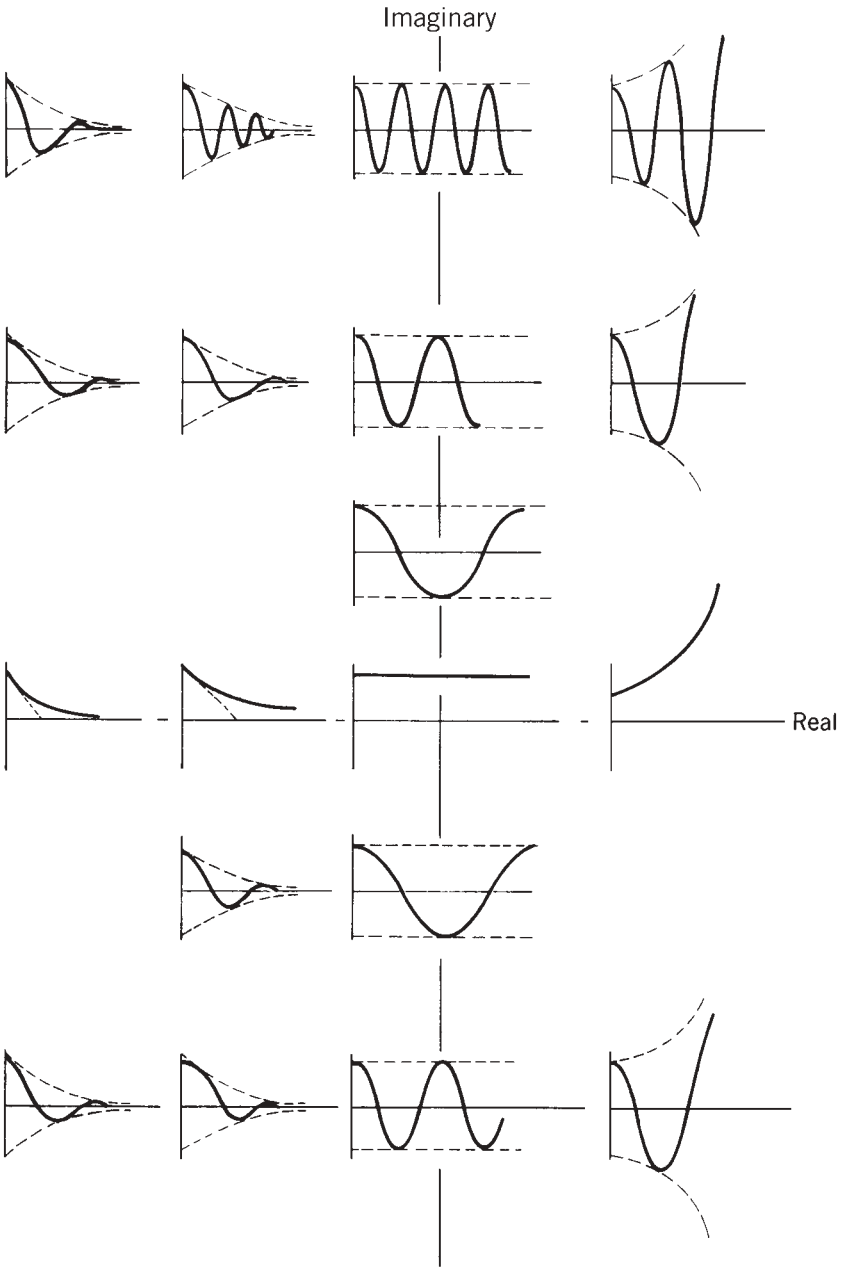


FIGURE 6.8. Correlation of time response with complex eigenvalues.

the system is unstable by looking for one or more eigenvalues with positive real parts. One can also look for lightly damped oscillations and find their frequencies by using representations like that of Figure 6.6, and one can find the fastest and slowest response components just by considering the locations of the various eigenvalues in the s-plane. Eigenvalues farthest from the origin of the s-plane in any direction represent the fastest response components, and those closest to the origin represent the slowest response components. Often one can make design decisions about a dynamic system simply by considering the eigenvalue locations.

6.4 TRANSFER FUNCTIONS

Section 6.3 developed the concept of eigenvalues and showed how the nature of the system's initial condition response is dictated by the eigenvalues. It is definitely worth the reader's time to become comfortable with deriving system equations directly from bond graphs, casting the state equations into the s-domain, and deriving the characteristic equation from which eigenvalues can be evaluated.

A second very important concept in linear system analysis is the transfer function. It is the ratio of an output to an input expressed in terms of the s variable rather than time. As mentioned earlier, the Laplace transform is often used to convert linear, constant coefficient differential equations in the time domain into equivalent algebraic equations in the s-domain. It is perhaps more physically intuitive to convert from the time domain to the s-domain using an exponential representation and we continue to do this for the development of transfer functions.

A transfer function is the ratio of output to input in the s-domain. Consider the first-order example from the previous section 6.3.1, where the equation of motion from Eq. (6.1) is

$$\dot{q} = -(k/b)q + (1/b)F_{\text{in}}. \tag{6.73}$$

To obtain the eigenvalue we previously set the forcing to zero. To pursue the transfer function relating the response $Q(s)$ to the input $F(s)$ we assume that the response and input have the form,

$$F_{\text{in}}(t) = F(s)e^{st},\ q(t) = Q(s)e^{st},\ \dot{q}(t) = sQ(s)e^{st}. \tag{6.74}$$

Substituting into Eq. (6.73), the transfer function becomes

$$\frac{Q(s)}{\mathrm{F}(s)} = \frac{1/b}{s + k/b} = \frac{1/b}{s + 1/\tau} = \frac{\tau/b}{\tau s + 1}. \tag{6.75}$$

In automatic control it is typical to represent transfer functions using blocks, as shown in Figure 6.9. The interpretation is that the input is multiplied by the contents of the block to produce the output.

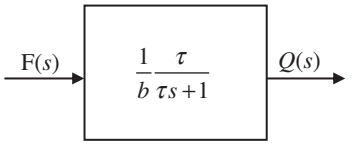


FIGURE 6.9. Block diagram of the transfer function from Eq. (6.75).

For the second-order example from Section 6.3.2 the state equations repeated here are

$$\begin{aligned} \dot{x} &= 0x - (1/m)p + 1V + 0F, \\ \dot{p} &= kx - (b/m)p + bV + 1F. \end{aligned} \tag{6.76}$$

or in matrix form

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 0 & -1/m \\ k & -b/m \end{bmatrix}\begin{bmatrix} x \\ p \end{bmatrix} + \begin{bmatrix} 1 \\ b \end{bmatrix} V + \begin{bmatrix} 0 \\ 1 \end{bmatrix} F. \tag{6.77}$$

For the eigenvalue analysis the forcing terms were set to zero. To obtain transfer functions relating desired outputs to the inputs, we assume

$$\begin{aligned} V(t) &= V(s)e^{st} \quad F(t) = F(s)e^{st}, \\ x(t) &= X(s)e^{st} \quad p(t) = P(s)e^{st}, \end{aligned} \tag{6.78}$$

and substitute into Eq. (6.77) with the result:

$$\begin{bmatrix} s & 1/m \\ -k & s + b/m \end{bmatrix}\begin{bmatrix} X(s) \\ P(s) \end{bmatrix} = \begin{bmatrix} 1 \\ b \end{bmatrix} V(s) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} F(s). \tag{6.79}$$

The astute reader will recognize that the matrix on the left side of Eq. (6.79) is identical to the matrix $[sI - A]$ from the eigenvalue analysis using Eq. (6.27).

A general expression for transfer functions obtained from starting points such as Eq. (6.79) will be shown below. If a specific input/output transfer function is desired, then Cramer's rule [4] can be used to obtain the result. For example, if you want the transfer function relating the output $X(s)$ to the input $V(s)$, first let the force input $F(s)$ be zero. Then Eq. (6.79) is two algebraic equations for two unknowns X and P with V as a known forcing quantity.

Cramer's rule states that the solution for the unknowns will always be the ratio of two determinants. The denominator will always be the determinant of the matrix, $\mathrm{Det}\,[sI - A]$. The numerator determinant will be the determinant of this matrix but with one column replaced by the forcing column of interest. In this first example, the forcing column is just $\begin{bmatrix} 1V \\ bV \end{bmatrix}$. This column is placed in the $[sI - A]$ matrix in the first column to find the first variable (X), in the second column to find the second variable (P), and so on for higher-order systems. Since X is the *first* entry in the response vector, the numerator comes from the

determinant of the left-side matrix having first substituted the forcing vector for the V input for the *first* column. The denominator for the transfer function is the determinant of the left-side matrix with no changes made. Since we desire the ratio X/V we can divide out the V in the forcing column, with the result

$$\frac{X}{V}(s) = \frac{\begin{vmatrix} 1 & 1/m \\ b & s + b/m \end{vmatrix}}{\begin{vmatrix} s & 1/m \\ -k & s + b/m \end{vmatrix}} = \frac{s}{s^2 + (b/m)s + (k/m)}. \tag{6.80}$$

If the response of X to the input F is desired, we assume that V is zero and use the forcing column appropriate for F. Cramer's rule then yields

$$\frac{X}{F}(s) = \frac{\begin{vmatrix} 0 & 1/m \\ 1 & s + b/m \end{vmatrix}}{\begin{vmatrix} s & 1/m \\ -k & s + b/m \end{vmatrix}} = \frac{-1/m}{s^2 + (b/m)s + (k/m)}. \tag{6.81}$$

Notice that the denominator is unchanged from Eq. (6.80).

For the response of $P(s)$ to the input $V(s)$ we put the forcing column for V in the *second* column of $[sI - A]$, because P is the *second*-state variable, to get

$$\frac{P}{V}(s) = \frac{\begin{vmatrix} s & 1 \\ -k & b \end{vmatrix}}{\begin{vmatrix} s & 1/m \\ -k & s + b/m \end{vmatrix}} = \frac{bs + k}{s^2 + (b/m)s + (k/m)}. \tag{6.82}$$

Notice again that the denominator of the transfer function in Eq. (6.82) is the same for all the transfer functions we have evaluated. The denominator in *all* the transfer functions is exactly the polynomial in s that we found in the characteristic equation, Eq. (6.29), for this second-order system. This is not a coincidence and the same idea proves to be true for all systems. Figure 6.10 shows these transfer functions in block diagram form.

6.4.1 The General Case for Transfer Functions

The general starting point for linear analysis is the matrix format,

$$\begin{aligned} [\dot{x}] &= [A][x] + [B][u], \\ [y] &= [C][x] + [D][u], \end{aligned} \tag{6.83}$$

where $[x]$ is the vector of state variables (p's and q's from a bond graph), $[u]$ is the vector of inputs (efforts and flows from source elements from a bond graph),

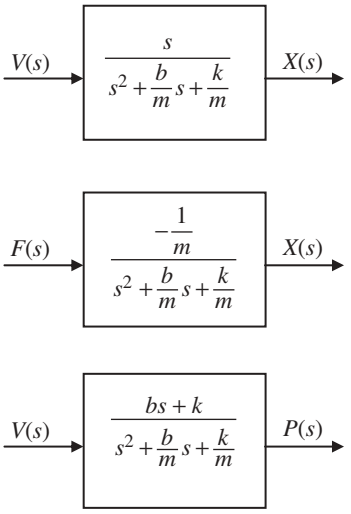


FIGURE 6.10. Block diagram representation of the transfer functions for the second-order example.

and $[y]$ is a vector of outputs (efforts and flows of interest on various bonds). To obtain transfer functions we let

$$\begin{aligned}[x(t)] &= [X(s)]e^{st},\\ [U(t)] &= [U(s)]e^{st}.\end{aligned} \tag{6.84}$$

It is convenient to consider one input at a time and one output at a time such that substituting (6.84) into (6.83) yields

$$[sI - A]\,[X(s)] = [b]U(s). \tag{6.85}$$

It is typical to use a lowercase $[b]$ to indicate a column vector when considering only one input at a time. In fact, (6.83) with (6.84) could be written in terms of the separate inputs as

$$[sI - A]\,[X(s)] = [b_1]\,U_1(s) + [b_2]\,U_2(s) + [b_3]\,U_3(s) + \cdots. \tag{6.86}$$

Also, when considering only one output at a time it is typical to write the output equation as

$$y = [c]^t[x], \tag{6.87}$$

where $[c]^t$ indicates a vector on its side, i.e., a row vector. For simplicity the single output y is assumed to not have dependence on any inputs. In other words, the components of the $[D]$ matrix in Eq. (6.83) pertaining to this output happen to be zero. Some other output certainly could depend explicitly on inputs, which would somewhat complicate the presentation here.

The s-domain representation of the output equation becomes

$$Y(s) = [c]^t[X](s). \tag{6.88}$$

We now can symbolically exhibit the transfer function between an output and input by starting with (6.85) and writing

$$[X(s)] = [sI - A]^{-1}[b]U(s), \tag{6.89}$$

and then use (6.88) with the result

$$Y(s) = [c]^t\,[X(s)] = [c]^t[sI - A]^{-1}[b]U(s). \tag{6.90}$$

Finally, the general expression for a transfer function comes from Eq. (6.90) as

$$\frac{Y(s)}{U(s)} = [c]^t[sI - A]^{-1}[b]. \tag{6.91}$$

Note that for each output and input under consideration there will be specific $[c]^t$ and $[b]$ vectors.

There is one further symbolic manipulation that can be made to Eq. (6.91) using the formula for the inverse of a matrix to arrive at the expression

$$\frac{Y(s)}{U(s)} = \frac{[c]^t\text{Adj}[sI - A][b]}{\text{Det}[sI - A]} \tag{6.92}$$

where Adj[$\cdot$] is the adjoint and Det [$\cdot$] is the determinant of the argument matrix.

Although we would rarely attempt carry out the indicated operations in Eq. (6.92) to derive a transfer function by hand, this formula shows that the Det$[sI - A]$ is the denominator of all input/output transfer functions for a given dynamic system and this denominator is the characteristic polynomial that was used to determine the eigenvalues in Eq. (6.70).

For a given dynamic system characterized by the linear Eqs. (6.83), the numerator of any input/output transfer function is different for each input and each output. For low-order systems, Cramer's rule can be used to determine the transfer function between ith-state variable X_i and the jth input U_j. It functionally appears as shown in Eq. (6.93):

$$\frac{X_i(s)}{U_j(s)} = \frac{\left|\begin{array}{l}\text{Determinant of } [sI - A] \\ \text{having first substituted the } \left[b_j\right] \text{ vector} \\ \text{for the } i\text{th column}\end{array}\right|}{|sI - A|}. \tag{6.93}$$

Then the transfer function for any desired output can be constructed as a combination of the state variable transfer functions according to the output equation.

This has been a brief introduction to the topic of transfer functions. We will be using transfer functions extensively when the topic of automatic control is discussed.

It is important to know that there are many linear analysis computer programs, such as Matlab [5], that can determine transfer functions even for very large systems. As an example, in Matlab, after creating a system model and deriving a linear set of governing equations in the matrix form of Eq. (6.83) we would first create a system using the Matlab state-space function with a command such as ''System=ss(A,b,c_t,0)''. The focus on a single input and output defines the b and c vectors as in Eqs. (6.85) and (6.87). This is followed by the command that turns the system into a transfer function ''TF_system=tf(System)''. The result, ''TF_system'', can be viewed symbolically as a ratio of polynomials in the s-domain as well as being used for other analysis applications. With programs that can handle linear systems, the translation of state equations and output equations to transfer functions is routine.

6.5 FREQUENCY RESPONSE

We now turn to the problem of finding solutions to Eqs. (6.1a) or (6.1b) that are compatible with given time functions for $u_1(t), u_2(t), \ldots, u_r(t)$. Again, the principle of superposition will prove useful for the linear system, since it will allow us to find solutions for each input separately and then permit the simple addition of the solutions to find a solution for all the inputs acting simultaneously.

In principle, any time function may be studied as an input, but in practice, sinusoidal forcing is of overwhelming importance. This is partly because periodic forcing can be decomposed into sinusoidal components using the Fourier series with the effects of the component forcing functions superposed using the principle of superposition, and partly because the testing of real devices is conveniently accomplished in many cases by using sinusoidal inputs at frequencies varied over a range of interest.

Most industries that do any kind of vibration testing will have shaker facilities that can generate sinusoidal inputs and monitor the response of system components. Automotive companies also have shaker facilities for testing suspensions, and even entire vehicles can be shaken in larger facilities. Computer and electronic components are tested using shaker tables that can sweep through many frequencies and determine if components break, and their electrical input–output properties are similarly tested using sinusoidal voltage inputs.

Frequency response is the response of various system outputs to fixed amplitude, fixed frequency inputs. One might guess that if the input to a system is a fixed-frequency sinusoid, then the outputs from that system will also be fixed-amplitude sinusoids at the same frequency as the input. It turns out that this is true for linear systems but not necessarily for nonlinear systems. For more realistic nonlinear systems the response to a fixed-frequency input can be a much distorted looking periodic wave and can even have a different fundamental frequency than the input. Nevertheless, the fixed-frequency sinusoidal input is used pervasively in engineering applications and yields useful approximate results even for most nonlinear systems.

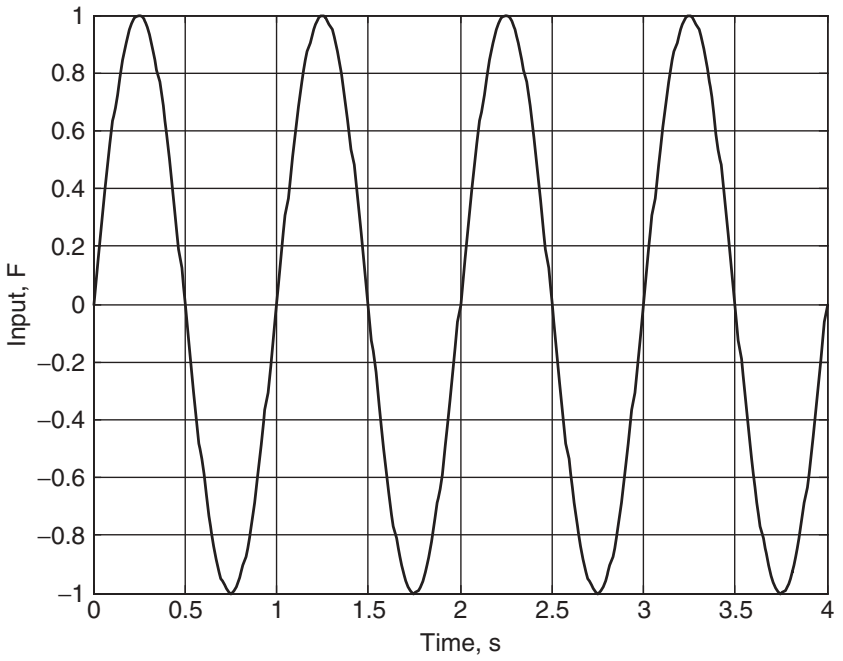


FIGURE 6.11. Fixed-frequency input force of 1 N amplitude at a frequency of 1 Hz.

One way to present a frequency response is to show the time histories of the input and response. For the second-order example system that generated Eq. (6.16) some parameters were chosen and the response was generated for a sinusoidal input at a frequency of 1 Hz. The fixed-amplitude input and the corresponding displacement response are shown in Figures 6.11 and 6.12. Notice that after a brief transient the output becomes a fixed-amplitude sinusoid. The term *frequency response* refers to the *steady-state response* of the system after transients have died away. (You can see why frequency-response analysis doesn't work for unstable systems.) Note that the response is not synchronized in time with the input. We will describe this time shift as the response having a *phase relationship* to the input. We could change the frequency and run the simulation again and the response would have a different amplitude and phase relationship. We could do this procedure for many frequencies and generate many plots of the time history of the response.

What is of more interest is the response amplitude and phase as they depend on the forcing frequency. For the second-order example system from Figure 6.3 the state equations were derived as Eqs. (6.16). This state representation is the preferred starting point for most analysis and simulation. However, for demonstration purposes, Figure 6.13 shows a simplified version of the second-order example of Figure 6.3. In terms of the displacement of the mass y due to the

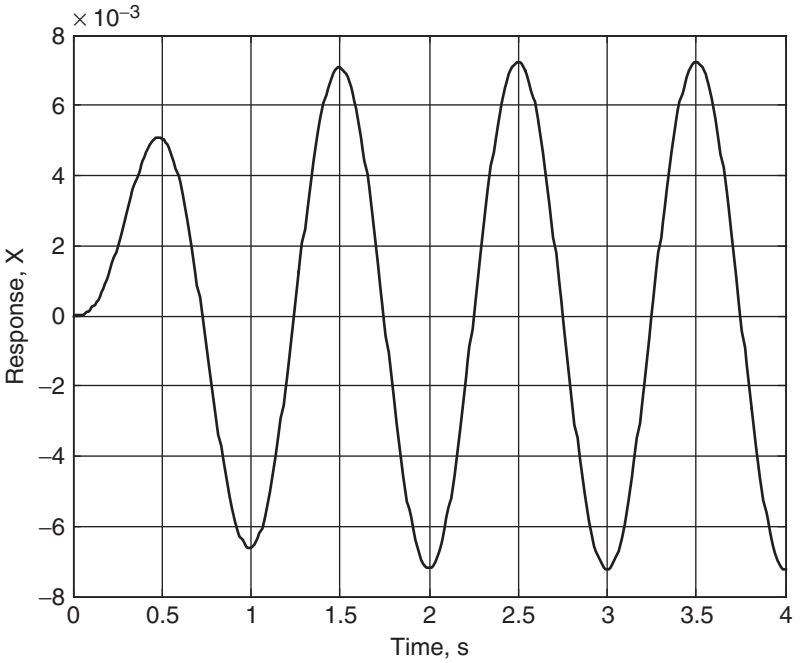


FIGURE 6.12. Response to a fixed-frequency force input.

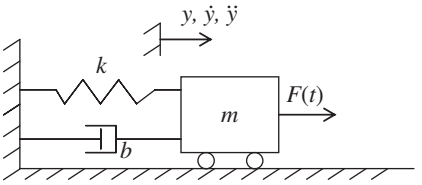


FIGURE 6.13. Simplified version of the system from Figure 6.3.

force input $F(t)$, the single equation of motion can be written as

$$\ddot{y} + (b/m)\dot{y} + (k/m)y = (1/m)F(t). \tag{6.94}$$

(Note that we now define the mass position to be the negative of the spring deflection as shown in the original bond graph, $y = -x$. Thus, the transfer function $Y(s)/F(s)$ from Eq. (6.94) is the same as $X(s)/F(s)$ in Eq. (6.81) except for the missing minus sign.)

If the input is

$$F(t) = F_0 \cos(\omega t), \tag{6.95}$$

then the steady-state response of mass displacement will be

$$y(t) = Y_0 \cos(\omega t + \phi). \tag{6.96}$$

We are not so much interested in plotting the time histories. We really would like to determine how Y_0 and ϕ depend on the forcing frequency ω. When the response amplitude divided by the input amplitude is plotted versus the forcing frequency, and the response phase angle is plotted versus the forcing frequency, the results are called the frequency-response plots of the system. Such a presentation is called a Bode plot when the amplitude vs. frequency plot is shown on log–log scales and the phase angle vs. frequency plot is shown on a linear-log scale. The goal of frequency-response analysis is to determine the amplitude and phase dependence on forcing frequency.

One way to determine a frequency response for the single second-order Eq. (6.94) is to use Eq. (6.95) for the input and Eq. (6.96) for the response and substitute into the governing equation. This will generate a lot of algebra to solve for Y_0 and ϕ as functions of ω. Another approach is to use the *real part of* operator introduced in Section 6.3.3,

$$F(t) = F_0 \cos(\omega t) = \mathrm{Re}(F_0 e^{j\omega t}) \tag{6.97}$$

and

$$y(t) = Y_0 \cos(\omega t + \phi) = \mathrm{Re}(Y_0 e^{j(\omega t+\phi)}) = \mathrm{Re}(Y_0 e^{j\phi} e^{j\omega t}) = \mathrm{Re}(H_y e^{j\omega t}), \tag{6.98}$$

where H_y is a complex number that can be expressed as

$$H_y = Y_0 e^{j\phi}. \tag{6.99}$$

The magnitude of this complex number is the response amplitude we desire and the angle of this complex number is the phase angle we desire. Thus,

$$\left|H_y\right| = Y_0 \quad \text{and} \quad \angle H_y = \phi. \tag{6.100}$$

In other words, if we can obtain the complex number H_y, then we can recover the response amplitude by taking the magnitude of H_y and the response phase angle by determining the angle of the complex number H_y. The complex number H_y is called the *complex frequency response function*. (Often instead of H_y, we will consider the ratio H_y/F_0, since for linear systems the response is always proportional to the input amplitude.)

It turns out that it is simple to compute the complex frequency-response function using the real operator and far more difficult to perform the algebra if one starts with the real-time functions from (6.95) and (6.96).

To obtain the complex frequency response function $H_y = Y_0 e^{j\phi}$ for the example system in Figure 6.13 we would assume that the input $F(t)$ is

$$F(t) = \mathrm{Re}(F_0 e^{j\omega t}) \tag{6.101}$$

and the response is

$$y(t) = \mathrm{Re}(H_y e^{j\omega t}), \tag{6.102}$$

then substitute these into the governing differential equation Eq. (6.94) and carry out the indicated operations to obtain H_y. It turns out that the *real part of* operator $\mathrm{Re}(\bullet)$ can be understood but suppressed during the operations necessary to find the frequency response. Thus, the input and output can be expressed simply as

$$F(t) = F_0 e^{j\omega t}, \quad y(t) = H_y e^{j\omega t}. \tag{6.103}$$

Comparing Eqs. (6.103) with the input/output assumption for transfer functions from Eqs. (6.78), we notice that these forms are identical if $s = j\omega$.

We now state without proof that the complex frequency response function is obtained by simply substituting $s = j\omega$ in the transfer function between any output and any input. For the example system from Figure 6.13 and governed by the equation of motion (6.94), to derive the transfer function between the output $Y(s)$ and input $F(s)$, let

$$\begin{aligned} F(t) &= F(s)e^{st}, \\ y(t) &= Y(s)e^{st}, \end{aligned} \tag{6.104}$$

with the result

$$\frac{Y}{F}(s) = \frac{1}{ms^2 + bs + k}. \tag{6.105}$$

If the input is a fixed-frequency cosine wave,

$$F = F_0 \cos \omega t, \tag{6.106}$$

then we expect the steady-state response to be

$$y(t) = Y \cos(\omega t + \phi). \tag{6.107}$$

The complex frequency response would be

$$\frac{Y(j\omega)}{\mathrm{F}(j\omega)} = \frac{1}{m(j\omega)^2 + bj\omega + k} = \frac{1}{k - m\omega^2 + jb\omega}. \tag{6.108}$$

The response amplitude divided by the input amplitude comes from the first of (6.100) as

$$\left|\frac{Y}{\mathrm{F}}\right| = \frac{1}{\left[(k - m\omega^2)^2 + (b\omega)^2\right]^{1/2}}, \tag{6.109}$$

where the reader should recall that the magnitude of a complex number comes from the square root of the sum of the squares of the real and imaginary parts.

The response phase angle comes from the second of (6.100) as

$$\angle\frac{Y}{\mathrm{F}} = 0 - \tan^{-1}\frac{b\omega}{k - m\omega^2}, \tag{6.110}$$

where we have made use of the fact that the associated angle of the ratio of complex numbers is the angle of the numerator complex number minus the angle of the denominator complex number. We have also used the fact that the angle of a complex number is the arctangent of the imaginary part divided by the real part. For this example the numerator has no imaginary part and the arctangent of zero is zero. We could now plot or sketch the magnitude and phase relationships versus the forcing frequency.

For low-order linear systems with only a few governing equations, carrying out the analytical steps shown here is a reasonable exercise. For larger systems we would resort to a computational solution. For example, the commercial program Matlab [5] can produce the frequency response directly from the matrix formulation for linear systems in the form of Eqs. (6.83). The Matlab program was used to produce the displacement frequency response shown in Figure 6.14.

6.5.1 Example Transfer Functions and Frequency Responses

In this section the analytical tools developed in the previous sections are demonstrated for several physical systems.

Simple Vehicle Model. Consider first the system shown in Figure 6.15. This model is often used to represent the vertical dynamics of an automotive suspension in its simplest form. The velocity input $v_{\text{in}}(t)$ at the base of the system represents the roadway unevenness and the control force $F_c(t)$ between the base and the mass represents an actuator that can be controlled to produce a desired response of the suspension. Also shown in Figure 6.15 is the bond graph for the system with state variables identified.

Using procedures from Chapter 5, the state equations come directly from the bond graph as

$$\begin{aligned} \dot{p}_m &= kq_k + bv_{\text{in}} - (b/m)p_m + F_c, \\ \dot{q}_k &= v_{\text{in}} - (1/m)p_m, \end{aligned} \tag{6.111}$$

or, in matrix form,

$$\frac{d}{dt}\begin{bmatrix} p_m \\ q_k \end{bmatrix} = \begin{bmatrix} -b/m & k \\ -1/m & 0 \end{bmatrix}\begin{bmatrix} p_m \\ q_k \end{bmatrix} + \begin{bmatrix} b \\ 1 \end{bmatrix} v_{\text{in}} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} F_c. \tag{6.112}$$

If we desire eigenvalues, transfer functions, and/or frequency response, then we start with

$$p_m(t) = P_m(s)e^{st},\ q_k(t) = Q_k(s)e^{st},\ v_{\text{in}}(t) = V_{\text{in}}(s)e^{st},\ F_c(t) = F_c(s)e^{st}. \tag{6.113}$$

This substitution into (6.112) yields

$$\begin{bmatrix} s + b/m & -k \\ 1/m & s \end{bmatrix}\begin{bmatrix} P_m \\ Q_k \end{bmatrix} = \begin{bmatrix} b \\ 1 \end{bmatrix} V_{\text{in}} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} F_c. \tag{6.114}$$

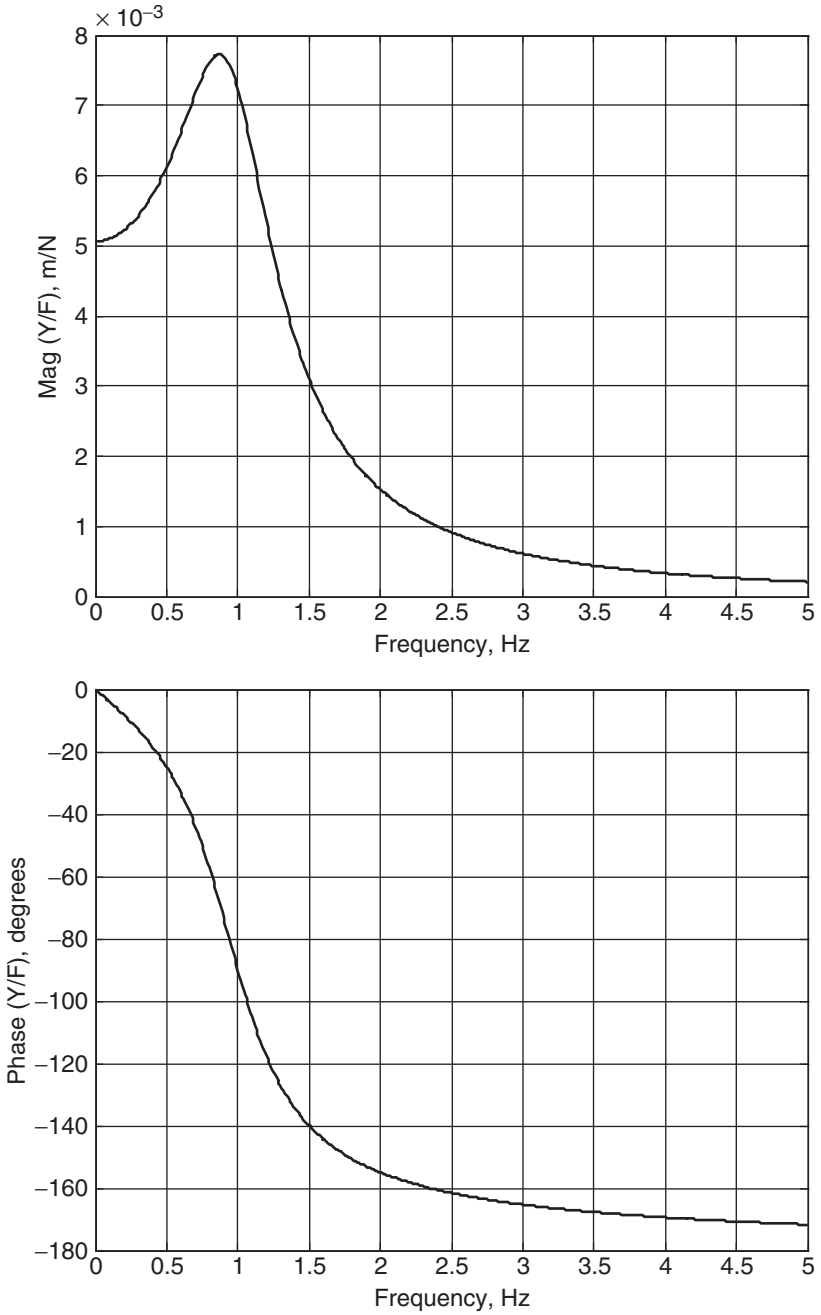


FIGURE 6.14. The frequency-response plots for the system of Figure 6.13.

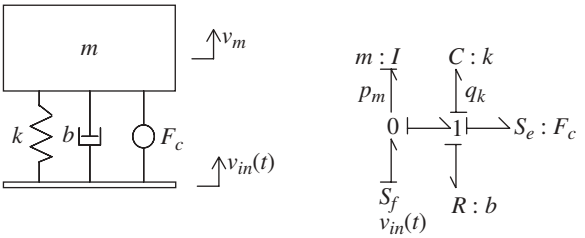


FIGURE 6.15. Simple vehicle suspension and bond graph.

The reader should recognize Eq. (6.114) as a particular version of the general form

$$[sI - A][X] = [b_{\text{in}}]\, V_{\text{in}} + [b_F]\, F_c. \tag{6.115}$$

Let's look at the transfer functions relating the mass velocity v_m to the base velocity v_{in} and the control force F_c. Since the mass velocity is the momentum state variable p_m divided by the mass parameter m, it is most convenient to derive the transfer function for P_m in Eq. (6.114) and simply divide the result by m. Using Cramer's rule,

$$\frac{P_m}{V_{\text{in}}}(s) = \frac{\begin{vmatrix} b & -k \\ 1 & s \end{vmatrix}}{s^2 + (b/m)s + k/m}, \quad \frac{P_m}{F_c}(s) = \frac{\begin{vmatrix} 1 & -k \\ 0 & s \end{vmatrix}}{s^2 + (b/m)s + k/m} \tag{6.116}$$

or

$$\frac{V_m}{V_{\text{in}}}(s) = \frac{(b/m)s + k/m}{s^2 + (b/m)s + k/m}, \quad \frac{V_m}{F_c}(s) = \frac{(1/m)s}{s^2 + (b/m)s + k/m}. \tag{6.117}$$

If we want to compute the eigenvalues, the characteristic equation is the denominator of the transfer functions,

$$s^2 + (b/m)s + k/m = 0. \tag{6.118}$$

If the frequency response is desired, then let $s = j\omega$ in the transfer functions and generate the complex frequency-response function. For the mass velocity response to the base velocity input the result using the first of Eqs. (6.117) is

$$\frac{V_m}{V_{\text{in}}}(j\omega) = \frac{(b/m)(j\omega) + k/m}{(j\omega)^2 + (b/m)(j\omega) + k/m} = \frac{k/m + j(b/m)\omega}{k/m - \omega^2 + j(b/m)\omega}. \tag{6.119}$$

The magnitude of the frequency response is the magnitude of the complex number from Eq. (6.119),

$$\left|\frac{V_m}{V_{\text{in}}}\right| = \frac{\left[(k/m)^2 + ((b/m)\omega)^2\right]^{1/2}}{\left[\left(k/m - \omega^2\right)^2 + ((b/m)\omega)^2\right]^{1/2}}, \tag{6.120}$$

and the phase angle between the mass velocity and the base input velocity is the angle of the complex number from Eq. (6.119),

$$\angle \frac{V_m}{V_{\text{in}}}(j\omega) = \tan^{-1} \frac{(b/m)\omega}{k/m} - \tan^{-1} \frac{(b/m)\omega}{k/m - \omega^2}, \tag{6.121}$$

which is the angle of the numerator complex number minus the angle of the denominator complex number.

Using the definitions of natural frequency and damping ratio from Section 6.3.4, the magnitude and phase can also be expressed as

$$\left| \frac{V_m}{V_{\text{in}}} \right| = \frac{\left[1 + (2\varsigma(\omega/\omega_n))^2\right]^{1/2}}{\left[\left(1 - (\omega/\omega_n)^2\right)^2 + (2\varsigma(\omega/\omega_n))^2\right]^{1/2}} \tag{6.122}$$

and

$$\angle \frac{V_m}{V_{\text{in}}}(j\omega) = \tan^{-1} 2\varsigma(\omega/\omega_n) - \tan^{-1} \frac{2\varsigma(\omega/\omega_n)}{1 - (\omega/\omega_n)^2}. \tag{6.123}$$

Figure 6.16 shows the magnitude and phase for a few values of the damping ratio.

Side Branch Accumulator. As another example consider the system shown in Figure 6.17. It is fluid-filled tube with a side branch consisting of a length of tube terminating into a volume that separates liquid from gas using a piston with a seal around the edge.

There is an orifice downstream of the side branch that is modeled as a linear resistance R_f. Upstream there is a pump that injects fluid into the tube, and the unsteady flow induced by the pump is represented as a specified input flow $Q_{\text{in}}(t)$. The side branch has fluid inertia I_f and the volume termination has volume compliance C_a. If the side branch tube has length L_t and cross-sectional area A_t, then a good estimate of the fluid inertia parameter is

$$I_f = \rho_0 L_t / A_t,$$

where ρ_0 is the density of the liquid. This formula was presented in Chapter 4.

Furthermore, if the air volume in the accumulator is V, then a good estimate of the compliance parameter is

$$C_a = V/\rho_a c^2$$

where ρ_a is the density of air and c is the acoustic wave speed in the fluid. This was also presented in Chapter 4.

The bond graph for this system is also shown in Figure 6.17 and the two state variables are identified on the bond graph. Following causality, the equations of

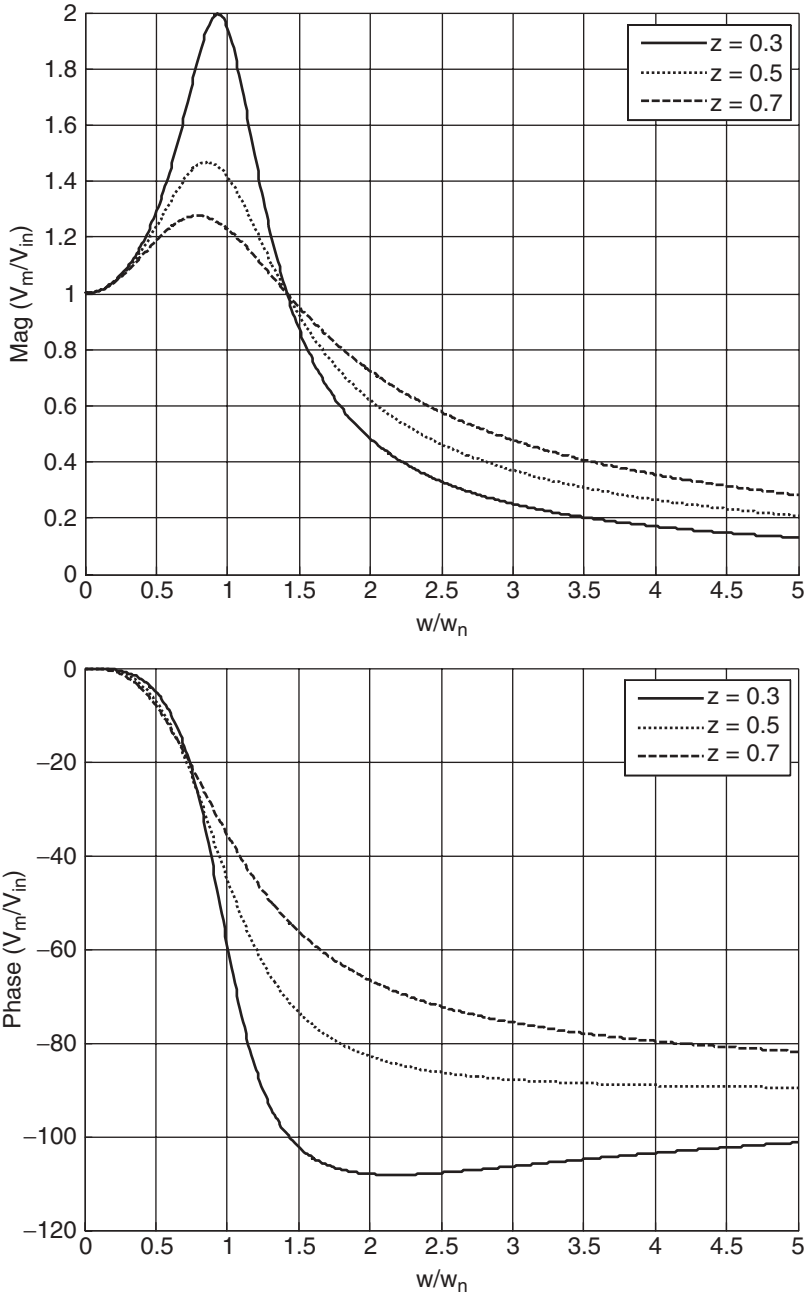


FIGURE 6.16. Frequency response for the system of Figure 6.15.

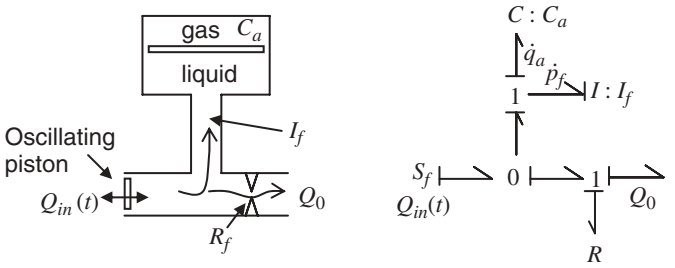


FIGURE 6.17. Side branch oscillator and bond graph.

motion come directly from the bond graph as

$$\begin{aligned} \dot{p}_f &= -(1/C_a)q_a + R_f Q_{\text{in}} - (R_f/I_f)\,p_f \\ \dot{q}_f &= (1/I_f)\,p_f \end{aligned} \tag{6.124}$$

The desired output is the flow through the orifice Q_0 and this can be expressed as

$$Q_0 = Q_{\text{in}} - (1/I_f)\,p_f. \tag{6.125}$$

Equations (6.124) can be expressed in the matrix form as

$$\frac{d}{dt}\begin{bmatrix} p_f \\ q_a \end{bmatrix} = \begin{bmatrix} -R_f/I_f & -1/C_a \\ 1/I_f & 0 \end{bmatrix}\begin{bmatrix} p_f \\ q_a \end{bmatrix} + \begin{bmatrix} R_f \\ 0 \end{bmatrix} Q_{\text{in}} \tag{6.126}$$

For eigenvalues, transfer functions, and frequency response, we always start with the s-domain; thus, let

$$Q_{\text{in}}(t) = Q_{\text{in}}(s)e^{st},\ \ p_f = P_f(s)e^{st},\ \ q_a = Q_a(s)e^{st}, \tag{6.127}$$

with the result

$$\begin{bmatrix} s + R_f/I_f & 1/C_a \\ -1/I_f & s \end{bmatrix}\begin{bmatrix} P_f \\ Q_a \end{bmatrix} + \begin{bmatrix} R_f \\ 0 \end{bmatrix} Q_{\text{in}}. \tag{6.128}$$

We desire the transfer function relating $Q_0(s)$ to $Q_{\text{in}}(s)$. From (6.125) we see that

$$\frac{Q_0}{Q_{\text{in}}}(s) = 1 - \frac{1}{I_f}\frac{P_f}{Q_{\text{in}}}(s). \tag{6.129}$$

The simplest way to derive the desired transfer function is to first derive the transfer function between $P_f(s)$ and $Q_{\text{in}}(s)$ and then carry out the operations indicated in Eq. (6.129) to obtain the final result. Thus,

$$\frac{P_f}{Q_{\text{in}}}(s) = \frac{\begin{vmatrix} R_f & 1/C_a \\ 0 & s \end{vmatrix}}{s^2 + (R_f/I_r)s + /I_f C_a} = \frac{R_f s}{s^2 + (R_f/I_r)s + /I_f C_a}. \tag{6.130}$$

Carrying out the operations from Eq. (6.129) yields the final transfer function as

$$\frac{Q_0}{Q_{\text{in}}}(s) = \frac{s^2 + 1/I_f C_a}{s^2 + (R_f/I_f)s + 1/I_f C_a}. \tag{6.131}$$

If we desire the eigenvalues, then we recognize the denominator polynomial as the characteristic equation,

$$s^2 + (R_f/I_f)s + 1/I_f C_a = 0. \tag{6.132}$$

For frequency response, the complex frequency-response function is generated by letting $s = j\omega$ in the transfer function, with the result

$$\frac{Q_0}{Q_{\text{in}}}(j\omega) = \frac{-\omega^2 + 1/I_f C_a}{-\omega^2 + (R_f/I_f)j\omega + 1/I_f C_a} = \frac{\omega_0^2 - \omega^2}{\omega_0^2 - \omega^2 + j2\varsigma_0\omega_0\omega}, \tag{6.133}$$

where a frequency ω_0 has been defined as

$$\omega_0^2 = 1/I_f C_a,$$

and a damping ratio has been introduced,

$$2\varsigma_0\omega_0 = R_f/I_f.$$

Recognizing that the numerator of (6.133) is a complex number with no imaginary part, the magnitude of the frequency response is

$$\left|\frac{Q_0}{Q_{\text{in}}}\right| = \frac{\left[(\omega_0^2 - \omega^2)^2\right]^{1/2}}{\left[(\omega_0^2 - \omega^2)^2 + (2\varsigma_0\omega_0\omega)^2\right]^{1/2}}. \tag{6.134}$$

This is an interesting result. Notice that there is a "special" forcing frequency where the oscillatory output flow goes to zero, because when $\omega = \omega_0$, the numerator of the magnitude response is zero. These side branch accumulators are used frequently in hydraulic systems to attenuate unsteadiness in the flow that is induced by reciprocating piston pumps. They are "tuned" to the proper frequency by adjusting the pressure and volume of air in the accumulator. Figure 6.18 shows this magnitude response for some selected parameter values.

6.5.2 Block Diagrams

Block diagrams were introduced briefly in Section 6.4 as a graphical representation of a transfer function. Such a representation is used frequently in the development of control systems. Consider the example system from Figure 6.3

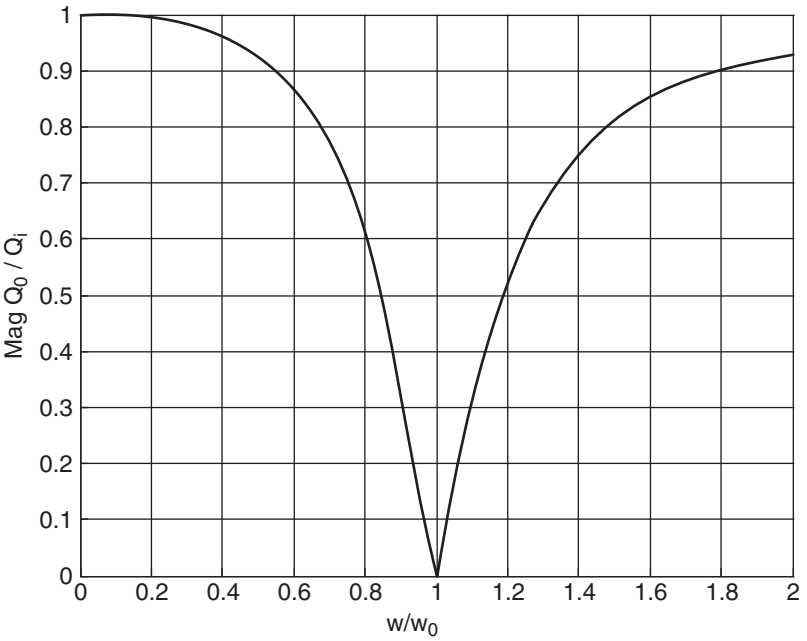


FIGURE 6.18. Magnitude response for the side branch accumulator.

with the state equations, repeated here from Eqs. (6.16),

$$\begin{aligned} \dot{x} &= 0x - (1/m)p + 1V + 0F, \\ \dot{p} &= kx - (b/m)p + bV + 1F. \end{aligned} \tag{6.135}$$

One possible output is the force indicated on the bond graph of Figure 6.3b, which is the sum of the forces in the spring and damper,

$$f = kx - (b/m)p + bV + 0F. \tag{6.136}$$

It is straightforward to derive the transfer function relating the output f to the input F as

$$\frac{f}{F}(s) = H_{f/F} = \frac{-((b/m)s + k/m)}{(s^2 + (b/m)s + k/m)} = \frac{-(2\varsigma\omega_n s + \omega_n^2)}{(s^2 + 2\varsigma\omega_n s + \omega_n^2)}. \tag{6.137}$$

Block diagrams are commonly used to represent transfer functions, and the block diagram shown in Figure 6.19*a* indicates that the output signal f is the product of the input signal F and the transfer function $H_{f/F}$.

For signals with the time dependence e^{st}, *differentiation is equivalent to multiplication by* s and *integration is equivalent to dividing by* s. Thus, the transfer function for an integrator is $1/s$. Using this notion, the block diagram of Figure 6.19*b* represents the state equations and output Eqs. (6.135) and (6.136) directly and shows that causally F is an input and f is an output.

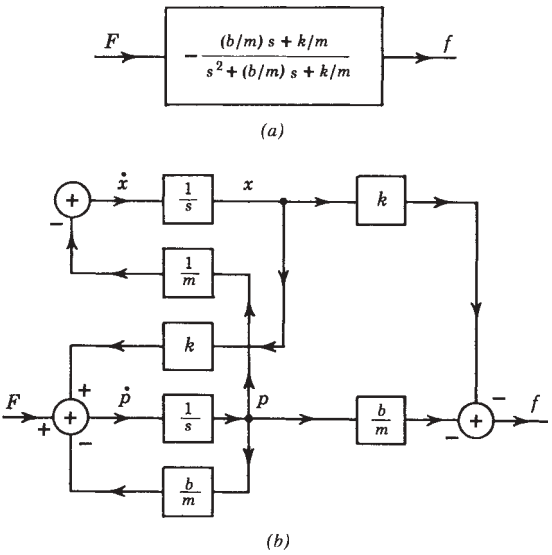


FIGURE 6.19. Block diagrams representing transfer functions between F and f: (a) combined transfer function; (b) direct representation of state and output equations.

The construction of the block diagram in Figure 6.19b is characteristic of the construction procedure for any set of state equations. One starts with an integrator and summing junction for each state variable. The derivative of the state variable is the input to the integrator and the state variable itself is the output from the integrator. The summing junctions add all the terms on the right sides of the state equations that are needed to construct the respective state derivatives. One can imagine ideal measurements of the state variables x and p in this example, which are indicated by "tapping" into the output of the respective integrators and "feeding" these measurements back to the summing junctions. For example, state variable p is measured and fed through the block containing $1/m$ such that $(1/m)p$ is fed into the upper summing junction with an appended minus sign. The output of the summing junction is $\dot{x} = -(1/m)p$, which is one of the state equations from Eq. (6.135).

It is possible to show that Figure 6.19b is equivalent to Figure 6.19a using essentially graphical means. The major tool is the well-known relation shown in the general input/output block diagram of Figure 6.20. As indicated in Figure 6.20a, the output y is related to the input x by

$$y(s) = H(s)e(s) = H(s)(x(s) - G(s)y(s)) \tag{6.138}$$

or

$$y(s) = \frac{H(s)}{1 + G(s)H(s)}x(s), \tag{6.139}$$

as indicated in Figure 6.20b.

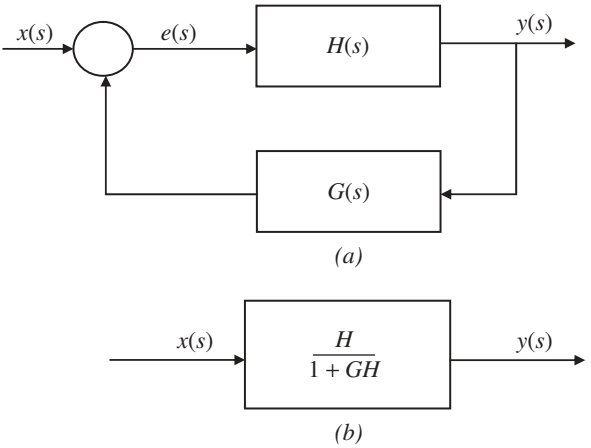


FIGURE 6.20. Block diagram reduction for feedback loops.

Although much of classical control theory and linear system analysis is concerned with block diagram algebra, it should be noted that since bond graphs generate highly organized state equations, only the simple algebra needed to solve equations such as (6.135) and (6.136) is required to generate transfer functions between any input and any output. There are some cases where block diagram reduction techniques provide convenient guides to accomplishing algebraic manipulation, but for large systems the computer-automated algebraic methods presented in Reference [2] for example, are far superior to graphical techniques in avoiding human error.

6.6 INTRODUCTION TO AUTOMATIC CONTROL

The primary emphasis of this book is physical understanding, modeling, and simulation of multienergy-domain engineering systems. Such understanding is essential to the design of actual hardware that is intended to accomplish some specific goal. We do not want to be confronted with vibration and noise problems, or thermal problems, or stability problems because of not understanding the physics of a proposed system. Hopefully, you are starting to appreciate that bond graphs help enormously in representing the physics of complex systems and in coming up with state representations of such systems.

Another primary reason for understanding the physics of a proposed system is so that an automatic controller can be developed that can change the behavior of the system in some desired fashion. An uncontrolled system (sometimes called an *open-loop* system) might not respond fast enough or with acceptable accuracy, and it may not be possible to correct such deficiencies by simply changing or redesigning the system. (Although having a model of the overall system will help a lot in assessing whether system changes can help). Automatic control uses

sensors, signal processing units, and actuators to change the inherent behavior of a system into a response that is specified and desirable. When a system incorporates automatic control it is often referred to as a *closed-loop* system.

As mentioned, when automatic control is proposed then sensors, signal processing, and actuators are needed. The sensors are used to measure specific system outputs, such as positions, velocities, pressures, and temperatures, that are going to be used as *error-generating* signals. These measurements are compared to some desired values of these outputs and the difference between the current value of an output and its desired value is the error which is sent to a signal processing unit. The signal processor is typically a computer algorithm that decides what actuation is needed to reduce the error signals. This unit then commands actuators that generate control forces, torques, pressures, etc., that move the system in such a way that error signals are reduced. Actuators require power to generate their output variables. Thus, power devices are also needed in a control system. All of the necessary devices for implementing automatic control can lead to substantial expense. Thus, a control system is rarely implemented without thorough analysis.

It is typical to start the design of an automatic control system using linear system models in which the controllable actuators can provide their respective outputs with no magnitude restrictions and without delay. There is initially no concern with power or speed of response limitations. In other words, we typically start the controller development as a complete fiction. It turns out that it is beneficial to first determine the best performance a controller can deliver before restricting the design to realistic actuators, realistic response times, and realistic power requirements. There is a wealth of literature on the topic of linear systems control and References [6], [7], and [8] are good starting points. In this chapter we emphasize control applications to physical systems and develop an understanding of what control can and cannot accomplish.

6.6.1 Basic Control Actions

Figure 6.21 shows a very simple system consisting of a mass acted upon by a control force. With respect to some inertial reference, the position, velocity, and acceleration are indicated.

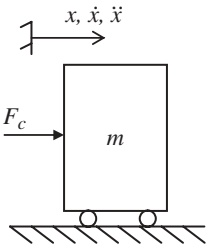


FIGURE 6.21. System consisting of a mass acted upon by a control force.

The equation of motion is simply

$$F_c = m\ddot{x}. \tag{6.140}$$

For $F_c(t) = F_c(s)e^{st}$, $x(t) = X(s)e^{st}$, the transfer functions between position and input force, velocity and input force, and acceleration and input force are

$$\frac{X}{F_c}(s) = \frac{1}{ms^2}, \quad \frac{V_m}{F_c}(s) = s\frac{X}{F_c}(s) = \frac{1}{ms}, \quad \frac{A_m}{F_c}(s) = s^2\frac{X}{F_c}(s) = \frac{1}{m}. \tag{6.141}$$

These transfer functions are shown as block diagrams in Figure 6.22.

Proportional Control. A very basic control action is called proportional control, where the control variable (F_c in this example) is proportional to the error between the system output and the desired value of the output. The controller block diagram is shown in Figure 6.23. Virtually all feedback control systems have proportional control as part of the control action.

Imagine first that we desire to control the acceleration of the mass. From Figure 6.22*c* we can derive the "closed loop" transfer function as

$$A_m = (1/m)F_c = (1/m)K_p(A_{m_{\text{des}}} - A_m) \tag{6.142}$$

or

$$\frac{A_m}{A_{m_{\text{des}}}} = \frac{K_p/m}{K_p/m + 1}. \tag{6.143}$$

Note that for this control action K_p has the dimensions of mass. As the proportional gain is increased such that $K_p/m \gg 1$, the response acceleration approaches the desired acceleration. Keep in mind that as K_p is increased, the control force is also increased and there are practical limitations that are not being considered at this juncture. For now, we simply note that proportional control works well for controlling acceleration of the idealized system.

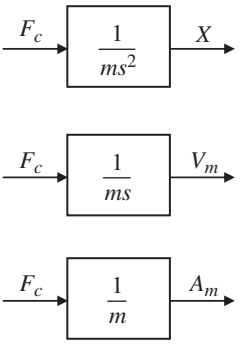


FIGURE 6.22. Block diagram representation of the transfer functions from Eqs. (6.141).

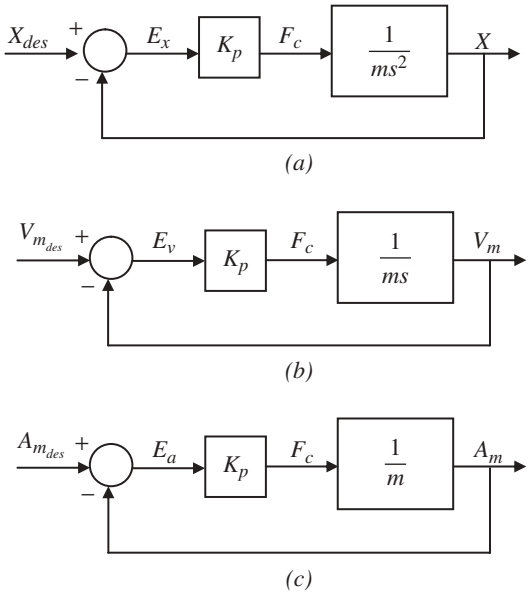


FIGURE 6.23. Controller block diagrams for proportional control of the transfer functions from Eqs. (6.141).

Next consider velocity control of the same single-mass system. For this controller, shown in Figure 6.23*b*, it proposed that the system velocity is measured and compared to some desired velocity to generate an error. A proportional controller is proposed such that the control force F_c is proportional to the velocity error. The closed-loop transfer function becomes

$$\frac{V_m}{V_{m_{\text{des}}}} = \frac{K_p/m}{s + K_p/m}. \tag{6.144}$$

For this idealized system the proportional gain K_p has the dimensions of a mechanical damping constant. The denominator of the *closed-loop* transfer function is the characteristic polynomial from which the *closed-loop* eigenvalue can be calculated; thus,

$$s + K_p/m = 0 \tag{6.145}$$

or

$$s = -K_p/m. \tag{6.146}$$

The eigenvalue is negative real for positive values of the controller gain. It is common to show closed-loop eigenvalues on a *root locus* plot, such as shown in Figure 6.24. This shows the path of the closed-loop eigenvalue as K_p is varied from zero to a large value. Obviously the *closed-loop eigenvalue* is equal to the *open-loop* eigenvalue when $K_p = 0$ and there is no control action. As the gain

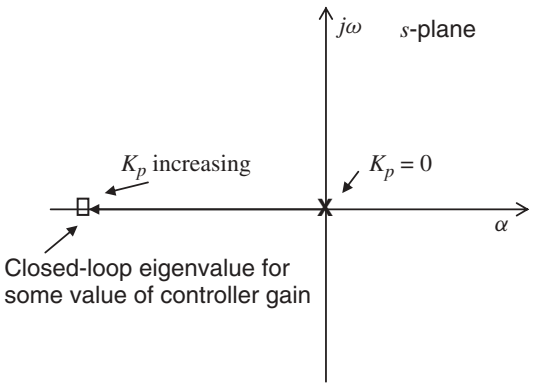


FIGURE 6.24. Root locus for the velocity controller of Eq. (6.144).

is increased the eigenvalue moves to the left and indicates a faster and faster system response.

Another way to view the effect of control is to rewrite Eq. (6.144) as

$$\frac{V_m}{V_{m_{\text{des}}}} = \frac{1}{\tau s + 1}, \tag{6.147}$$

where the time constant is $\tau = m/K_p$. This transfer function is sometimes referred to as a first-order lag or a low-pass filter. Its frequency response comes from letting $s = j\omega$, generating the magnitude response

$$\left| \frac{V_m}{V_{m_{\text{des}}}} \right| = \frac{1}{\left[1 + (\tau\omega)^2\right]^{1/2}}. \tag{6.148}$$

This response is shown in Figure 6.25 for a few values of the time constant τ.

The term *low-pass filter* comes from the fact that if $V_{m_{\text{des}}}$ is a low-frequency signal, $\left|V_m / V_{m_{\text{des}}}\right| \simeq 1$ and the magnitude of V_m is almost equal to the magnitude of $V_{m_{\text{des}}}$, but as frequency is increased the response becomes attenuated relative to the input magnitude $\left|V_m / V_{m_{\text{des}}}\right| \ll 1$. For long time constants (control gain K_p low) the attenuation begins at fairly low frequencies, while as the time constant is shortened (K_p increased) the attenuation begins at higher frequencies.

Figure 6.26 shows the step response of the closed-loop transfer function for several values of the time constant.

For short time constants (K_p high) the response rises to its final value quickly and rises less quickly as the time constant is increased (K_p low). For velocity control this proportional controller does an excellent job of tracking the desired velocity. Figures 6.25 and 6.26 were generated computationally using Matlab [5]. With a little experience one could have sketched these responses by hand for this low-order system.

Returning to the example system of Figure 6.21, it is now proposed to try proportional control for the position of the mass. It is assumed that mass position

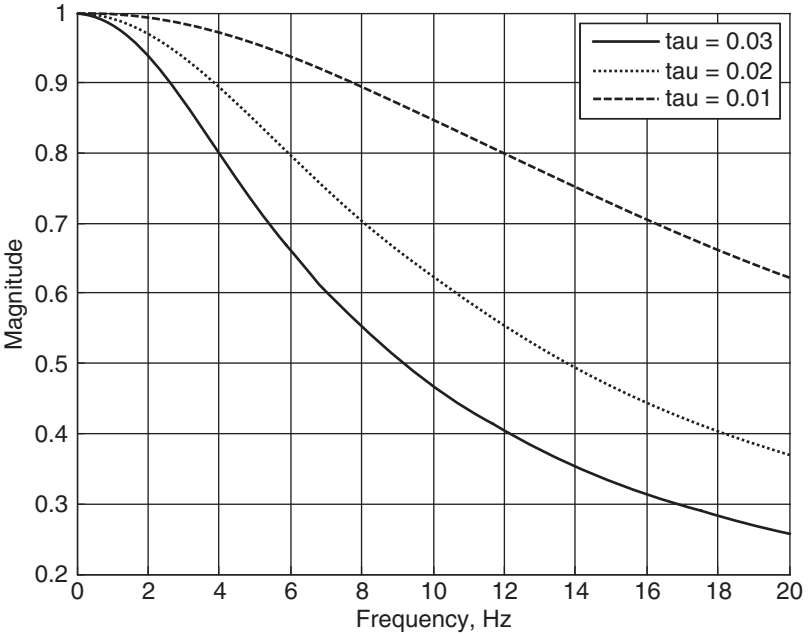


FIGURE 6.25. Magnitude response of the first-order lag transfer function.

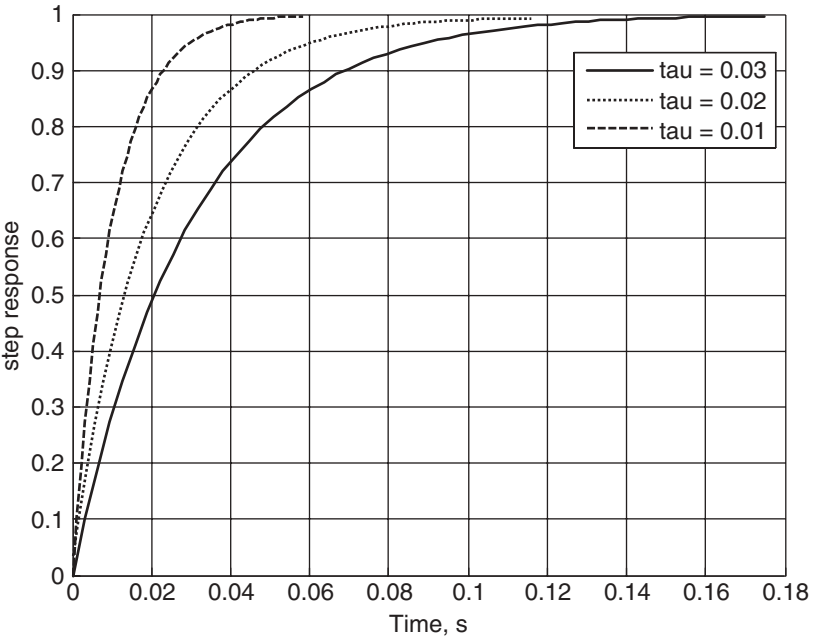


FIGURE 6.26. Step response of the first-order lag transfer function.

is measured and compared to some desired position and then the control force is set by multiplying the position error by a gain K_p. This control is shown in Figure 6.23*a*. The block diagram reduction yields the closed-loop transfer function as

$$\frac{X}{X_{\text{des}}} = \frac{K_p/m}{s^2 + K_p/m}. \tag{6.149}$$

This time the controller gain K_p has the dimensions of a spring stiffness, and the closed-loop eigenvalues come from setting the denominator polynomial to zero, with the result

$$s_{1,2} = \pm j\omega_n, \tag{6.150}$$

where $\omega_n = \sqrt{K_p/m}$. We know this indicates a continuously oscillatory response at the undamped natural frequency. We can affect the frequency by adjusting the controller gain, but we cannot make the response damp out. The root locus is shown in Figure 6.27.

As a final step in this first example of a control system let's recognize that the position is not actually measured and controlled but instead is the output from a transducer that measures a position signal and exhibits some time lag in this measurement. If the position measurement device had a characteristic such as

$$\frac{X_{\text{meas}}}{X}(s) = \frac{1}{\tau s + 1}, \tag{6.151}$$

then the measured position would respond to the actual position identically to the step response of Figure 6.26. The control system would appear as shown in Figure 6.28.

Block diagram reduction yields the closed-loop transfer function as

$$\frac{X_{\text{meas}}}{X_{\text{des}}} = \frac{K_p/m}{\tau s^3 + s^2 + K_p/m} \tag{6.152}$$

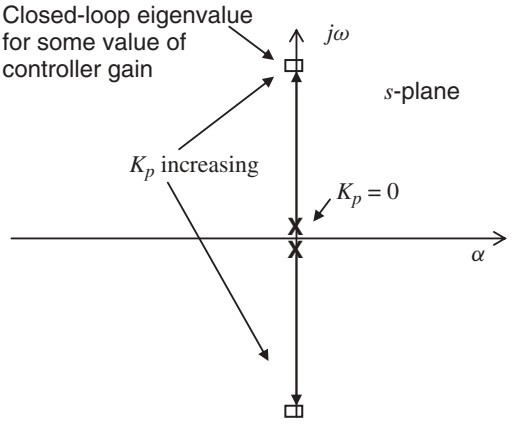


FIGURE 6.27. Root locus for the position controller of Eq. (6.149).

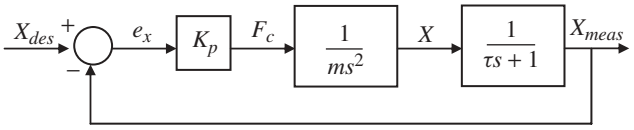


FIGURE 6.28. Position control with some measurement lag.

Notice that if $\tau = 0$ such that there is no delay in position measurement, then (6.152) reduces to (6.149). The characteristic equation again is the denominator of the transfer function set equal to zero. It can be shown that since the s-term is missing, this closed loop, controlled system is unstable using proportional control. See *Routh stability criterion* in Reference [1]. This instability is demonstrated in the root locus of Figure 6.29. Notice that a complex conjugate pair of closed-loop eigenvalues immediately move into the right half of the s-plane as the controller gain is increased from zero.

While proportional control is part of virtually all control systems, it is clearly not all that is needed as systems become more dynamically complex.

Proportional plus Derivative Control. A second control action that is often part of a control system is called *derivative control*. This action sets the control variable (force, torque, voltage, current, etc.) proportional to the derivative of the error,

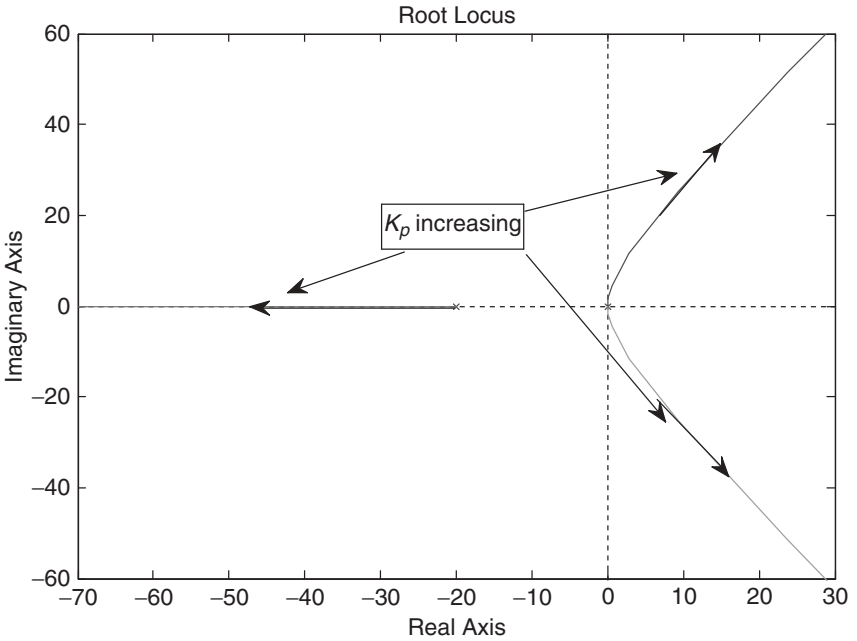


FIGURE 6.29. Root locus of the position control system with measurement lag.

such that the control force for the example system of Figure 6.21 would be

$$F_c = K_d \frac{d}{dt} e \tag{6.153}$$

or in the s-domain,

$$F_c(s) = K_d s E(s). \tag{6.154}$$

Derivative control is somewhat anticipatory in that the control variable reacts to the rate of change of the error rather than the error itself. If at some time, the error is passing through zero with a positive slope there will be no control action from a proportional controller but there will be some control action from the derivative control. The controller anticipates that the error is about to become larger and some control action is taken to slow the *growth* of the error.

If position control is the main function of the controller, then derivative action is identical to proportional control of velocity since velocity is the time derivative of the position. In a mechanical system a force component that opposes the velocity is a damping-like force, so it is often said that derivative control adds damping to a system.

Derivative control action is often combined with proportional control, resulting in the PD controller (proportional plus derivative control),

$$F_c(s) = (K_p + K_d s) E(s). \tag{6.155}$$

For the position controller from Figure 6.23*a*, this control action is shown as a block diagram in Figure 6.30.

Using block diagram reduction the closed-loop transfer function becomes

$$\frac{X}{X_{\text{des}}} = \frac{K_p/m + (K_d/m)s}{s^2 + (K_d/m)s + K_p/m}. \tag{6.156}$$

The closed-loop eigenvalues come from setting the denominator polynomial to zero, with the result

$$s_{1,2} = -\varsigma\omega_n \pm j\omega_d, \tag{6.157}$$

where $\omega_n = \sqrt{K_p/m}$, $\varsigma = K_d/2m\omega_n$, and $\omega_d = \omega_n\sqrt{1-\varsigma^2}$. The proportional control gain allows adjustment of the system frequency, and the derivative gain allows adjustment of the system damping. For this second-order, open-loop system, PD control allows arbitrary placement of the closed-loop eigenvalues

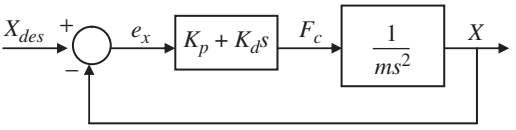


FIGURE 6.30. Controller block diagram for proportional plus derivative control of position modified from Figure 6.23*a*.

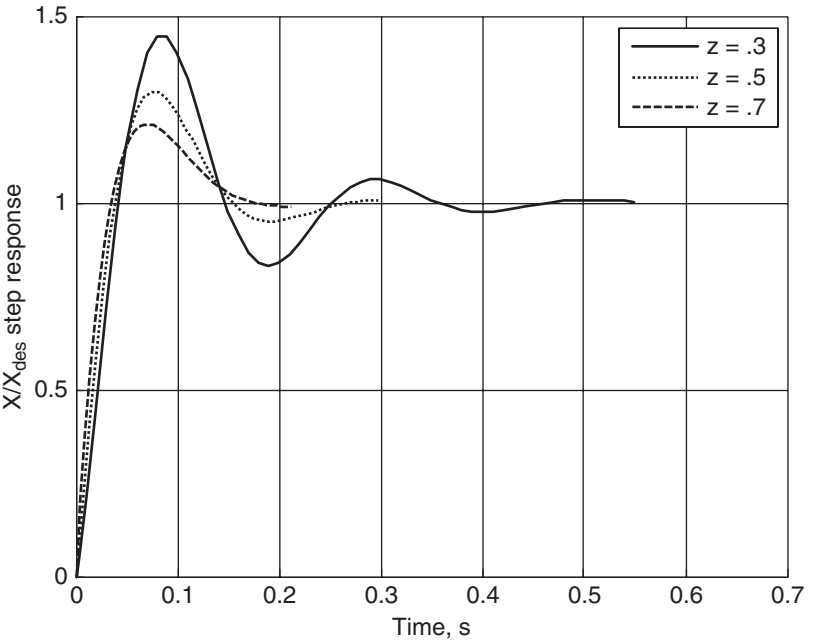


FIGURE 6.31. Step response of position controller with PD control.

anywhere in the s-plane. The derivative control term does appear in the numerator of the closed loop system and certainly does affect the dynamic response. But total adjustability of the closed-loop eigenvalues assures system stability. Figure 6.31 shows the step response of the closed-loop system for several values of the damping ratio ς and a natural frequency $f_n = 5$ Hz. Notice that for increased damping, i.e., higher values of the derivative gain K_d, there is diminished overshoot and faster settling to the final response value.

An alternative way to represent PD control from Eq. (6.155) is

$$F_c(s) = (K_p + K_d s)E(s) = K_d(s + K_p/K_d)E(s) = K(s + z)E(s), \quad (6.158)$$

where an overall system gain K is introduced along with a parameter z. Instead of specifying the proportional and derivative gains K_p and K_d, we specify the parameter z and the overall system gain K. The control system from Figure 6.30 is shown with this alternative controller in Figure 6.32.

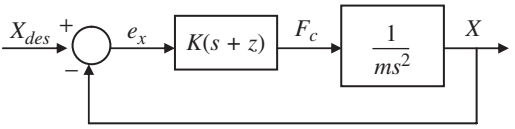


FIGURE 6.32. Controller block diagram for alternative representation of proportional plus derivative control of position modified from Figure 6.30.

Block diagram reduction yields the closed-loop transfer function as

$$\frac{X}{X_{\text{des}}} = \frac{(K/m)(s+z)}{s^2 + (K/m)s + (K/m)z} \tag{6.159}$$

Specification of the system gain and the parameter z allows arbitrary placement of the closed-loop eigenvalues in much the same way as independent specification of K_p and K_d. There is really no difference in the two representations of PD control, but the appearance of the numerator polynomial in Eq. (6.159) allows introduction of the concept of *zeros* and *poles*.

Transfer functions are always ratios of polynomials, with the general form

$$\frac{Y}{U}(s) = \frac{K(s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m)}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}. \tag{6.160}$$

The solutions to the denominator polynomial set to zero are the eigenvalues or *poles* of the system. For the uncontrolled system there are *open-loop poles* and for the controlled system there are *closed-loop poles*. The solutions for the numerator polynomial set to zero are the system *zeros*, and there can be *open-* and *closed-loop zeros* for a system.

Much like the poles or eigenvalues, the zeros can be real or they can appear as complex conjugate pairs. It is typical to show the open-loop poles and zeros on the complex s-plane, as is done in Figure 6.33 for the example system.

When the control system is represented as shown in Figure 6.32 it is typical to generate the root locus as the overall system gain is varied. For this example we fix the location of the zero and then vary the overall gain to determine the path of the closed-loop poles. This root locus is shown in Figure 6.34.

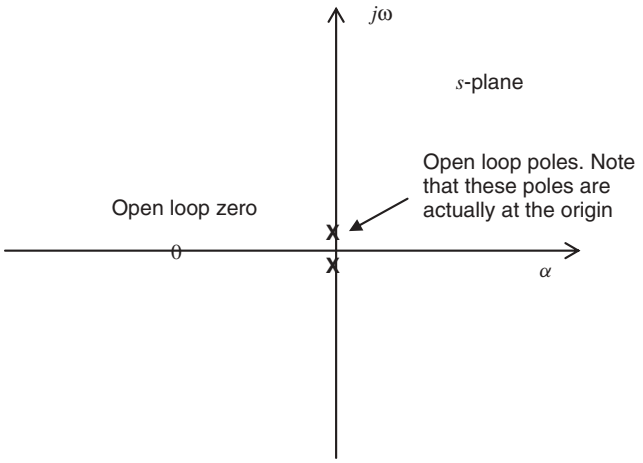


FIGURE 6.33. Open-loop poles and zeros for the control system of Figure 6.32.

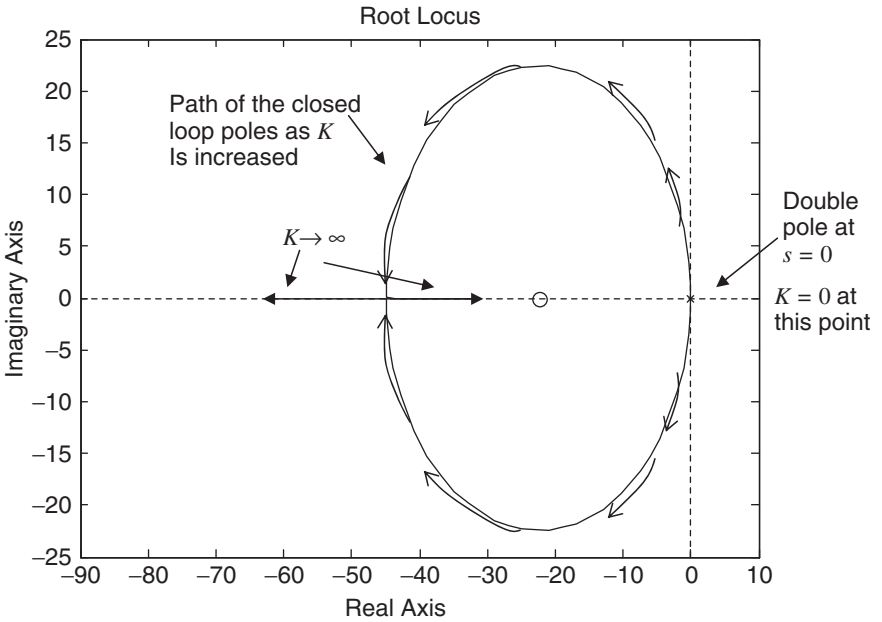


FIGURE 6.34. Root locus for PD position control of the system from Figure 6.32.

The closed-loop poles start at the open-loop poles for $K = 0$ and then become complex conjugate poles as the gain is increased. Ultimately, the root locus rejoins the real axis and we get repeated poles followed by one pole moving toward the open-loop zero and the other pole moving toward infinity as the gain is increased. It is useful to be able to sketch a root locus by hand to determine the possibility of attaining desirable closed-loop behavior. The rules for doing this are discussed in the next section.

Looking again at the step response for the closed-loop system, Figure 6.31, we see that the final value of the response is equal to the magnitude of the step input. For the demonstration system the step input magnitude is unity. The value of the response as time goes toward infinity is called the *final value*. This value can be determined by simply letting $s = 0$ in the closed-loop transfer function from Eq. (6.156) or Eq. (6.159). Evaluating the final value of a response formally comes from the final value theorem and the interested reader can see the derivation of this theorem in Reference [8]. For the example used here to develop the concepts of feedback control, the response had the desirable attribute of having its final value equal to the desired step input. This is common but not always the case.

Proportional plus Derivative plus Integral Control. Consider the example system shown in Figure 6.35. It consists of a mass-spring-damper with a control force F_c.

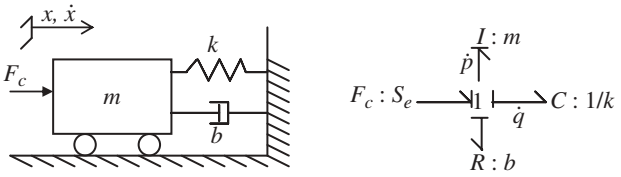


FIGURE 6.35. Example system and bond graph.

It is desired to design a position controller for the mass. The bond graph for the system is shown in the figure and the state equations come from the bond graph:

$$\begin{aligned} \dot{p} &= F_c - kq - (b/m)p, \\ \dot{q} &= (1/m)p. \end{aligned} \tag{6.161}$$

In the s-domain these become

$$\begin{bmatrix} s + b/m & k \\ -1/m & s \end{bmatrix} \begin{bmatrix} P(s) \\ Q(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} F_c(s). \tag{6.162}$$

Recognizing that the mass displacement x is the state variable q we derive the transfer function relating F_c to q using Cramer's rule, with the result

$$\frac{X}{F_c}(s) = \frac{1/m}{s^2 + (b/m)s + k/m}. \tag{6.163}$$

Equation (6.163) is the open-loop transfer function, shown with PD control in Figure 6.36.

Block diagram reduction yields the closed-loop transfer function between desired and actual response as

$$\frac{X}{X_{\text{des}}}(s) = \frac{(K_d/m)s + K_p/m}{s^2 + (b/m + K_d/m)s + k/m + K_p/m} \tag{6.164}$$

From the denominator we see that the closed-loop eigenvalues can be placed arbitrarily in the s-plane, which suggests good adjustment of the system response,

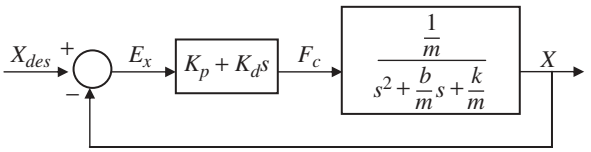


FIGURE 6.36. Controller block diagram for proportional plus derivative control of position for the system from Figure 6.35.

but letting $s \to 0$ in the transfer function shows that the final value of the response to a step change in desired position is *not* unity:

$$\frac{X}{X_{\text{des}}}|_{s\to 0} = \frac{K_p/m}{k/m + K_p/m}. \tag{6.165}$$

From the block diagram of Figure 6.36 we can also derive the error between X_{des} and X, with the result

$$\frac{E_x}{X_{\text{des}}}(s) = \frac{s^2 + (b/m)s + k/m}{s^2 + (b/m + K_d/m)s + k/m + K_p/m}. \tag{6.166}$$

Letting $s \to 0$ yields the final value of the error and we note that its final value is nonzero:

$$\frac{E_x}{X_{\text{des}}}|_{s\to 0} = \frac{k/m}{k/m + K_p/m}. \tag{6.167}$$

As the gain K_p is made larger and larger, the final value of the response approaches the desired response and the error is driven toward zero.

The issue with this physical system is that when the mass motion comes to rest, the control force must oppose the spring force. Using PD control, the only component of the control force that is nonzero when the system is at rest is the part due to proportional control, and proportional control can generate a force only when an error is present. Thus, there is finite error when the mass is at its final position.

To alleviate this problem we introduce the concept of *integral control*. Integral control generates a control force proportional to the integral of the error such that

$$F_c = K_I \int e\,dt \tag{6.168}$$

or, in the s-domain,

$$F_c(s) = (K_I/s)E(s). \tag{6.169}$$

If a component of the control force is related to the integral of the error then that component will continue to get larger as long as an error is present and, with proper usage, will drive the error toward zero.

When integral control is added to the proportional plus derivative control we obtain proportional plus integral plus derivative control or PID control. It is represented as

$$F_c(t) = K_p e + K_d \dot{e} + K_I \int e\,dt, \tag{6.170}$$

in the time domain or in the s-domain as

$$F_c(s) = [K_p + K_d s + (K_I/s)]E(s) \tag{6.171}$$

For the example system of Figure 6.35, the controller block diagram using a PID controller is shown in Figure 6.37.

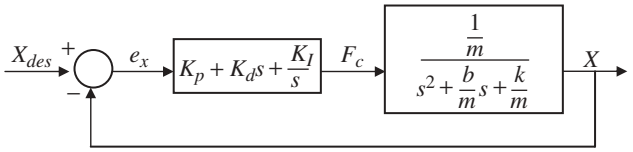


FIGURE 6.37. Controller block diagram for proportional plus derivative plus integral control of position for the system from Figure 6.35.

It is straightforward to derive the closed-loop transfer function, with the result

$$\frac{X}{X_{\text{des}}}(s) = \frac{(K_d/m)s^2 + (K_p/m)s + K_I/m}{s^3 + (b/m + K_d/m)s^2 + (k/m + K_p/m)s + K_I/m}. \tag{6.172}$$

To test for the final value of a step response, let $s = 0$ in this transfer function, with the result

$$\frac{X}{X_{\text{des}}}|_{s\to 0} = \frac{K_I/m}{K_I/m} = 1 \tag{6.173}$$

Thus, the integral control component of the control force moved the mass position to the desired final position while driving the position error to zero. However, another result of the integral control component was to increase the order of the system from 2 to 3. Notice that the denominator of the closed-loop system from Eq. (6.172) is now third order, thus there are three closed-loop eigenvalues where the open-loop system has only two. It turns out that the higher-order systems are more easily made unstable due to the action of control systems, thus integral control should be introduced with careful consideration.

The s-domain representation of PID control can be put into the form

$$F_c(s) = (K_p + K_d s + K_I/s)E(s) = [(K_d s^2 + K_p s + K_I)/s]E(s), \tag{6.174}$$

and this can be further reduced to

$$F_c(s) = [K(s^2 + 2\varsigma_0\omega_0 s + \omega_0^2)/s]E(s), \tag{6.175}$$

where

$$K = K_d,\ 2\varsigma_0\omega_0 = \frac{K_p}{K_d},\ \omega_0^2 = \frac{K_I}{K_d}. \tag{6.176}$$

Rather than independent specification of the three controller gains K_P, K_d, K_I, we can think of PID control as allowing placement of two arbitrary zeros anywhere in the s-plane plus an overall system gain K. There is a pole at the origin that always accompanies PID control. If we choose to have a complex zero, then we find that the complex conjugate will also be a zero. The path of the closed-loop poles as the system gain is varied over a broad range is shown in Figure 6.38 for the case where the physical system is lightly damped and oscillatory at 5 Hz and the zeros of the PID controller are placed as shown. Notice that there is a range of overall system gain K where the closed-loop system is

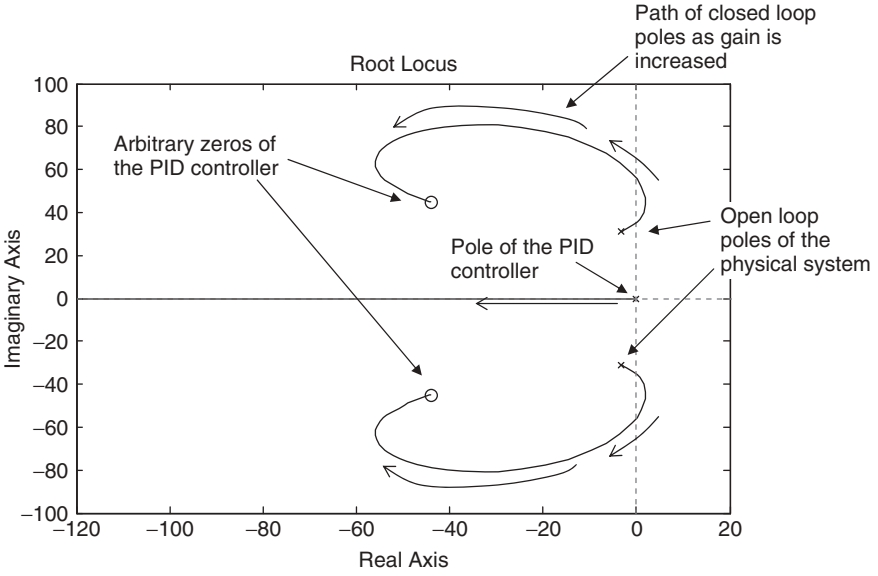


FIGURE 6.38. Root locus for PID control of position for the system from Figure 6.35.

unstable, i.e., closed-loop eigenvalues are in the right half of the s-plane. It is also true that PID control can produce desirable closed-loop poles and the integral control component assures zero steady-state error.

6.6.2 Root Locus Concept

In the example systems used to demonstrate the concepts of automatic control the concept of a *root locus* was introduced. The root locus shows the path of the closed-loop eigenvalues in the complex s-plane as some system gain is varied. Since we now have a good understanding of the meaning of an eigenvalue with respect to speed of system response, tendency to oscillate, stability, and so forth, the root locus is a very useful tool for visualizing the potential response of a system for a given control action.

In 1950 W. R. Evans developed a formal graphical procedure for creating a root locus from knowledge of the open-loop poles and zeros of the system and the poles and zeros of the proposed controller. References [9] and [10] are the original publications of his method and References [6], [7], and [8] all show how to apply the rules of construction. Before the availability of modern linear system analysis computer programs, the root locus method was used to actually determine the gain values needed to place closed-loop poles in desired locations. With programs such as Matlab [5] there is no need to use Evan's formalism and we can obtain a root locus easily. However, it is still useful to be able to sketch a root locus and one does not need all the formal development to be able to do this. The basic rules are presented next. The interested reader should see any of the above-mentioned references for a more complete development.

The Starting Point. A bond graph model of a linear physical system allows derivation of state and output equations useful for control studies in the form

$$\begin{aligned}[\dot{x}] &= [A][x] + [b_d]\,u_d + [b_c]\,u_c,\\ y &= [c]^t[x],\end{aligned} \tag{6.177}$$

where $[x]$ is the vector of state variables composed of momentum and displacement variables (p's and q's) from I- and C-elements in integral causality, u_d is a single disturbance input (effort or flow), i.e., an input over which we have no control, and u_c is a single control input (effort or flow) over which we do have control. The single output y is some combination of the states and we want to control this output to have some desired response. Obviously, in any given system there can be multiple disturbance inputs and/or multiple controlled inputs. There can also be multiple outputs we would like to control. For now we propose a single control variable u_c and a single output variable y and develop concepts for *single input–single output* (SISO) systems.

In Section 6.4 the concept of an input/output transfer function was introduced. The transfer function between the output y and control input u_c is called the *plant transfer function*,

$$G_p(s) = \frac{Y(s)}{U_c(s)}. \tag{6.178}$$

This transfer function is typically the ratio of two polynomials, such that

$$G_p(s) = \frac{K(s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m)}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}. \tag{6.179}$$

We know that the denominator polynomial is the plant characteristic equation, which delivers the plant or open-loop eigenvalues when set to zero and solved. The eigenvalues are called *poles* of the transfer function. The numerator, when set to zero and solved, yields the open-loop *zeros*. Thus, Eq. (6.179) can be represented as

$$G_p(s) = \frac{K_p(s+z_1)(s+z_2)\cdots(s+z_m)}{(s+p_1)(s+p_2)\cdots(s+p_n)} = K_p\frac{N_p(s)}{D_p(s)}, \tag{6.180}$$

where the zeros and poles, z_i and/or p_i, can be real or complex conjugate pairs.

The controller is also typically represented by a transfer function. (See Eq. (6.175) for a PID controller for example.) The closed-loop system can be represented in block diagram form as shown in Figure 6.39 in which the controller can also be expanded into its poles and zeros:

$$\begin{aligned}G_c(s) &= \frac{K_c(s^{m_c} + b_{1_c}s^{m_c-1} + \cdots + b_{m_c-1}s + b_{m_c})}{s^{n_c} + a_{1_c}s^{n_c-1} + \cdots + a_{n_c-1}s + a_{n_c}}\\ &= \frac{K_c(s+z_{c_1})(s+z_{c_2})\cdots(s+z_{c_m})}{(s+p_{c_1})(s+p_{c_2})\cdots(s+p_{c_n})} = K_c\frac{N_c(s)}{D_c(s)}.\end{aligned} \tag{6.181}$$

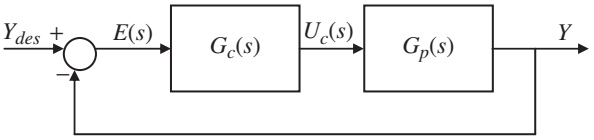


FIGURE 6.39. Block diagram of the plant and controller transfer functions.

The fundamental rule for determining a root locus comes from the block diagram reduction of Figure 6.39,

$$\frac{Y}{Y_{\text{des}}}(s) = \frac{KG_c(s)G_p(s)}{1 + KG_c(s)G_p(s)}, \tag{6.182}$$

where $K = K_c K_p$ is the overall system gain, $G_c G_p$ is called the open-loop transfer function, and the poles and zeros from Eqs. (6.180) and (6.181) are called the open-loop poles and open-loop zeros, respectively. The denominator of (6.182) is the closed-loop characteristic equation, and the closed-loop eigenvalues come from the solution of

$$1 + KG_c(s)G_p(s) = 0$$

or

$$KG_c(s)G_p(s) = -1. \tag{6.183}$$

Since G_c and G_p can each be written in terms of numerator and denominator polynomials as shown in Eqs. (6.180) and (6.181), Eq. (6.183) can also be written as

$$K\frac{N_c}{D_c}\frac{N_p}{D_p} + 1 = K\frac{N}{D} + 1 = 0, \tag{6.184}$$

where $N = N_c N_p$, $D = D_c D_p$. This can be further expanded into

$$D + KN = 0 \tag{6.185}$$

Equation (6.185) indicates that if the overall system gain K is zero, then $D = 0$ and the closed-loop eigenvalues equal the open-loop eigenvalues, and the root locus starts at the open-loop eigenvalues. As $K \to \infty$, D is no longer important in (6.185) and the only solution is $N = 0$, which can be true only if the closed-loop eigenvalues equal the open-loop zeros. Thus, we say that the root locus starts at the open-loop poles and ends at the open-loop zeros. We will see that the root locus may approach infinity in the s-plane as the gain is increased, so we consider $s \to \infty$ as a possible open-loop zero location.

Equation (6.183) indicates that if some complex number $s = \alpha + j\omega$ is a closed-loop eigenvalue, then

$$\begin{aligned} &\left|KG_c(s)G_p(s)\right| = 1 \\ &\angle KG_c(s)G_p(s) = n180^\circ, n = \pm 1, \pm 3, \pm 5, \ldots . \end{aligned} \tag{6.186}$$

In words this states that for any complex number s substituted into the open-loop transfer function $KG_c(s)G_p(s)$, that complex number is also a closed-loop eigenvalue if the magnitude of the open-loop transfer function is unity and the angle of the open-loop transfer function is 180° or an appropriate multiple of 180°. Keep in mind that, from Eqs. (6.180) and (6.181),

$$\left|KG_c(s)G_p(s)\right| = \frac{K\left|s+z_{c_1}\right|\left|s+z_{c_2}\right|\cdots|s+z_1|\,|s+z_2|}{\left|s+p_{c_1}\right|\left|s+p_{c_2}\right|\cdots|s+p_1|\,|s+p_2|} \tag{6.187}$$

and

$$\begin{aligned}\angle KG_c(s)G_p(s) = {} & \angle(s+z_{c_1})+\angle(s+z_{c_2})+\cdots+\angle(s+z_1)+\angle(s+z_2)+\cdots \\ & -\angle(s+p_{c_1})-\angle(s+p_{c_2})-\cdots-\angle(s+p_1) \\ & -\angle(s+p_2)-\cdots.\end{aligned} \tag{6.188}$$

A graphical interpretation of the magnitude and angle contribution of a pole or zero is shown in Figure 6.40.

For this example a real pole is used. This real pole is shown at $s=-p$ in the complex s-plane. Also shown is the complex number $(p+j0)$ and the "test" point s. Vectors are drawn from the origin to the complex number $+p$ and from the origin to the test point. A vector sum is taken producing the vector $s+p$. The length of this vector is $|s+p|$ and the angle relative to the positive real axis is $\angle(s+p)$. The magnitude and angle would then be used as their respective contributions to Eqs. (6.187) and (6.188) to determine if the test point is a closed-loop eigenvalue and thus on the root locus.

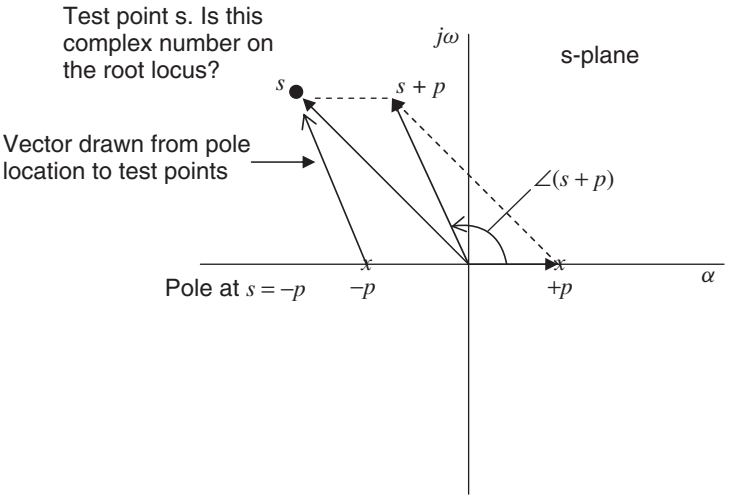


FIGURE 6.40. Graphical interpretation of magnitude and angle of a pole or zero.

There is a much simpler way to determine the magnitude and angle contribution of the poles and zeros. Figure 6.40 also shows a vector drawn from the pole location $-p$ to the test point. This vector is identical to the vector $s + p$. Thus, the simplest way to determine the magnitude and angle contribution of each open-loop pole and zero is to draw a vector from each to a test point. Then the length of each vector is the respective magnitude contribution and the angle of each vector is the angle contribution. This is also true for complex poles and it is left as an exercise for the reader to show this. Figure 6.41 shows a few poles and zeros and their magnitude and angle contributions for a test point.

A root locus sketch starts by showing the open-loop poles and zeros on the complex s-plane. As mentioned at the beginning of this section, in the past the poles and zeros were represented accurately in the s-plane and magnitudes and angles were calculated or measured so that the root locus could be precisely drawn. We are now able to obtain a root locus using computational procedures. But it is still useful to be able to sketch an approximate root locus to determine the utility of a proposed control action. Here are several rules for sketching a root locus.

Rule 1. As shown in Figure 6.41 and repeated here, the root locus starts at the open-loop poles for $K = 0$ and terminates at the open-loop zeros (including possibly the zeros at infinity) for $K \to \infty$.

Rule 2. The root locus exists on the real axis only to the left of an odd number of real poles and/or zeros. This requirement comes from the angle condition of Eq. (6.186). Consider the poles and zeros shown in Figure 6.42.

There is a real pole and real zero and a complex conjugate pair of poles. One test point is located between the real pole and real zero. The angle contribution of the real pole to the test point is 180° and the angle contribution of the zero is zero degrees. Notice that the complex poles (or complex zeros) always contribute a total of 360° to any test point on the real axis and thus need not be considered

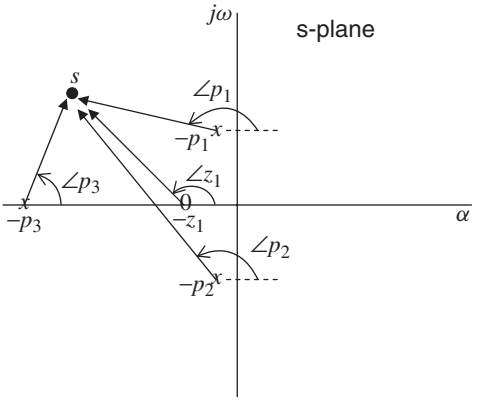


FIGURE 6.41. Graphical interpretation of magnitude and angle contribution of several poles and zeros.

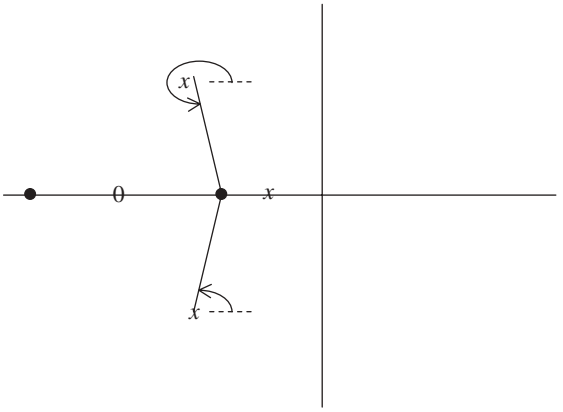


FIGURE 6.42. Angle condition for the real axis.

further. If we add the total contribution of the zeros (1 in this case) and subtract the total contribution of the poles according to Eq. (6.188), we see that the resulting angle at the first test point is -180°, which shows that this part of the real axis is on the root locus.

The second test point is to the left of the real zero in Figure 6.42. The angle contribution of the real pole is 180° and the angle contribution of the real zero is 180°. By adding the angle contributions of the zero and subtracting the angle contribution of the pole, the resulting total is 0 degrees, which does not meet the angle requirement for being on the root locus. Thus, to the left of an even number of poles and/or zeros the root locus does not exist on the real axis.

Rule 3. Complex poles and zeros, i.e., those with imaginary parts, will always appear in complex conjugate pairs. Thus, the root locus is symmetrical with respect to the real axis.

Rule 4. Remember that the root locus starts at the open-loop poles and ends at the open-loop zeros, which might include zeros at infinity. If there are several more poles than zeros then parts of the root locus must approach infinity. They do so on asymptotes and it is useful to know how many asymptotes there are, what angle these asymptotes make with respect to the real axis, and where these asymptotes intersect the real axis.

In general, the open-loop transfer function has the appearance shown in Eqs. (6.179) and (6.181), where these two transfer functions are multiplied to generate the final open-loop transfer function. As the test point $s \to \infty$, the open-loop transfer function would approach

$$KG_c(s)G_p(s)|_{s\to\infty} \to \frac{K}{s^{(n_{\text{tot}}-m_{\text{tot}})}}, \tag{6.189}$$

where $m_{\text{tot}} = m_c + m$ is the total number of finite zeros, and $n_{\text{tot}} = n_c + n$ is the total number of poles.

The angle of the open loop transfer function is the angle of the asymptote and to be on the root locus,

$$-(n_{\text{tot}} - m_{\text{tot}})\angle s = \pm n180 \quad n = 1, 3, 5, \ldots. \tag{6.190}$$

Thus, the angles of the asymptotes become

$$\angle s = \frac{\pm n180^\circ}{(n_{\text{tot}} - m_{\text{tot}})} \quad n = 1, 3, 5, \ldots. \tag{6.191}$$

The quantity $n_{\text{tot}} - m_{\text{tot}}$ is called the pole/zero excess and the number of asymptotes is equal to this quantity.

Rule 5. The asymptotes intersect the real axis at the *centroid* of the pole/zero configuration. Stated without proof, the centroid is located on the real axis at

$$\sigma = \frac{\sum p_i - \sum z_j}{n_{\text{tot}} - m_{\text{tot}}}, \tag{6.192}$$

where $\sum p_i$ is the sum of all the poles and $\sum z_i$ is the sum of all the zeros. Note that the imaginary part of any pole or zero does not contribute to this calculation since the conjugate is also part of the calculation. Also note that real parts of poles located in the left half plane move the centroid to the left while real parts of zeros located in the left half plane move the centroid to the right.

Figure 6.43 shows some pole/zero configurations. In part (*a*) there are two poles and no zeros. Thus, the pole/zero excess is 2 and there are two asymptotes. The angle of the asymptotes from Eq. (6.191) is

$$\angle s = \frac{\pm 180^\circ}{2} = \pm 90^\circ$$

and the centroid from Eq. (6.192) is

$$\sigma = \frac{-p + (-p)}{2} = -p.$$

This shows that for large values of the system gain the root locus (i.e., the closed loop poles) will approach infinity on these asymptotes.

Figure 6.43*b* shows two poles and one zero, thus the pole/zero excess equals 1. This tells us there is 1 asymptote and the angle of this asymptote is

$$\angle s = \frac{\pm 180^\circ}{1} = 180^\circ.$$

The centroid has no significance in this case.

Figure 6.43*c* shows three poles and no zeros, thus the pole/zero excess equals 3. There are three asymptotes and their angles are

$$\angle s = \frac{\pm 180^\circ}{3} = \pm 60^\circ \text{ (this accounts for 2 of the 3 asymptotes)}$$

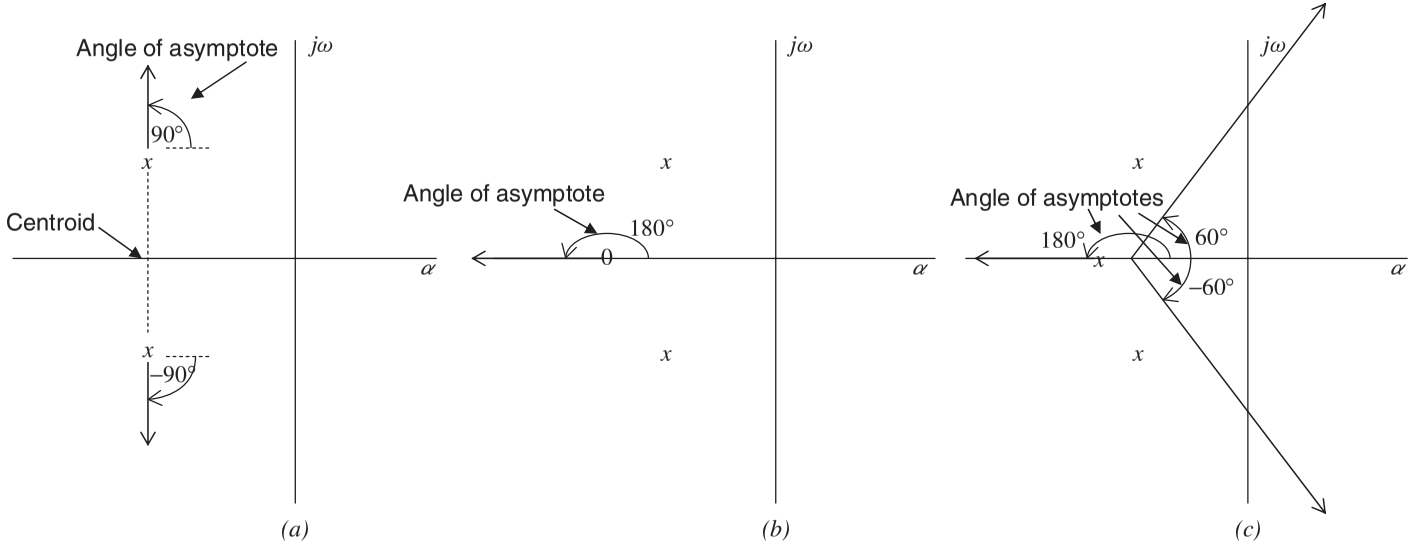


FIGURE 6.43. Some pole/zero configurations and associated asymptotes.

and

$$\angle s = \frac{\pm 3(180^\circ)}{3} = 180^\circ.$$

The centroid of the asymptotes comes from (6.192) with the approximate result as shown.

Rule 6. If the root locus exists between two poles on the real axis then the root locus, which starts at these poles, must *break away* from the real axis at some point between them. If the root locus exists between two zeros on the real axis then a *break-in* point must exist. This rule follows from the fact that the root locus starts (for low gain values) at the open-loop poles and ends (for large gain values) at the open-loop zeros. If the root locus lies between two poles on the real axis, then, as the gain is increased, the locus must escape the real axis to move toward the zeros.

There are formal ways to determine the exact location of breakaway or break-in points, but for purposes of sketching the root locus we need only realize that such points exist.

There are more features of root loci that could be developed. For example, the root locus must depart from complex conjugate poles and must arrive at complex conjugate zeros. It is possible to determine the angle of departure from the poles and angle of arrival at zeros. The interested reader should consult the references at the end of this chapter for details. For sketching purposes we need only realize that such departures and arrivals exist.

Let's revisit the physical system consisting of the position control of a single mass where we postulate the use of PID control. The physical system and block diagram are shown in Figure 6.44. The PID controller is shown in the form presented earlier in Eq. (6.175). We would first choose the damping and frequency ς_0, ω_0 and then vary the gain K to sketch the root locus.

Figure 6.45 sketches the root locus using the rules just presented.

First note that the system contributes a double pole at zero and the controller contributes an additional pole at zero. These poles are all together at zero, but it is

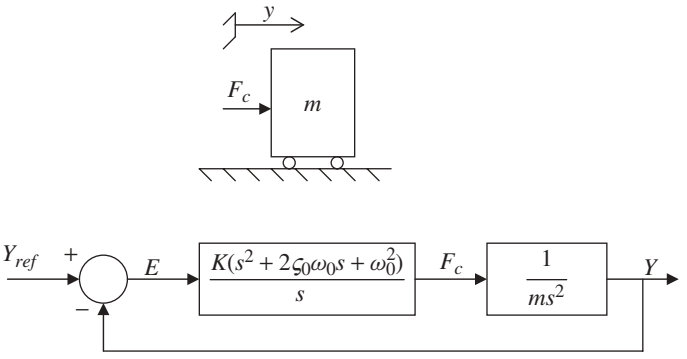


FIGURE 6.44. Physical system and block diagram for PID control of mass position.

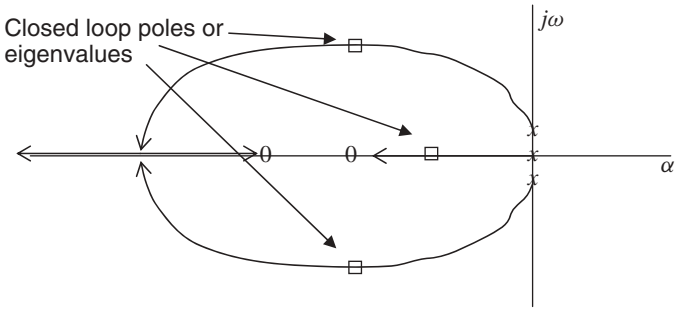


FIGURE 6.45. Sketch of root locus for the system of Figure 6.44.

easier to sketch the root locus if we "think" of them as being very close together at zero. The two zeros of the PID controller are shown in arbitrary positions and could be moved to other locations to see their effect on the location of the closed-loop eigenvalues.

Application of rule 2 indicates that the root locus exists on the real axis between the real pole at the origin and the first zero, and between the second zero and the zero at infinity. Since the pole/zero excess equals 1, rule 4 indicates that there is one asymptote and its angle is 180°. Finally, from rule 6 there must be a break in point between the second zero on the real axis and the zero at infinity. As the system gain is increased the root locus must leave the open-loop poles and move toward the open loop zeros. The closed loop eigenvalues must lie on the root locus, and the root locus will appear as sketched in Figure 6.45. Figure 6.46 shows the sketch of the root locus if the zeros of the PID controller are chosen as complex conjugate roots.

The root locus exists on the real axis between the pole at the origin and the zero at infinity. There is still only one asymptote. The root locus must leave the open-loop poles for small gain values and arrive at the open-loop zeros as the gain is increased. Thus, the root locus will appear as shown in the sketch and the closed-loop poles will lie on this root locus.

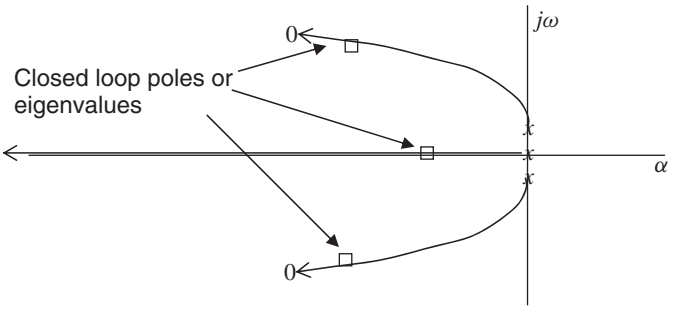


FIGURE 6.46. Sketch of root locus from Figure 6.45 with different controller zero locations.

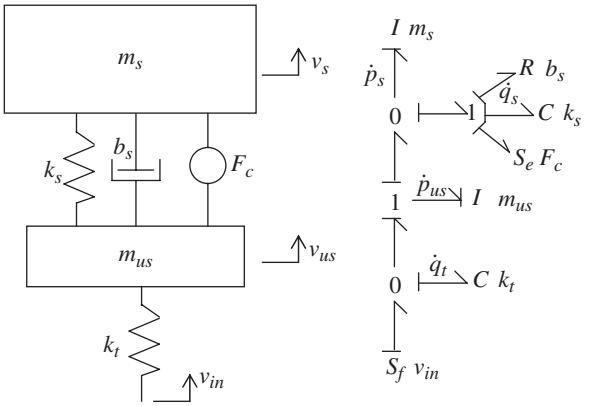


FIGURE 6.47. Schematic and bond graph of a quarter-car model with active suspension.

As a final example of obtaining a root locus consider the quarter-car model shown in Figure 6.47. The model consists of sprung and unsprung mass elements m_s, m_{us}, suspension stiffness and damping k_s, b_s, base disturbance velocity input v_{in}, and controllable force F_c. Ultimately we need to model a real actuator that can generate this fictitious force, but for now we simply propose that we can generate and control a prescribed force.

A bond graph of the model is also included in Figure 6.47, and causality indicates that there are four state variables, which are labeled on the bond graph. The equations of motion come directly from the bond graph as

$$\frac{d}{dt}\begin{bmatrix} p_s \\ p_{\text{us}} \\ q_s \\ q_t \end{bmatrix} = \begin{bmatrix} -b_s/m_s & b_s/m_{\text{us}} & k_s & 0 \\ b_s/m_s & -b_s/m_{\text{us}} & -k_s & k_t \\ -1/m_s & 1/m_{\text{us}} & 0 & 0 \\ 0 & -1/m_{\text{us}} & 0 & 0 \end{bmatrix} \begin{bmatrix} p_s \\ p_{\text{us}} \\ q_s \\ q_t \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} F_c + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} v_{\text{in}}. \tag{6.193}$$

If these equations are put into the s-domain, the transfer function between the sprung mass velocity v_s and the control force F_c can be derived as

$$\frac{V_s}{F_c}(s) = \frac{(1/m_s)s(s^2 + k_t/m_{\text{us}})}{s^4 + (b_s/m_s + b_s/m_{\text{us}})s^3 + (k_s/m_{\text{us}} + k_t/m_{\text{us}} + k_s/m_s)s^2 + (b_s k_t/m_s m_{\text{us}})s + k_s k_t/m_s m_{\text{us}}}. \tag{6.194}$$

If we are interested in the frequency response of the sprung mass velocity to a harmonic input of the control force, we would let $s = j\omega$ and follow the procedure from Section 6.5. Without carrying out all the steps we can observe that the denominator indicates there are two distinct natural frequencies (if we think of the damping being temporarily equal to zero). From the numerator dynamics there is a special frequency where the sprung mass will remain motionless under

the action of the control force. This frequency is

$$\omega_0^2 = k_t/m_{\text{us}}$$

and is often referred to as the wheelhop frequency, since it involves the tire stiffness and the unsprung mass.

For the root locus we observe that the open-loop transfer function has three zeros and four poles and they are distributed approximately as shown in Figure 6.48. There is a zero at the origin and complex conjugate zeros on the imaginary axis. There are two sets of complex conjugate poles, which are shown in Figure 6.48 as lightly damped.

Since the pole/zero excess is unity, there is only one asymptote along the real axis. The root locus would look something like that shown in the figure.

In the area of vibration control we often look at the response of a system to a disturbance input rather than a desired input. For the quarter-car the disturbance input is the ground input $v_{\text{in}}(t)$. Equations (6.193) yield the open-loop transfer function between the sprung mass velocity v_s and the disturbance velocity:

$$\frac{V_s}{V_{\text{in}}}(s) = \frac{(k_t/m_{\text{us}})[(b_s/m_s)s + k_s/m_s]}{s^4 + (b_s/m_s + b_s/m_{\text{us}})s^3 + (k_s/m_{\text{us}} + k_t/m_{\text{us}} + k_s/m_s)s^2 + (b_s k_t/m_s m_{\text{us}})s + k_s k_t/m_s m_{\text{us}}}. \quad (6.195)$$

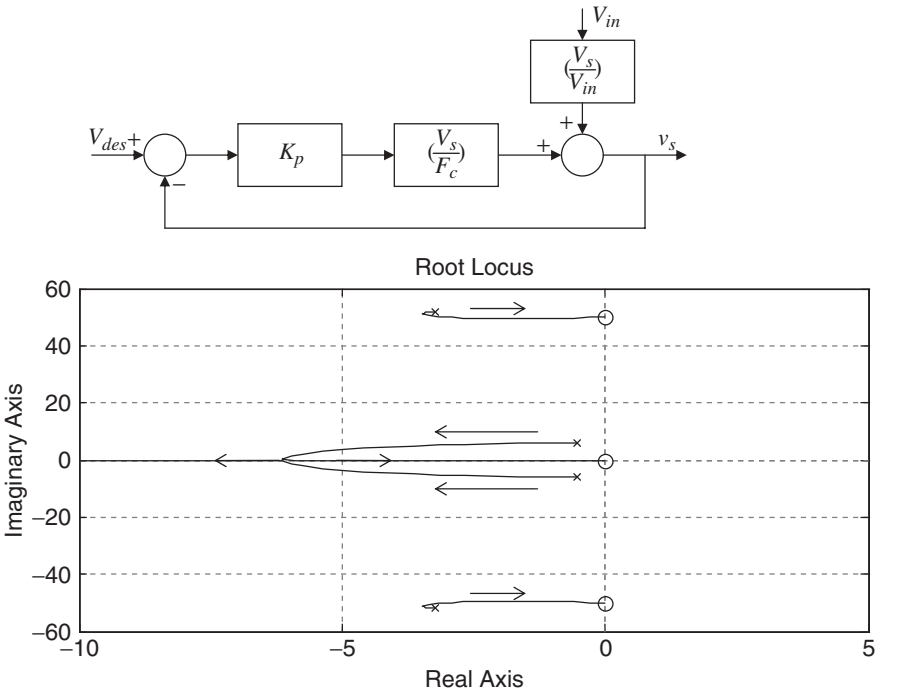


FIGURE 6.48. Block diagram representation and root locus for the quarter-car model with proportional control.

We can see that the passive damping b_s will control the resonant behavior dictated by the denominator but will also affect the high-frequency behavior dictated by the numerator. With proportional control the closed-loop transfer function becomes

$$\frac{V_s}{V_{\text{in}}}|_{\text{CL}} = \frac{\left(\frac{V_s}{V_{\text{in}}}\right)}{1 + K_p\left(\frac{V_s}{F_c}\right)}$$

$$= \frac{(k_t/m_{\text{us}})[(b_s/m_s)s + k_s/m_s]}{s^4 + (b_s/m_s + b_s/m_{\text{us}} + K_p/m_s)s^3 + (k_s/m_{\text{us}} + k_t/m_{\text{us}} + k_s/m_s)s^2 + [(b_s k_t + K_p k_t)/m_s m_{\text{us}}]s + k_s k_t/m_s m_{\text{us}}}. \tag{6.196}$$

We see that the proportional control effects the damping terms of the denominator and thus this type of control may prove useful for ride control of a vehicle suspension.

Lead-lag Control. Before leaving this section on root locus it is appropriate to introduce another control action that can prove useful. It is called lead-lag control and is represented by the transfer function,

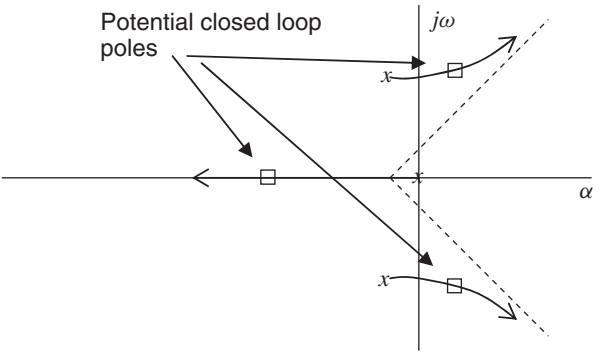
$$G_c(s) = \frac{K(s+z)}{(s+p)}, \tag{6.196a}$$

where the pole p is located in the complex plane much further to the left than the zero z. This control action does not effect the pole/zero excess as one pole and one zero are added to the open-loop transfer function of the system. However, from Eq. (6.192) we see that the centroid of the asymptotes is moved to the left by the action of this controller. Thus, this control action tends to extend the range of system gain K over which the closed-loop system is stable. Consider part (a) of Figure 6.48A, where there are three open loop poles and no zeros. There are three asymptotes located at angles 180° and at $\pm 60^\circ$. The centroid is located as shown. Proportional control will move the closed-loop poles immediately into the unstable right half plane.

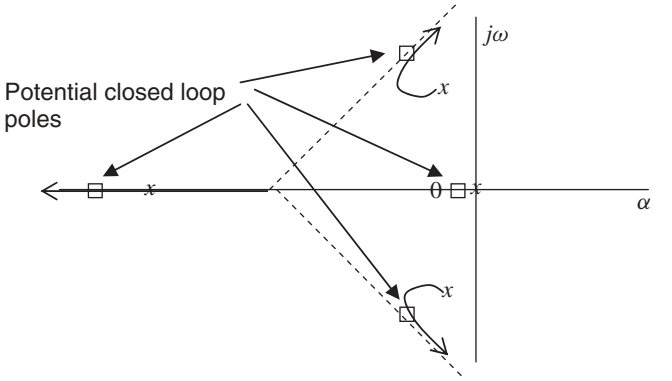
Now consider the addition of the lead-lag controller with added pole and zero located as shown in part (b) of Figure 6.48A. We see that the centroid is moved to the left and there is a range of system gain that yields stable closed loop behavior.

6.6.3 General Control Considerations

In the previous section the concept of automatic control was introduced using position control of a single mass acted upon by a controllable force. The concept of proportional, derivative, and integral control actions was presented. Virtually all control systems will have these actions as part of the overall control system.



(a) Proportional control of a 3-pole open-loop system



(b) Proportional control of a 3-pole open-loop system with added lead-lag control

FIGURE 6.48A. The effect of lead-lag control on the root locus.

Also introduced by example was the concept of a *root locus*, where the path of the closed-loop eigenvalues is shown in the complex s-plane as some controller gain is varied. It is useful to be able to sketch a root locus to determine if a proposed controller has the potential to deliver a desired closed-loop response.

In this section the concept of automatic control is developed more generally and more formally. This development is followed by some realistic physical examples.

Setting Up the Problem. As presented previously, it is proposed that some physical system can be justifiably modeled as a linear system, and a bond graph model of this system yields the state equations in the form

$$\begin{aligned} [\dot{x}] &= [A][x] + [b_d]\, u_d + [b_c]\, u_c, \\ y &= [c]^t [x], \end{aligned} \tag{6.197}$$

where $[x]$ is the vector of state variables composed of momentum and displacement variables (p_s and q_s) from I- and C-elements in integral causality, u_d is a single disturbance input (effort or flow), over which we have no control, and u_c is a single control input (effort or flow), over which we have control. The single output y is some combination of the states and we want to control this output to have some desired response. When we have one controllable input and one controlled output the control system is classified as a SISO (single input–single output) system.

Next, derive the transfer functions relating the output y to the controllable input u_c and the output to the disturbance input u_d; thus derive,

$$G_p(s) = [c]^t \frac{\text{Adj}[sI - A]}{\text{Det}[sI - A]} [b_c] = \frac{Y(s)}{U_c(s)} = \frac{N_p}{D},$$
$$G_d(s) = [c]^t \frac{\text{Adj}[sI - A]}{\text{Det}[sI - A]} [b_d] = \frac{Y(s)}{U_d(s)} = \frac{N_d}{D}. \tag{6.198}$$

The block diagram representation of the control system is shown in Figure 6.49.

The controller is represented by the transfer function $KG_c(s)$ and could be as simple as a proportional gain or a PID controller, a complex transfer function, or even a series of instructions or rules. The basic concept is that when an error exists between the desired output and the actual output, a control action takes place that reduces the error and results in a desirable system response.

The plant transfer function is $G_p(s)$ and the open-loop or *forward*-loop transfer function is $KG_c(s)G_p(s)$. The disturbance transfer function is $G_d(s)$. Block diagram algebra yields the closed-loop transfer function relating Y to Y_{des} as

$$\frac{Y}{Y_{\text{des}}}(s) = \frac{KG_cG_p}{1 + KG_cG_p} \tag{6.199}$$

and the closed-loop transfer function relating the output to the disturbance input as

$$\frac{Y}{U_d}(s) = \frac{G_d}{1 + KG_cG_p}. \tag{6.200}$$

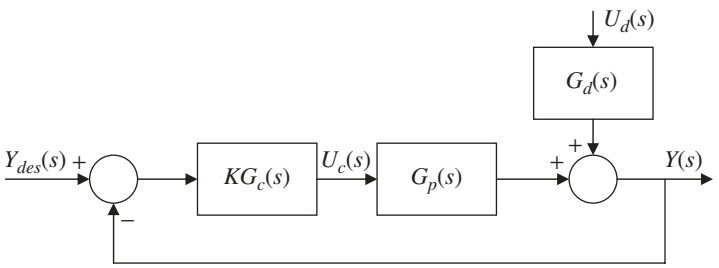


FIGURE 6.49. Starting block diagram for controller development.

Note that if it were possible to have the overall controller gain K become arbitrarily large then Y would approach Y_{des} and the response to the disturbance would approach zero. In general, the overall system gain is limited by power and stability considerations.

From Section 6.6.2 dealing with root locus we know that the main effect of feedback control is to place the closed-loop poles in some desirable locations. We know that the closed-loop poles equal the open-loop poles when there is no control and that they approach the open-loop zeros as the system gain is increased. A goal of the controller design is to have some dominant poles with desired frequency and damping and to have the remaining poles be significantly faster and well damped (further to the left in the complex plane). Normally, the fast poles have little significant influence on the closed-loop response. It is typical to assess the overall system response by considering the time response of the output to a step change in the desired output and by evaluating the closed-loop frequency response.

Consider a plant transfer function,

$$G_p(s) = \frac{Y}{U_c} = \frac{\omega_p^2}{s^2 + 2\zeta_p\omega_p s + \omega_p^2}, \tag{6.201}$$

with $\zeta_p = 0.1$ and $f_p = \dfrac{\omega_p}{2\pi} = 1.0$ Hz. Lets propose the use of a PID controller such that

$$KG_c(s) = K\frac{(s^2 + 2\zeta_0\omega_0 s + \omega_0^2)}{s}, \tag{6.202}$$

where we specify values for ζ_0 and ω_0. A sketch of the root locus is shown in Figure 6.50 and it indicates that we can place the closed-loop poles in desirable locations with the proposed controller.

Figure 6.51 shows the step response for several values of the system gain K with $\zeta_0 = 0.7$ and $f_0 = \omega_0/2\pi = 2.0$ Hz. Notice that we can affect the speed of

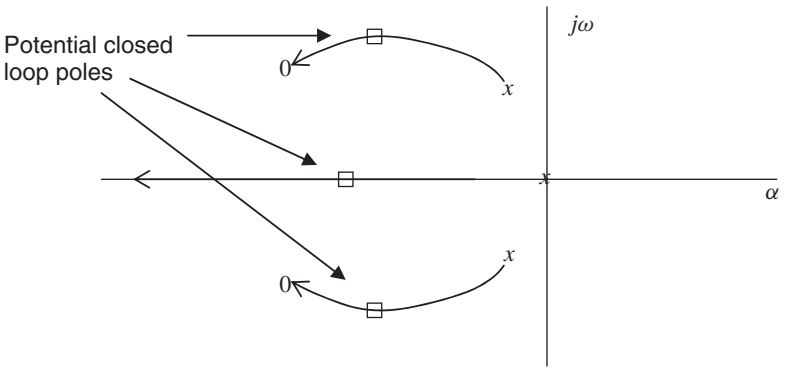


FIGURE 6.50. Approximate root locus for PID control of a second-order plant.

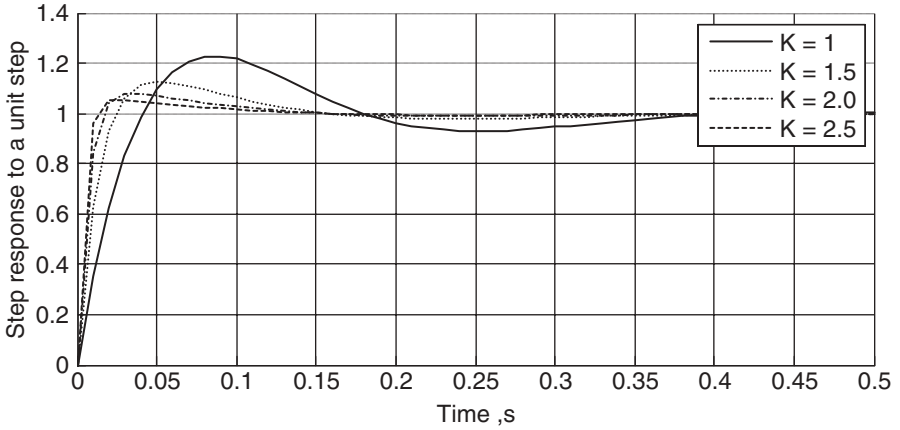


FIGURE 6.51. Step response of the second-order plant with PID control.

response, the overshoot of the final value, and the settling time by adjusting the controller parameters.

Our first concern in designing the control action is that the system be stable over some range of system gain. The next concern is that a desirable closed-loop response results. Considering that we are designing controllers for physical systems, we should keep in mind the power required to accomplish the closed-loop response we specify. Since the system model is a linear representation of a "real" system we should ultimately test the controller with a more realistic system model to determine the range over which our controller will be effective. This test model will often be nonlinear even when we continue to use a linear controller.

Motion Control. A main use of feedback control is to have the output from a system follow some desired output. Another use of such a control system is to keep the system at the desired output when a disturbance acts on the system. Such control systems are called "regulators" and come under the broad category of motion control.

Consider the physical system shown in Figure 6.52. It consists of a dc motor driving a rotary inertia. The dc motor has winding resistance R_w and coupling constant T. The rotary inertia has moment of inertia J. Also shown in the figure is a bond graph of the system with causality and state variable identified.

The single state equation from the bond graph is

$$\dot{p}_J = \tau_d + (T/R_w)(e_c - (T/J)p_J) \tag{6.203}$$

We need the plant transfer function $G_p(s)$ relating rotational angle to the control voltage e_c and we also need the disturbance transfer function $G_d(s)$ relating the rotational angle to the disturbance torque τ_d. These transfer functions are most easily derived by first deriving the transfer functions for the angular

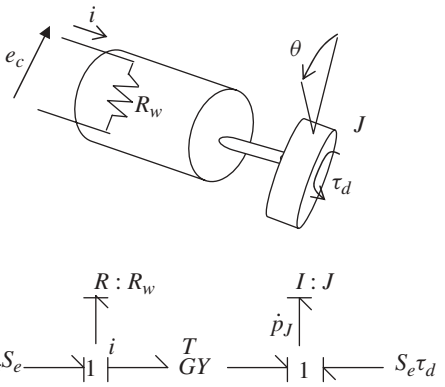


FIGURE 6.52. Schematic and bond graph of an angular position control system.

velocity and then simply *integrating* the angular velocity by dividing by s. The angular velocity ω is the angular momentum p_J divided by moment of inertia J. Thus,

$$\frac{\Omega}{E_c}(s) = \frac{T/R_w J}{s + T^2/R_w J}, \quad \frac{\Omega}{T_d}(s) = \frac{1/J}{s + T^2/R_w J} \tag{6.204}$$

and, therefore,

$$G_p(s) = \frac{\Theta}{E_c}(s) = \frac{T/R_w J}{s(s + T^2/R_w J)}, \quad G_d(s) = \frac{\Theta}{T_d}(s) = \frac{1/J}{s(s + T^2/R_w J)}. \tag{6.205}$$

Figure 6.53 shows the approximate root locus if we propose PD control for this system where the controller $G_c(s)$ is

$$G_c(s) = K(s + z). \tag{6.206}$$

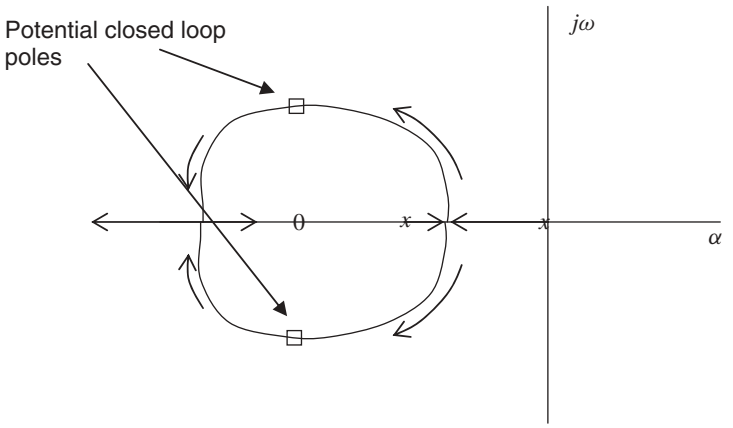


FIGURE 6.53. Approximate root locus of the angular position control system.

For this low-order system we can derive the closed-loop transfer functions using block diagram algebra as indicated by Eqs. (6.199) and (6.200). Thus,

$$\frac{\Theta}{\Theta_{\text{des}}}|_{\text{CL}}(s) = \frac{(KT/R_wJ)(s+z)}{s^2 + [(T^2+KT)/R_wJ]s + (KT/R_wJ)z},$$
$$\frac{\Theta}{T_d}|_{\text{CL}}(s) = \frac{1/J}{s^2 + [(T^2+KT)/R_wJ]s + (KT/R_wJ)z}. \tag{6.207}$$

The closed-loop denominator shows that we are able to adjust the constant term through our choice of Kz and we can adjust the s-term through our choice of K. Thus, the closed-loop eigenvalues are completely adjustable using this PD controller. Furthermore, if we set $s = 0$ we can determine the final value of the response to a step input in θ_{des} or τ_d. The first of Eqs. (6.207) shows that $\theta \to \theta_{\text{des}}$, and the second equation indicates that

$$\frac{\Theta}{T_d}|_{\text{CL}|_{s\to 0}} = \frac{1}{(KT/R_w)z}. \tag{6.208}$$

For θ to have a very small response to the disturbance torque the system gain will have to be large, and this may force a compromise with the desired dynamic response of the system or the power required to accomplish the control. This can be demonstrated computationally using the parameters listed in Table 6.1.

Figure 6.54 shows the step response of the controlled system for two values of the overall system gain. It appears that we can make the system arbitrarily fast by simply increasing the gain.

Figure 6.55 shows the step response of the rotational angle to a disturbance torque. Increasing the system gain indicates that we can reduce the response to a disturbance to an arbitrarily small value.

Figure 6.56 shows the control voltage needed to accomplish the control. We see that there is a price to be paid for increasing the system performance. The voltage required increases, as does the current and the power. Thus, in a real controller

TABLE 6.1. Parameters for the control system of Figure 6.52

Winding resistance, $R_w = 1\ \Omega$
Coupling constant, $T = 0.5$ Nm/A
Mass of rotating disk, $m = 5/2.2$ kg
Radius of rotating disk, $R = 4.0(0.0254)$ m
$J = \frac{mR^2}{2}$ kg-m^2
Location of the controller zero, $z = 20$
Overall system gain, $K = 0.1, 0.2$

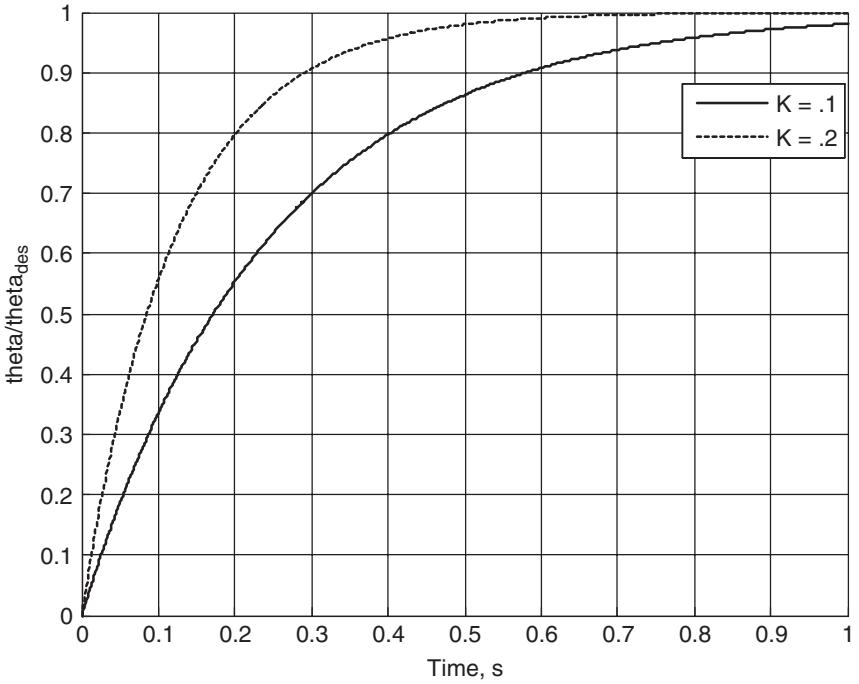


FIGURE 6.54. Step response of the angular position control system for two values of overall system gain.

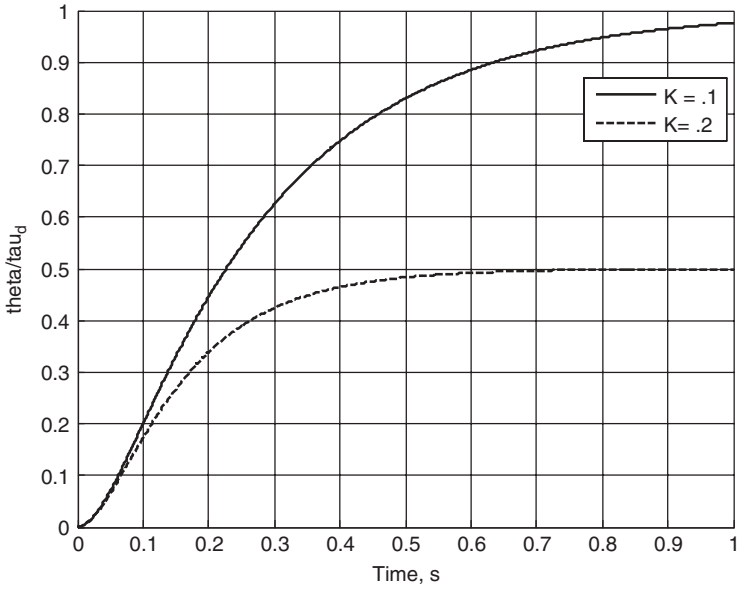


FIGURE 6.55. Step response to a disturbance torque input of the angular position control system for two values of overall system gain.

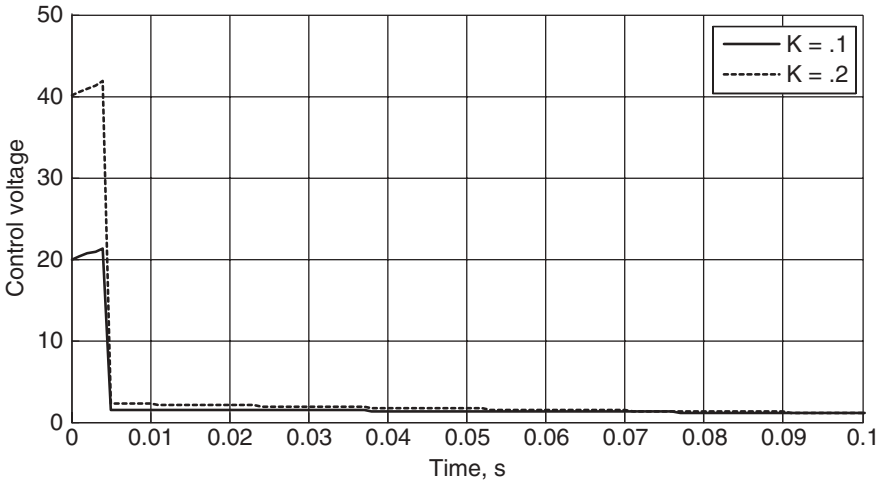


FIGURE 6.56. Control voltage for two values of overall system gain.

design we must consider the trade-offs between desired system performance, actuator capabilities, and power limitations.

Consider now a modified version of the system from Figure 6.52. There is now twisting of the shaft that couples the motor to the rotating inertia and it has stiffness k_τ Nm/rad. The schematic and bond graph are shown in Figure 6.57.

There are two state variables, p_J, the angular momentum of the rotating mass, and q_τ the twist angle in the shaft. Using the derivation procedure from Chapter 5, the state equations are

$$\frac{d}{dt}\begin{bmatrix} p_J \\ q_\tau \end{bmatrix} = \begin{bmatrix} 0 & k_\tau \\ -1/J & -R_w k_\tau / T^2 \end{bmatrix}\begin{bmatrix} p_J \\ q_\tau \end{bmatrix} + \begin{bmatrix} 0 \\ 1/T \end{bmatrix} e_c + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tau_d. \qquad (6.209)$$

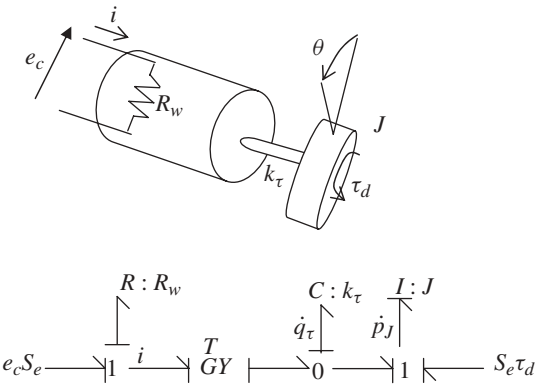


FIGURE 6.57. Modified physical system from Figure 6.52.

And in the s-domain these become

$$\begin{bmatrix} s & -k_\tau \\ 1/J & s + R_w k_\tau / T^2 \end{bmatrix} \begin{bmatrix} P_J(s) \\ Q_\tau(s) \end{bmatrix} + \begin{bmatrix} 0 \\ 1/T \end{bmatrix} E_c(s) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} T_d(s). \tag{6.210}$$

The transfer functions relating the angular velocity of the rotating mass to the control voltage and the disturbance torque are straightforwardly derived as

$$\frac{\Omega}{E_c}(s) = \frac{k_\tau / JT}{s^2 + (R_w k_\tau / T^2)s + k_\tau / J}, \quad \frac{\Omega}{T_d}(s) = \frac{(1/J)(s + R_w k_\tau / T^2)}{s^2 + (R_w k_\tau / T^2)s + k_\tau / J}, \tag{6.211}$$

and the transfer functions for the angular displacement are derived by simply dividing by s (integrating ω), yielding the plant transfer function and the disturbance transfer function as

$$\begin{aligned} G_p(s) &= \frac{\Theta}{E_c}(s) = \frac{k_\tau / JT}{s[s^2 + (R_w k_\tau / T^2)s + k_\tau / J]}, \\ G_d(s) &= \frac{\Theta}{T_d}(s) = \frac{(1/J)(s + R_w k_\tau / T^2)}{s[s^2 + (R_w k_\tau / T^2)s + k_\tau / J]}. \end{aligned} \tag{6.212}$$

The open-loop poles of the plant are sketched in Figure 6.58 along with the additional poles and zeros associated with proposed control actions. It should be clear that proportional control will yield an unstable system. With three open-loop poles and no open-loop zeros, there are three asymptotes with angles of $\pm 60^\circ$ and 180°. The intersections of the asymptotes with the real axis is going to be very near the cluster of open-loop poles, as indicated by Eq. (6.192).

It is proposed to combine PD control and lead-lag control in order to have stable closed-loop poles for a range of system gain. The PD control allows placement of an additional zero, which reduces the pole zero excess from 3 to 2 and makes the angle of the asymptotes $\pm 90^\circ$. But the intersection with the real axis is moved further to the right as indicated by Eq. (6.192). By adding the lead-lag control action, the asymptotes are moved to the left, yielding a range of system gain where the closed-loop poles are in an acceptable location. Thus, the control action will be

$$G_c(s) = K \frac{s + z}{s + p}(s + z_{\text{pd}}), \tag{6.213}$$

where z and p are the zero and pole of the lead-lag controller and z_{pd} is the zero associated with the PD control action.

The system parameters are the same as given in Table 6.1 with the addition of the torsional stiffness. Let's assume that the torsional spring is fairly soft and the combined rotary inertia and spring yield a frequency of 1 Hz, such that $\omega_n^2 = (2\pi 1)^2$ and $k_\tau = J\omega_n^2$.

The program Matlab [5] was used to evaluate the closed-loop step response of the system as shown in Figure 6.59. The controller poles and zeros were selected through some numerical experimentation. As can be seen in the figure, there is

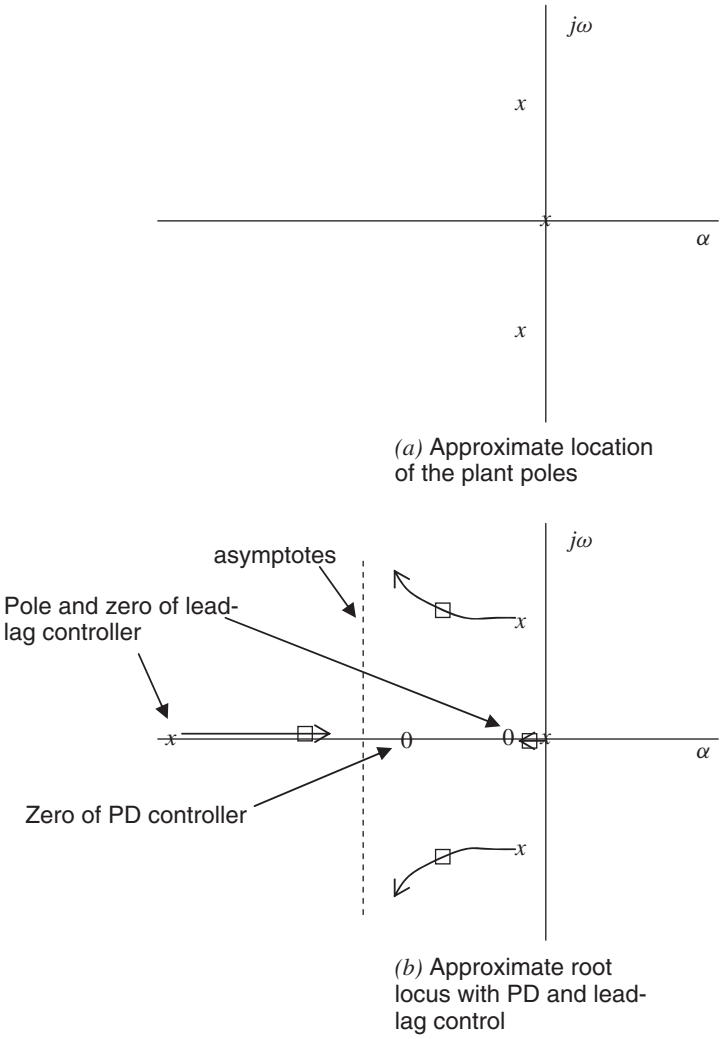


FIGURE 6.58. Approximate root locus for the system of Figure 6.57 using PD and lead-lag control.

some overshoot at the beginning of the response followed by a settling to the final value of $\frac{\theta}{\theta_{\text{des}}} \rightarrow 1$. It would appear to be possible to obtain a faster response using different controller parameters and a different system gain, but the physical system itself indicates that this is really not possible. We are trying to control the position of a rotary inertia by twisting the end of a relatively soft spring. A fast response of such a system requires much twisting of the shaft, as shown in Figure 6.60. This example shows that it is important to couple one's physical system understanding with the principles of automatic control to obtain a realistic overall system design.

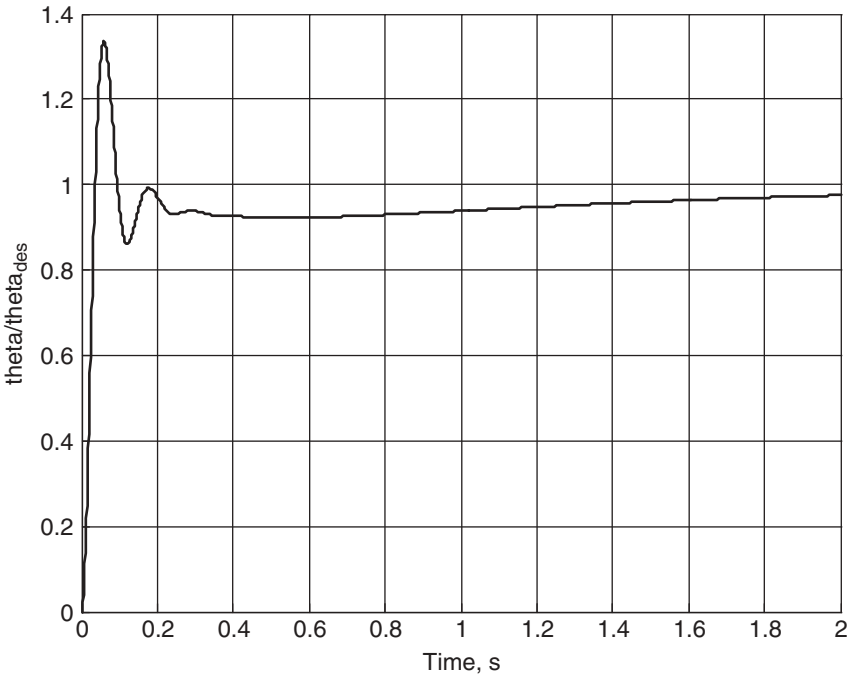


FIGURE 6.59. Step response of the angular position control system.

Motion Control Using State Variable Feedback. In classical control we measure one output and compare the measurement to some desired output and generate an error signal upon which the controller compensator operates to generate actuator commands. The system is composed of perhaps many states and if all the states can be continuously measured, then certainly they contain more information than just a single output measurement. Consider our typical starting point for linear system analysis,

$$[\dot{x}] = [A][x] + [b_d]\,u_d + [b_c]\,u_c,$$
$$y = [c]^t[x], \tag{6.214}$$

where there is one control input u_c, one disturbance input u_d, and one output y.

Let's propose that the control input depends on the state variables such that

$$u_c = K(y_{\text{des}} - [k]^t[x]), \tag{6.215}$$

where K is an overall system gain and $[k]^t$ is a row vector of gains multiplying the state variables. Using this control action in Eqs. (6.214) yields

$$[\dot{x}] = [A - [b_c]\,K[k]^t][x] + [b_c]\,K\,y_{\text{des}} + [b_d]\,u_d,$$
$$y = [c]^t[x]. \tag{6.216}$$

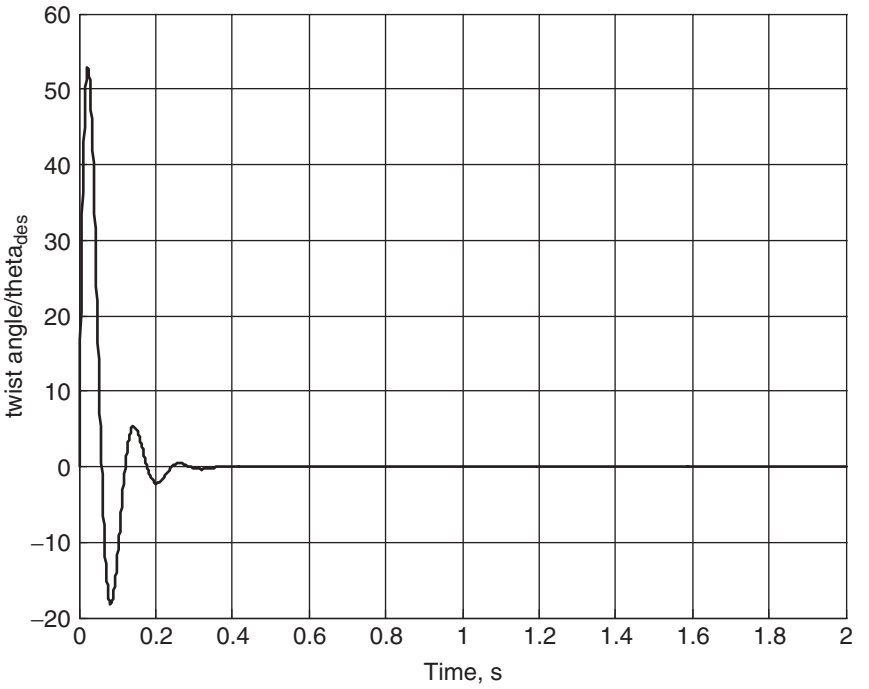


FIGURE 6.60. Step response of the shaft twist angle corresponding to the position response of Figure 6.59.

From Eq. (6.216) we see that state variable feedback modifies the original $[A]$-matrix into the closed-loop $[A]$-matrix,

$$[A]^* = [A - [b_c]\, K[k]^t], \tag{6.217}$$

such that the closed-loop eigenvalues come from the closed-loop characteristic equation

$$\text{Det}[sI - A^*] = 0. \tag{6.218}$$

It is straightforward to show that state variable feedback allows arbitrary placement of closed-loop poles anywhere in the s-plane (see References [6] and [7]). It is demonstrated here using the angular displacement control system from Figure 6.57.

The state equations were derived previously as Eqs. (6.209) and are repeated here with the substitution of state variable feedback control,

$$e_c = K(\theta_{\text{des}} - [k]^t[x]) = K\left(\theta_{\text{des}} - [k_{p_J} \quad k_{q_\tau} \quad k_\theta]\begin{bmatrix} p_J \\ q_\tau \\ \theta \end{bmatrix}\right). \tag{6.219}$$

One modification to state equations (6.209) is the inclusion of the rotational angle of the rotary inertia as an additional state variable.

Using (6.219) in the state equations results in

$$\frac{d}{dt}\begin{bmatrix} p_J \\ q_\tau \\ \theta \end{bmatrix} = \begin{bmatrix} 0 & k_\tau & 0 \\ -1/J & -R_w k_\tau/T^2 & 0 \\ 1/J & 0 & 0 \end{bmatrix}\begin{bmatrix} p_J \\ q_\tau \\ \theta \end{bmatrix} - \begin{bmatrix} 0 \\ 1/T \\ 0 \end{bmatrix} K \begin{bmatrix} k_{p_J} & k_{q_\tau} & k_\theta \end{bmatrix}\begin{bmatrix} p_J \\ q_\tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 1/T \\ 0 \end{bmatrix} K\theta_{\text{des}} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\tau_d \tag{6.220}$$

or

$$\frac{d}{dt}\begin{bmatrix} p_J \\ q_\tau \\ \theta \end{bmatrix} = \begin{bmatrix} 0 & k_\tau & 0 \\ -1/J - Kk_{p_J}/T & -R_w k_\tau/T^2 - Kk_{q_\tau}/T & -Kk_\theta/T \\ 1/J & 0 & 0 \end{bmatrix}\begin{bmatrix} p_J \\ q_\tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 1/T \\ 0 \end{bmatrix} K\theta_{\text{des}} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\tau_d, \tag{6.221}$$

yielding

$$\mathbf{A}^* = \begin{bmatrix} 0 & k_\tau & 0 \\ -1/J - Kk_{p_J}/T & -R_w k_\tau/T^2 - Kk_{q_\tau}/T & -Kk_\theta/T \\ 1/J & 0 & 0 \end{bmatrix}. \tag{6.222}$$

The closed-loop characteristic equation is

$$\text{Det}[sI - A^*] = \text{Det}\begin{bmatrix} s & -k_\tau & 0 \\ 1/J + Kk_{p_J}/T & s + R_w k_\tau/T^2 + Kk_{q_\tau}/T & Kk_\theta/T \\ -1/J & 0 & s \end{bmatrix} = 0$$

or

$$s^3 + (R_w k_\tau/T^2 + Kk_{q_\tau}/T)s^2 + k_\tau(1/J + Kk_{p_J}/T)s + k_\tau Kk_\theta/JT = 0 \tag{6.223}$$

Since the adjustable gains $k_{p_J}, k_{q_\tau}, k_\theta$ appear in the coefficients of s^2, s^1, s^0, these coefficients are totally adjustable and the three-closed loop eigenvalues can be placed arbitrarily.

It is instructive to complete this derivation by looking at the closed-loop transfer functions relating θ to θ_{des} and θ to τ_d. This comes from using Cramer's rule on the s-domain representation of Eqs. (6.221). The result of the algebra is

$$\frac{\Theta}{\Theta_{\text{des}}} = \frac{Kk_\tau/TJ}{s^3 + (R_w k_\tau/T^2 + Kk_{q_\tau}/T)s^2 + k_\tau(1/J + Kk_{p_J}/T)s + k_\tau Kk_\theta/JT},$$
$$\frac{\Theta}{T_d} = \frac{(1/J)(s + R_w k_\tau/T^2 + Kk_{q_\tau}/T)}{s^3 + (R_w k_\tau/T^2 + Kk_{q_\tau}/T)s^2 + k_\tau(1/J + Kk_{p_J}/T)s + k_\tau Kk_\theta/JT}. \tag{6.224}$$

One observation is that state variable feedback does not insure the absence of steady-state error. In the first of Eq. (6.224) the final value of the step response to a unit change in θ_{des} is obtained from the transfer function by letting $s \to 0$. This results in

$$\frac{\theta}{\theta_{\text{des}}}|_{s\to 0} \to \frac{1}{k_\theta}, \tag{6.225}$$

which is not unity unless $k_\theta = 1$. While not proven here, it is also a true that state variable feedback does not change the location of any open-loop zeros for the closed-loop response to the reference input. The first of Eqs. (6.212) is the open-loop plant transfer function. Notice that the numerator of this transfer function is a constant. Using state variable feedback, the closed-loop response to the reference is the first of Eqs. (6.224). The numerator is again a constant, unaffected by the feedback gains. It is provable that the open-loop and closed-loop zeros are the same when using state variable feedback. This observation is not true for the disturbance response. As seen from the second of Eqs. (6.224), the feedback gains do affect the numerator of the disturbance transfer function.

Figure 6.61 shows the step response of the angular position control system having chosen the state variable gains to yield an acceptable response with the gain $k_\theta = 1$, such that there is zero steady-state error.

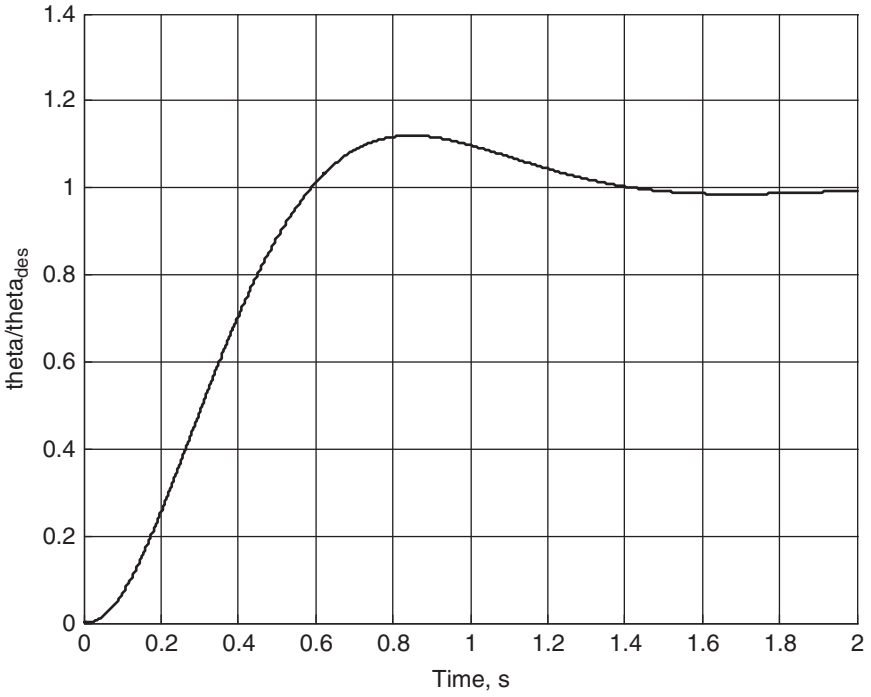


FIGURE 6.61. Step response of the angular control system to a unit step change in desired angular position.

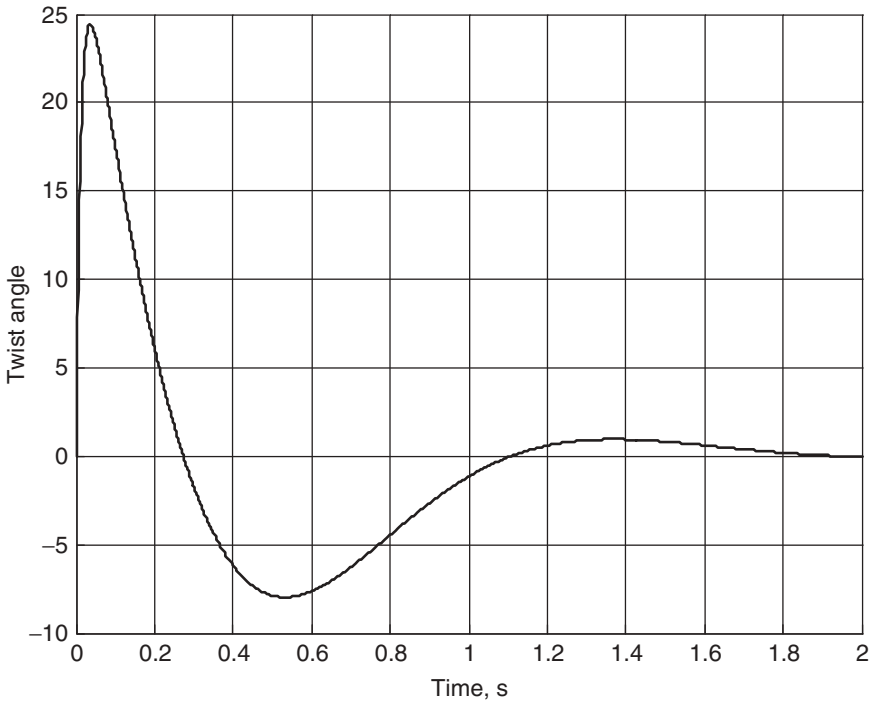


FIGURE 6.62. Step response of the shaft twist angle to a unit step change in desired angular position.

Figure 6.62 shows the twist angle in the shaft and we see that we are able to accomplish the position control with a more reasonable twist angle than was needed for the compensated system of Figure 6.60.

This example indicates that state variable feedback is a powerful tool for designing control systems. Since closed-loop poles can be arbitrarily specified, system stability can be guaranteed. The practical problem in using state variable feedback for complex multistate systems is that each state must be measured in order to be fed back as part of the control system. As the number of sensors grows so does cost of implementation. Keep in mind that state variables are "man-made" and are part of a model that is a representation of a physical system. It may not be possible to measure a system state, as that state may be just a figment of the imagination of the modeler. Furthermore, it is not practical to expect multiple sensors to remain calibrated and operational over the long term. Nevertheless, state variable feedback is often part of an overall control system design, as it is indicative of the level of control that is theoretically attainable and provides an excellent target for more practical control approaches.

The interested reader should see References [3] and [6]–[8] to learn more about state variable feedback and the possibility of estimating some states in order to reduce the number of actual measurements. The computer program Matlab [5] contains many procedures to assist in the design of state variable controllers.

Vibration Control. In the previous section on motion control the idea was to use controlled actuators to force a system to follow a desired trajectory while simultaneously responding as little as possible to a disturbance input. Another use of automatic control is to force a structure to behave in a prescribed way to external disturbances. This use of automatic control comes under the heading of vibration control. When a structure requires isolation from its environment it is typical to first use passive means to provide the isolation. Commercial "isolators" are available in all shapes and sizes to provide such passive isolation. These devices all have compliance and dissipative properties in the form of springs, rubber pads, shock absorbers, air springs, and so forth. When isolation beyond that possible with passive elements is required, then automatic control becomes a candidate solution.

The approach to developing the control is somewhat different than that for motion control and is demonstrated using the example from Figure 6.63. This example system was used earlier in Section 6.5.1 to demonstrate frequency response. It is repeated here for convenience. The system consists of a mass to be isolated from the ground motion with a proposed controlled force element located between the mass and ground input. Also shown in the figure is the "desired" system as well as system bond graphs.

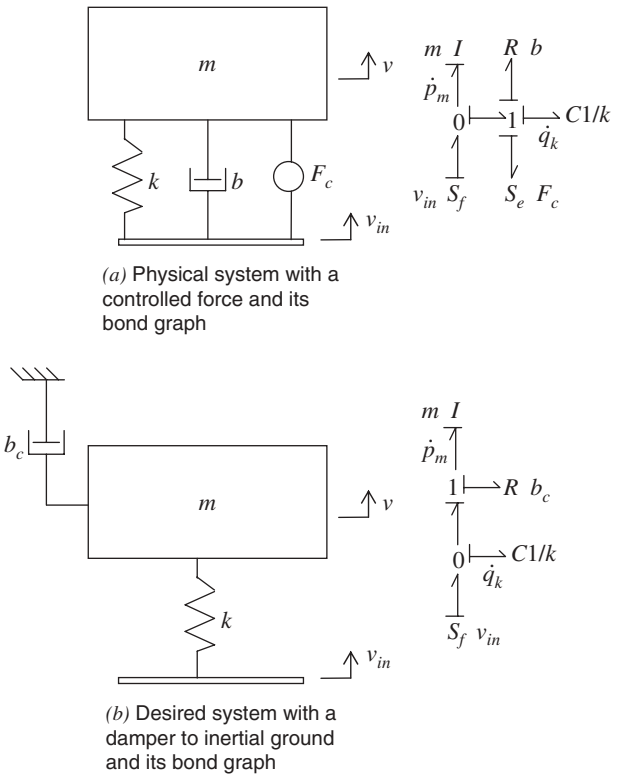


FIGURE 6.63. Vibration isolation system.

In vibration control it is typical to specify the desired performance in the frequency domain. We are not so much interested in having the response to a disturbance approach zero but rather we would like the system to respond in a prescribed way over some range of frequencies.

The state equations for the controlled system come directly from the bond graph as

$$\begin{aligned} \dot{p}_m &= bv_{\text{in}} - (b/m)\frac{p_m}{m} + kq_k + F_c, \\ \dot{q}_k &= v_{\text{in}} - (1/m)p_m, \end{aligned} \tag{6.226}$$

where p_m is the momentum of the mass and q_k is the displacement across the spring (with the spring, damper, and control force all positive in compression). In the s-domain these equations become

$$\begin{bmatrix} s + b/m & -k \\ 1/m & s \end{bmatrix} \begin{bmatrix} P_m(s) \\ Q_k(s) \end{bmatrix} = \begin{bmatrix} b \\ 1 \end{bmatrix} V_{\text{in}}(s) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} F_c(s), \tag{6.227}$$

yielding the transfer functions

$$\begin{aligned} \frac{V}{V_{\text{in}}}(s) &= \frac{(b/m)s + k/m}{s^2 + (b/m)s + k/m} = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \\ \frac{V}{F_c}(s) &= \frac{(1/m)s}{s^2 + (b/m)s + k/m} = \frac{(1/m)s}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \end{aligned} \tag{6.228}$$

where the damping ratio ζ and natural frequency ω_n have been defined previously.

For the desired system shown in Figure 6.63*b* the state equations are

$$\begin{aligned} \dot{p}_m &= -(b_c/m)p_m + kq_k \\ \dot{q}_k &= v_{\text{in}} - (1/m)p_m, \end{aligned} \tag{6.229}$$

with an s-domain representation of

$$\begin{bmatrix} s + b_c/m & -k \\ 1/m & s \end{bmatrix} \begin{bmatrix} P_m(s) \\ Q_k(s) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} V_{\text{in}}(s), \tag{6.230}$$

yielding the transfer functions

$$\begin{aligned} \left(\frac{V}{V_{\text{in}}}\right)_{\text{des}} &= \frac{k/m}{s^2 + (b_c/m)s + k/m} = \frac{\omega_n^2}{s^2 + 2\zeta_c\omega_n s + \omega_n^2}, \\ \left(\frac{V_{\text{rel}}}{V_{\text{in}}}\right)_{\text{des}} &= \frac{s(s + b_c/m)}{s^2 + (b_c/m)s + k/m} = \frac{s(s + 2\zeta_c\omega_n)}{s^2 + 2\zeta_c\omega_n s + \omega_n^2}, \end{aligned} \tag{6.231}$$

where ζ_c is the desired damping ratio.

Note that the transfer function for $V_{\text{rel}}(s)$ comes from first deriving the transfer function for $Q_k(s)$, the relative displacement across the suspension, and then

multiplying by s to take the derivative and yield the relative velocity, $V_{\text{rel}}(s) = sQ_k(s)$. We make use of this transfer function below.

The reason that the system in Figure 6.63*b* is the desired system becomes apparent from the frequency responses shown in Figure 6.64. Remember, the frequency response is obtained from the transfer functions by letting $s = j\omega$ and carrying out the steps shown in Section 6.5. The program Matlab [5] can numerically perform these steps and produce the responses shown in Figure 6.64.

In Figure 6.63*a*, the damper is located in the conventional location between the ground excitation and the structure mass. As a result, the frequency response from Figure 6.64*a* shows increasing high-frequency response as the damping is increased to reduce the resonant response. The system in Figure 6.63*b* has the damper attached to the mass at one end and to inertial ground at the other. This inertial damping produces the frequency response of Figure 6.64*b*. In this case we can control the resonant behavior and retain excellent high-frequency isolation. For most applications it is not possible to attach one end of a damper to inertial space; thus, this system is "desired" but not attainable using passive devices.

The challenge is to use the control force to make the actual physical system perform like the desired system over some frequency range. Because the systems are linear, superposition dictates that

$$V(s) = \left(\frac{V}{V_{\text{in}}}\right) V_{\text{in}}(s) + \left(\frac{V}{F_c}\right) F_c(s), \tag{6.232}$$

where the terms in parentheses are the transfer functions from Eqs. (6.228). We want the response to be identical to the first of Eqs. (6.231); thus,

$$\left(\frac{V}{V_{\text{in}}}\right) V_{\text{in}}(s) + \left(\frac{V}{F_c}\right) F_c(s) = \left(\frac{V}{V_{\text{in}}}\right)_{\text{des}} V_{\text{in}}(s). \tag{6.233}$$

Solving for the control yields

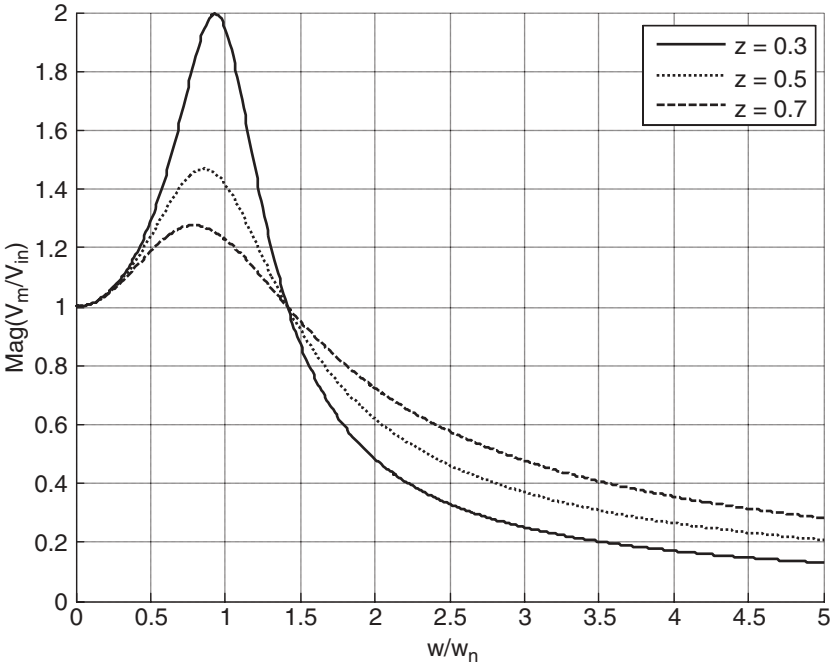
$$F_c(s) = \frac{\left(\left(\frac{V}{V_{\text{in}}}\right)_{\text{des}} - \left(\frac{V}{V_{\text{in}}}\right)\right)}{\left(\frac{V}{F_c}\right)} V_{\text{in}}(s). \tag{6.234}$$

Identify the numerator and denominator of the transfer functions from Eqs. (6.228) and (6.231) as

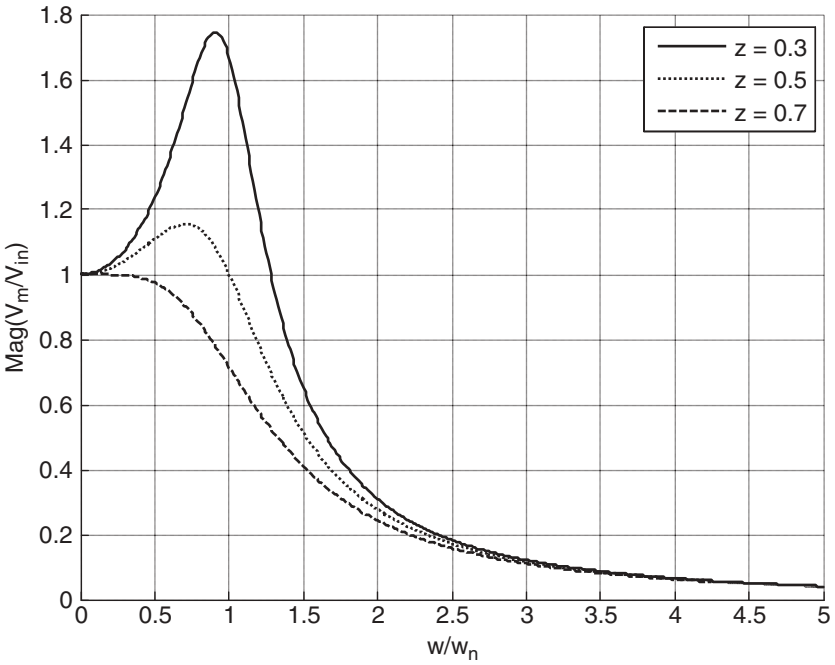
$$\text{Numerator of } \left(\frac{V}{V_{\text{in}}}\right) = N(s) = 2\zeta\omega_n s + \omega_n^2 \text{ from Eq. (6.228)}$$

$$\text{Numerator of } \left(\frac{V}{F_c}\right) = \frac{1}{m}s \text{ from Eq. (6.228)}$$

$$\text{Denominator of } \left(\frac{V}{V_{\text{in}}}\right) = D(s) = s^2 + 2\zeta\omega_n s + \omega_n^2 \text{ from Eq. (6.228)}$$



(a) Frequency response of the passive system



(b) Frequency response of the desired system

FIGURE 6.64. Frequency response of the passive and desired systems.

$$\text{Numerator of } \left(\frac{V}{V_{\text{in}}}\right)_{\text{des}} = \omega_n^2 \text{ from Eq. (6.231)}$$

$$\text{Denominator of } \left(\frac{V}{V_{\text{in}}}\right)_{\text{des}} = s^2 + 2\zeta_c\omega_n s + \omega_n^2 \text{ from Eq. (6.231)}$$

Use these definitions in Eq. (6.234) and obtain

$$F_c(s) = -\frac{2\zeta_c\omega_n^3 s}{(1/m)s(s^2 + 2\zeta_c\omega_n s + \omega_n^2)} V_{\text{in}}(s) - \frac{2\zeta\omega_n s^2(s + 2\zeta_c\omega_n)}{(1/m)s(s^2 + 2\zeta_c\omega_n s + \omega_n^2)} V_{\text{in}}(s) \tag{6.235}$$

The interpretation of this result is that if the input disturbance could be measured and then processed through the filter of Eq. (6.235), then the output force would cause the controlled system to behave identically to the desired system. It is not practical to measure the disturbance and there are simplifications to be carried out.

The first simplification is to recognize that if the control force of (6.235) were applied then the resulting response is the desired one of Eq. (6.231). Thus, solving for $V_{\text{in}}(s)$, $V_{\text{in}} = \dfrac{s^2 + 2\zeta_c\omega_n s + \omega_n^2}{\omega_n^2} V(s)$, and substituting into the first term of Eq. (6.235) yields

$$F_c(s) = -2\zeta_c\omega_n m V(s) - \frac{2\zeta\omega_n s^2(s + 2\zeta_c\omega_n)}{(1/m)s(s^2 + 2\zeta_c\omega_n s + \omega_n^2)} V_{\text{in}}(s). \tag{6.236}$$

The first term on the right-hand side indicates that one part of the control force simply opposes the absolute velocity of the structure mass. From the second of Eqs. (6.231), solving for $V_{\text{in}}(s)$ yields

$$V_{\text{in}}(s) = \frac{s^2 + 2\zeta_c\omega_n s + \omega_n^2}{s(s + 2\zeta_c\omega_n)} V_{\text{rel}}(s),$$

such that Eq. (6.236) reduces to

$$F_c(s) = -2\zeta_c\omega_n m V(s) - 2\zeta\omega_n m V_{\text{rel}}(s) \tag{6.237}$$

The end result is one that we could have observed from the physical systems directly. The second term in (6.237) is the force produced by the passive damper and the control action is to simply cancel this force and substitute a force (the first term in (6.237)) that would be generated by a damper attached to inertial space.

The idea of using active control to generate a force on a structure like a damper to inertial ground turns out to be a very good strategy in vibration control. Commercial devices for ground vehicle suspensions use this strategy as a major part of their control action.

When it is not possible to locate a controlled actuator between two moving parts of a structure, then the only way to generate a controlled force on the structure is to use an inertial actuator. A major application area for inertial

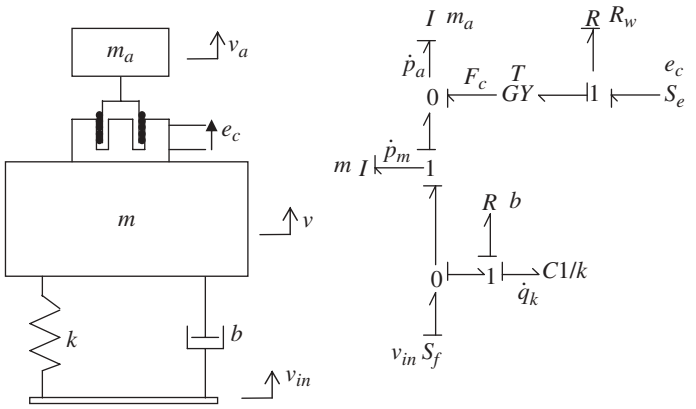


FIGURE 6.65. Vibrating structure with an inertial actuator for control.

actuators is for vibration control on helicopters and fixed-wing aircraft. Consider the system of Figure 6.63 with the voice coil actuator attached to the structure mass as shown in Figure 6.65.

A voice coil actuator is attached to the structure mass and, through electromagnetic coupling, can accelerate the proof-mass m_a, thus creating a reaction force on the structure. The bond graph of the physical system is also shown in Figure 6.65 and state variables are identified on the bond graph along with causality. The state equations come directly from the bond graph as

$$\begin{aligned}
\dot{p}_m &= kq_k + bv_{\text{in}} - (b/m)p_m - T/R_w\,(e_c - T(p_a/m_a - p_m/m))\,,\\
\dot{p}_a &= (T/R_w)\,(e_c - T(p_a/m_a - p_m/m))\,,\\
\dot{q}_k &= v_{\text{in}} - p_m/m,
\end{aligned} \tag{6.238}$$

or in matrix form

$$\frac{d}{dt}\begin{bmatrix} p_m \\ p_a \\ q_k \end{bmatrix} = \begin{bmatrix} -b/m - T^2/R_w m & T^2/R_w m_a & k \\ T^2/R_w m & -T^2/R_w m_a & 0 \\ -1/m & 0 & 0 \end{bmatrix}\begin{bmatrix} p_m \\ p_a \\ q_k \end{bmatrix} + \begin{bmatrix} -T/R_w \\ T/R_w \\ 0 \end{bmatrix} e_c + \begin{bmatrix} b \\ 0 \\ 1 \end{bmatrix} v_{\text{in}}. \tag{6.238a}$$

The actuator force F_c is labeled on the bond graph and, by following causality, can be related to the state variables by

$$F_c = (T/R_w)\,(e_c - T(p_a/m_a - p_m/m))\,. \tag{6.239}$$

From the previous example it has been established that a good strategy in vibration control is to have the structure behave as though it has inertial damping

b_c. For the sign convention shown in Figure 6.65, this requires the control force to be

$$F_c = (b_c/m)\,p_m = (T/R_w)\,(e_c - T(p_a/m_a - p_m/m))\,, \tag{6.240}$$

which allows the control voltage to be derived as

$$(T/R_w)e_c = (b_c/m)\,p_m + (T^2/R_w)(p_a/m_a - p_m/m). \tag{6.241}$$

There are two ways to look at this control action. In terms of the velocities of the structure mass and proof-mass we obtain,

$$(T/R_w)e_c = b_c v + (T^2/R_w)v_{\text{rela}} \tag{6.242}$$

where v_{rela} is the relative velocity between the proof-mass and the structure mass, positive in extension. By measuring the absolute velocity of the structure mass and the relative velocity across the actuator, the control voltage can be established such that inertial damping results.

Equation (6.241) allows the control voltage to be set using state variable feedback as described above. The control voltage can be written as

$$(T/R_w)e_c = \begin{bmatrix} b_c/m - T^2/R_w m & T^2/R_w m_a & 0 \end{bmatrix} \begin{bmatrix} p_m \\ p_a \\ q_k \end{bmatrix}, \tag{6.243}$$

which modifies the state equations to

$$\frac{d}{dt}\begin{bmatrix} p_m \\ p_a \\ q_k \end{bmatrix} = \begin{bmatrix} -b/m - b_c/m & 0 & k \\ b_c/m & 0 & 0 \\ -1/m & 0 & 0 \end{bmatrix} \begin{bmatrix} p_m \\ p_a \\ q_k \end{bmatrix} + \begin{bmatrix} b \\ 0 \\ 1 \end{bmatrix} v_{\text{in}}. \tag{6.244}$$

In the s-domain the controlled system representation becomes

$$\begin{bmatrix} s + b/m + b_c/m & 0 & -k \\ -b_c/m & s & 0 \\ 1/m & 0 & s \end{bmatrix} \begin{bmatrix} P_m(s) \\ P_a(s) \\ Q_k(s) \end{bmatrix} = \begin{bmatrix} b \\ 0 \\ 1 \end{bmatrix} v_{\text{in}}. \tag{6.245}$$

The transfer function relating the structure mass velocity to the input velocity comes from using Cramer's rule, with the result

$$\frac{V}{V_{\text{in}}}(s) = \frac{s((b/m)s + k/m)}{s(s^2 + (b/m + b_c/m)s + k/m)}. \tag{6.246}$$

The appearance of the control parameter b_c in the denominator indicates that the frequency response can be reduced at the resonant frequency while also reducing the response at high frequency. The control voltage needed to accomplish this control comes from Eq. (6.243) as

$$\frac{T}{R_w}\frac{E_c}{V_{\text{in}}}(s) = \left(\frac{b_c}{m} - \frac{T^2}{R_w m}\right)\frac{P_m}{V_{\text{in}}}(s) + \frac{T^2}{R_w m_a}\frac{P_a}{V_{\text{in}}}(s). \tag{6.247}$$

Again using Cramer's rule, the appropriate transfer functions can be derived and the result is

$$\frac{T}{R_w}\frac{E_c}{V_{\text{in}}}(s) = \frac{\left(\frac{b}{m}s + \frac{k}{m}\right)\left((b_c - T^2/R_w)s + b_c T^2/R_w m_a\right)}{s(s^2 + (b/m + b_c/m)s + k/m)}. \tag{6.248}$$

Using the definitions, $\frac{b}{m} = 2\zeta\omega_n$, $\frac{b_c}{m} = 2\zeta_c\omega_n$, and $\omega_n^2 = \frac{k}{m}$, Eqs. (6.246) and (6.248) can be placed in the nondimensional format,

$$\frac{V}{V_{\text{in}}}(s) = \frac{2\zeta\frac{s}{\omega_n} + 1}{\frac{s^2}{\omega_n^2} + (2\zeta + 2\zeta_c)\frac{s}{\omega_n} + 1},$$

$$\frac{T}{mR_w\omega_n}\frac{E_c}{V_{\text{in}}}(s) = \frac{\left(2\zeta\frac{s}{\omega_n} + 1\right)\left(\left(2\zeta_c - \frac{T^2}{mR_w\omega_n}\right)\frac{s}{\omega_n} + 2\zeta_c\frac{T^2}{mR_w\omega_n m_a/m}\right)}{\frac{s}{\omega_n}\left(\frac{s^2}{\omega_n^2} + (2\zeta + 2\zeta_c)\frac{s}{\omega_n} + 1\right)}. \tag{6.249}$$

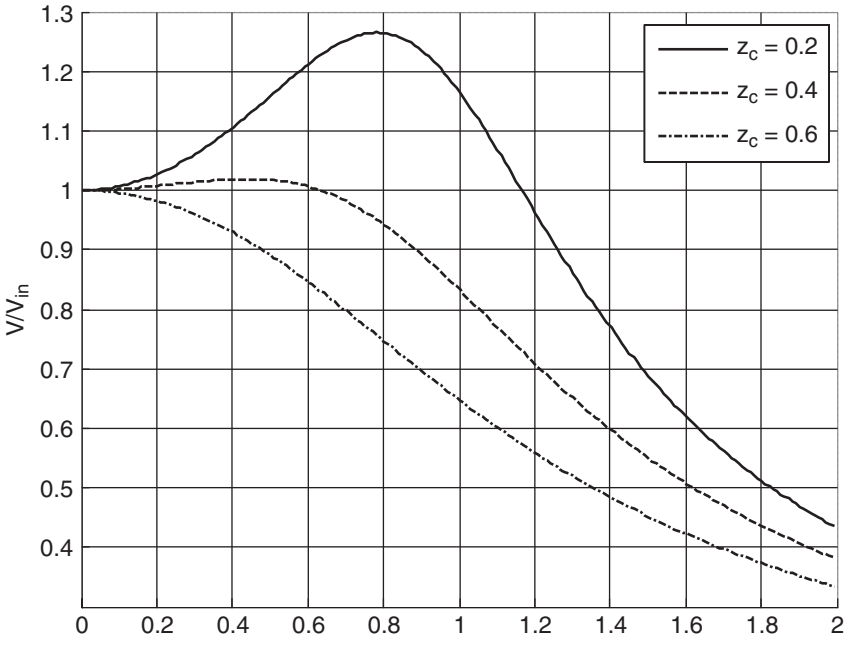
Table 6.2 lists some parameters for the physical system of Figure 6.65 and the program Matlab [5] was used to generate some system responses.

Figure 6.66*a* shows that the inertial actuator is quite capable of generating a force on the structure that emulates an inertial damper. Figure 6.66*b* shows the nondimensional control voltage required to produce the respective level of isolation.

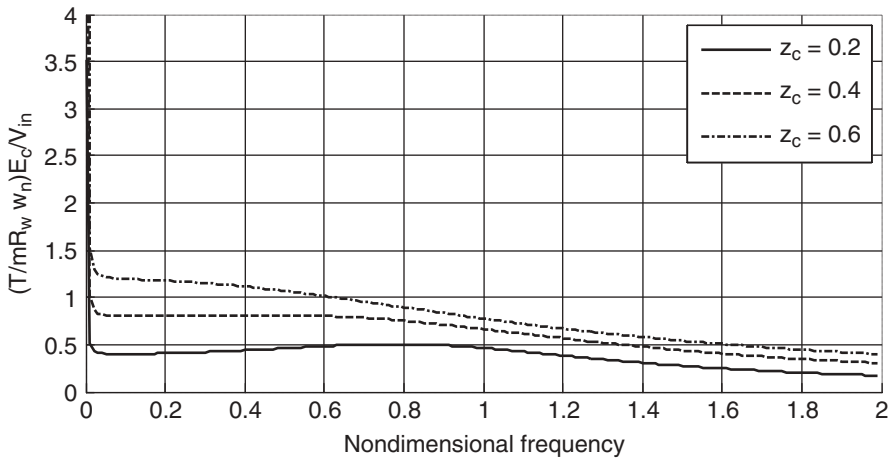
If this was an analysis performed for system design purposes, we would next perform extensive parameter studies by assessing the effect of the electromagnetic coupling constant, winding resistance, size of the proof mass, and so forth. Such an analysis proves invaluable in the initial design and prototyping of a vibration control system.

TABLE 6.2. Parameters for the system of Figure 6.65

Structure mass, $m = 50/2.2$ kg
Natural frequency, $\omega_n = 2\pi 2$ rad/s, such that
$k = m\omega_n^2$
Passive damping ratio, $\zeta = 0.3$, such that
$b = 2\zeta\omega_n m$
Coupling constant for the inertial actuator, $T = 0.5$ N/A
Winding resistance, $R_w = 2\ \Omega$
Ratio of actuator proof mass to structure mass, $\frac{m_a}{m} = \frac{1}{20}$



(a) Velocity response for various control gains



(b) Input voltage to actuator for various control gains

FIGURE 6.66. Frequency response of the controlled system from Figure 6.65.

6.7 SUMMARY

This chapter was devoted to linear system analysis. It is an important topic because some aspects of the behavior of many engineering systems can be profitably studied using linear models. For such systems an abundance of analytical tools are available, which have been conveniently packaged into a number of computer programs.

In the experience of the authors, there are three extremely important linear system concepts for engineers dealing with dynamic physical systems: eigenvalues, transfer functions, and frequency response. Eigenvalues have nothing to do with system inputs but only with the inherent system dynamics. For low-order systems we can obtain analytical expressions for the eigenvalues and determine system natural frequencies and damping. When a system is modeled and state variables are identified, we know how many eigenvalues there are and how each contributes to the system response. Eigenvalues reveal stability issues before any system responses are generated. It is worth one's time to understand the message of eigenvalues.

Transfer functions are the ratio of an output to an input in the s-domain. They are straight forwardly derived from the state-space representation of a system. Transfer functions and their block diagram representations are the starting point for virtually all development of automatic control.

Frequency response is the response of a system output variable to a fixed-frequency system input variable. It is a convenient test to perform on actual physical systems because a fixed-frequency sinusoidal input is relatively easy to generate in a laboratory environment. By substitution of $s = j\omega$ in a transfer function, the complex frequency-response function is generated and from the complex frequency-response function it is straightforward to obtain magnitude and phase information. The frequency response reveals troublesome resonances, high-frequency behavior, frequencies where the response is small, and more.

The final part of this chapter introduced the topic of automatic control, which deals with using sensors and actuators to change the inherent behavior of a physical system. Control systems typically act on the error between a desired response and the actual response, and controller design deals with the control action that operates on the error signal and outputs commands to the actuators, which then move the system so as to reduce the error. The "controller" can take many forms, but in the initial design phase it tends to be a transfer function with characteristics that come from the design concepts presented in this chapter.

This book is devoted to modeling physical engineering systems and the first chapters have been devoted to linear system concepts. The fact is there are no truly linear systems. The second half of this book deals with more realistic nonlinear systems modeled using the same procedures as already presented. So the reader should be careful when modeling a system as linear or linearizing a nonlinear system so that linear analysis tools can be used. It is typical in model development to linearize a nonlinear model, develop controls for the linear system, and then test the ability of the controller on the nonlinear system using computer simulation.

REFERENCES

[1] K. Ogata, *System Dynamics*, 3rd ed., Upper Saddle River, NJ: Prentice-Hall, 1998.

[2] J. L. Melsa and S. K. Jones, *Computer Programs for Computational Assistance in the Study of Linear Control Theory*, 2nd ed., New York: McGraw-Hill, 1973.

[3] W. J. Grantham and T. L. Vincent, *Modern Control Systems Analysis and Design*, New York: Wiley, 1993.

[4] Wylie, C. R. *Advanced Engineering Mathematics*, New York: McGraw-Hill, 1960.

[5] Matlab and Simulink software, Natick, MA: Mathworks Co.

[6] K. Ogata, *Modern Control Engineering*, Upper Saddle Rriver, NJ: Prentice Hall, 2010.

[7] G. Franklin, J. D. Powell, and Emami-Naeini, Feedback Control of Dynamic Systems, 6th ed., Upper Saddle River, NJ: Prentice Hall, 2010.

[8] R. C. Dorf and R. H. Bishop, *Modern Control Systems*, 11th ed., Upper Saddle River, NJ: Prentice Hall, 2008.

[9] Evans, W. R., Control system synthesis by root locus method, *Trans. AIEE*, vol. 69, 1950.

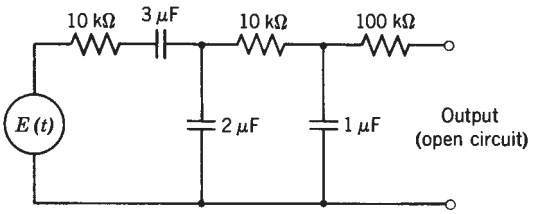
[10] Evans, E. R., The use of zeros and poles for frequency response or transient response, *Trans. ASME*, vol. 76, 1954.

PROBLEMS

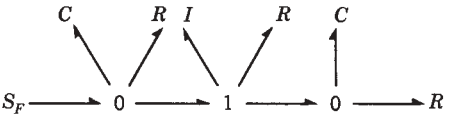
6-1. Consider the undamped oscillator with natural frequency of 10 rad/s. See Eq. (6.16). Determine the free response starting from zero initial position and an initial velocity.

Apply Euler's integration method, Eq. (6.6), for a few steps. What would be a reasonable way to pick Δt so that the digital solution would be accurate?

6-2. For the filter circuit shown, write state and output equations with the voltage $E(t)$ as the input and the terminal voltage as the output. Write the characteristic equation from which the eigenvalues could be determined.



6-3. Consider the bond graph shown below:



Assuming that all elements are linear, write state equations. Pick two variables that could be output variables, and write output equations for them. Identify the coefficients in Eqs. (6.1a) and (6.2a) with your equation coefficients, as was done in Eq. (6.18). Sketch mechanical, electrical, and hydraulic systems that could be modeled with this bond graph.

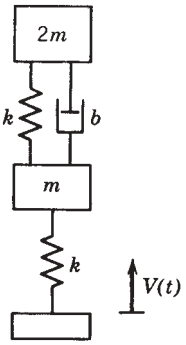
6-4. Using equations such as Eq. (6.43) or sketches in the complex plane, demonstrate that

$$\frac{d}{dt}\,\mathrm{Re}(e^{j\omega t}) = \mathrm{Re}\left(\frac{d}{dt}e^{j\omega t}\right).$$

6-5. If an automobile weighs 3000 lb and has an undamped natural frequency of 1.0 Hz when it oscillates on its main suspension springs, what should the total stiffness of the four shock absorbers be if a damping ratio of 0.707 is to be achieved? Consider the shock absorbers to be linear dampers even though, in practice, shock absorbers are distinctly nonlinear.

6-6. Sketch a frequency-response plot for the output voltage divided by the input voltage for the system of Problem 6-2.

6-7. In the system shown, consider $V(t)$ to be a velocity input. Set up state equations for the system, and let the acceleration of the top mass be a system output. (Note that the acceleration is proportional to the total force on the mass.) Find the ratio of the acceleration to the input velocity when the velocity is sinusoidal, and sketch the frequency response plot.

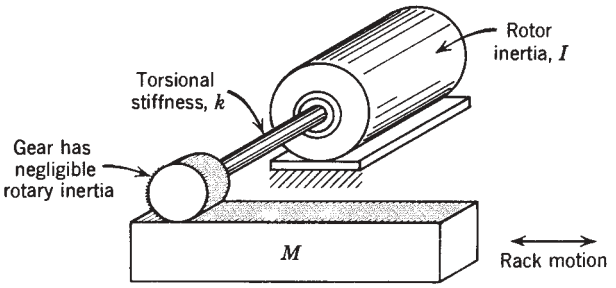


6-8. If a first-order system is excited by a cosine wave starting from a zero initial condition at $t = 0$, sketch the forced response, free response, and total response.

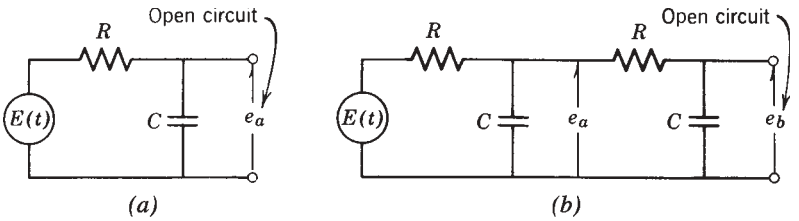
6-9. By labeling signals and writing equations implied by the blocks, demonstrate the validity of the block diagram identity of Figure 6.20.

6-10. Using Figure 6.20, reduce Figure 6.19*a* to Figure 6.19*b*.

6-11. The rotor of the motor shown is connected through a flexible shaft to the pinion, which drives the rack and load mass M. Assuming the pinion may be modeled as having no rotary inertia, construct a bond graph for the free motion of the system. Find the eigenvalues of the system and interpret them.



6-12. Figures *a* and *b* below depict two linear filters that are supposed to remove unwanted high-frequency noise from a voltage signal $E(t)$. In Figure *a* the output voltage measured when no current is allowed to flow is e_a; in Figure *b* the output is e_b. All resistors have the same resistance R, and all capacitors have the capacitance C.

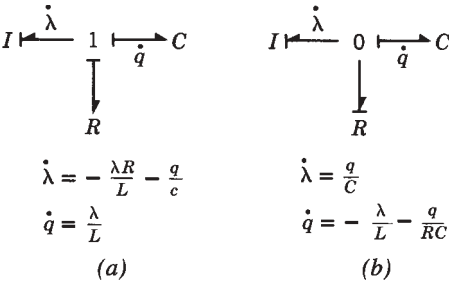


(a) For Figure *a* derive state equations, and relate the output, e_a, to the state variables. For Figure *b* find state equations, and relate both e_a and e_b to the state variables.

(b) For Figure *a* find the transfer function $H_{e_a/E}(s)$ relating e_a and E. For Figure *b* find transfer functions for $H_{e_a/E}$ and $H_{e_b/E}(s)$ e_a and e_b.

(c) Should the two transfer functions for e_a in Figures a and b not be equal? Should the transfer function for e_b not be the square of the transfer function for e_a? If so, why?

6-13. Consider the bond graphs and state equations below.



(a) Draw the corresponding circuits.

(b) Describe qualitatively the effect in both cases when $R \to 0$ and $R \to \infty$.

(c) Find relations for the system eigenvalues in both cases, using the following procedure:

i. Let $\dot{x} = Ax$ be the state equation.

ii. If $x = Xe^{st}$, then $sX = AX$, or $[sI - a][X] = [0]$ if $[I]$ is the unit matrix.

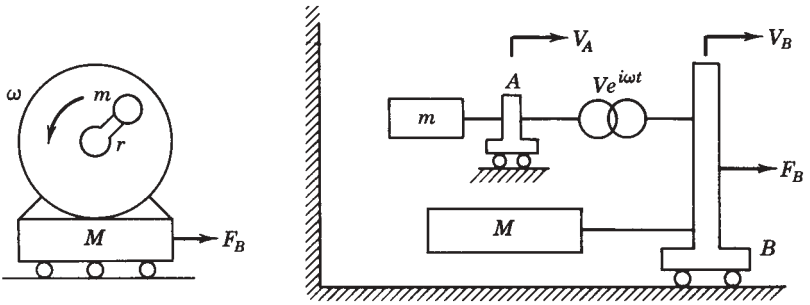
iii. The eigenvalues are values of s for which $\text{Det}[sI - A] = 0$, since only when the determinant vanishes can X be nontrivial.

Note that A is given, so all that is required is to form $sI - A$ and set its determinant equal to zero. The results should square with (b).

6-14. The sketches below show a schematic diagram of a rotating eccentric vibrator and a "mechanical network" representation of the device. The assumption is that the motor produces a simple harmonic relative horizontal velocity of

$$V \cos \omega t = r\omega \cos \omega t = \text{Re}\left(V e^{j\omega t}\right)$$

between the rotating mass m and the main mass M. Because of the assumed constraints, we do not discuss vertical forces or vertical motion.



(a) Construct a bond graph for the system, and fully augment it. Assume F_B is a known force applied to the system.

(b) Write a state equation for the system, and write output equations for V_A and V_B.

(c) Verify, using the results of (b), that the following terms sometimes used to describe the exciter are valid when $Ve^{j\omega t}$ is the complex representation of $V \cos \omega t$ for the velocity source:

i. The *internal mobility* looking back at B is the ratio of velocity to force when the velocity generator is stopped and locked in position:

$$\frac{V_B}{F_B} = \frac{1}{j\omega (M + m)}.$$

ii. The *blocked force* output at B is F_B when some external agent prevents motion of B:

$$F_{B(\text{blocked})} = j\omega m V.$$

iii. The *free-velocity* output at B is the velocity at B when there is no force F_B:

$$V_{B(\text{free})} = \frac{m}{M + m} V.$$

iv. A general network theorem states that the internal mobility is the quotient of the free velocity and the blocked force. See H. H. Skilling, *Electrical Engineering Circuits*, New York: John Wiley & Sons, 1957, p. 339. Verify the theorem for this example.

6-15. Reinterpret the bond graph in Figure 6.3*b* as if it referred to an electric circuit with parameters R, C, and L. After showing the circuit, redraw the Bode plots of Figure 6.14 showing the significant points on the plot in terms of the electrical rather than the mechanical parameters.

6-16. A transfer function between an output Y and an input X is

$$\frac{Y(x)}{X(s)} = \frac{G(s^2 + \omega_0^2)}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad \omega_n < \omega_0.$$

(a) Let $s = j\omega$ and write down the complex frequency-response function.

(b) If $x(t)$ is a harmonic input,

$$x = x \cos(\omega t);$$

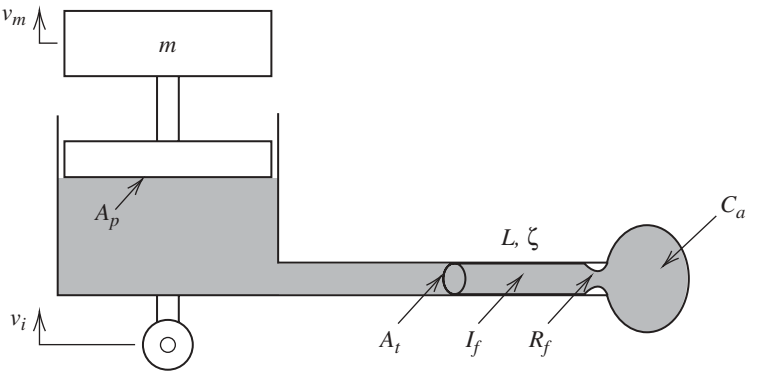
then, for a linear system, we know

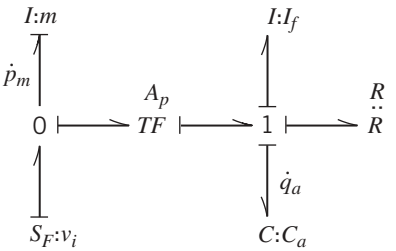
$$Y = Y \cos(\omega t + \varphi).$$

Derive expressions for Y and φ.

(c) Sketch Y/X versus ω for low, medium, and high ζ, and imagine a physical system that would behave like this.

6-17. The system below is supposed to isolate the mass, m, from ground motion, $v_i(t)$. The mass sits on a fluid-filled displacer that is attached to an air compliance using a long, fluid-filled tube. The tube has area A_t, length L, and fluid density ρ. The nominal air volume is V. Construct a bond graph for this device and show that yours is the same as the one given.





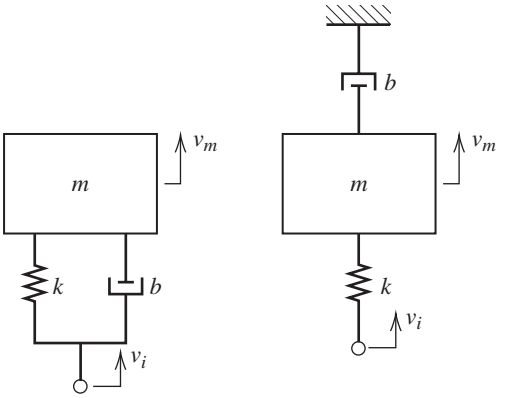
(a) Derive the state equations. There is some derivative causality, so use procedures from Chapter 5 to deal with this problem.

(b) Derive the transfer function relating output, v_m, to input, v_i.

(c) Compare this transfer function with the one from Problem 6-16. What do you think of this device as a vibration isolator?

(d) From your knowledge of how the fluid inertia, I_f, relates to physical system parameters, can you suggest design changes that will improve the high-frequency isolation?

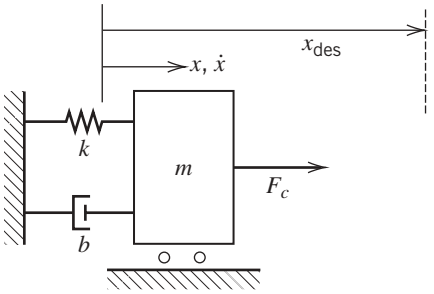
6-18. For the two devices below, construct bond graph models, derive state equations, and derive transfer functions relating $v_m(s)$ to $v_i(s)$. Turn these transfer functions into frequency-response functions and sketch $|v_m/v_i|$ versus frequency. Comment on how these devices compare with respect to vibration isolation.



6-19. In control systems, an error is generated between a desired output and an actual output, and this error is passed through a controller filter, the output of which is an effort or flow that moves the system in a direction that reduces the error. A typical control filter is the PID controller, which produces a component of actuator output proportional (P) to the error, proportional to the integral (I) of the error, and proportional to the derivative (D) of the error. In the s-domain, this is given by

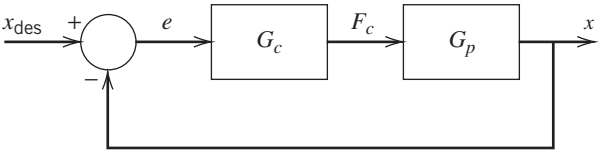
$$G_c(s) = K_P + K_D s + \frac{K_I}{s}.$$

It is proposed to use the PID controller in a position control system. The physical system is modeled as shown below, where we have control of the force, F_c.



(a) Construct a bond graph model of this system and derive the transfer function relating mass position, x, to force, F_c. Call this transfer function $G_p(s)$.

(b) A block diagram of the entire control system is shown. Derive the closed-loop transfer function relating the output, x, to the desired position, x_{des}.



(c) Compare the system eigenvalues without control (open loop) to those when the controller is operating (closed loop). Show how the controller gains can change the eigenvalues.

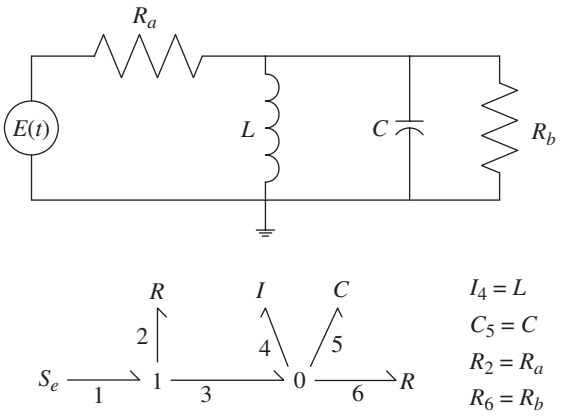
(d) Sketch the response of x to a step change in x_{des} and comment on the ability of the PID controller to produce a reasonable response.

6-20. The figure shows a relatively simple electric circuit and a numbered bond graph. The circuit has two energy storage elements, an inductor and a capacitor, so unless there is derivative causality involved, this should be a second-order system. Assume that all elements are linear, with coefficients as shown in general or physical form in the table to the right of the bond graph.

(a) Apply causality to the bond graph and write the state equations in matrix form using either version of the coefficients. The results should be of the form of Eq. (6.1b) and the matrices should resemble those in Eq. (6.18).

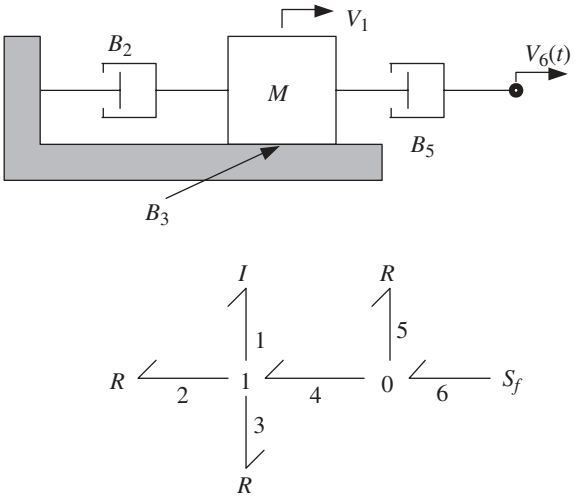
(b) Find the characteristic equation as was done in Eqs. (6.29) and (6.30) by setting the determinant of the matrix $[sI - A]$ to zero.

(c) By matching your characteristic equation coefficients to the coefficients in the version shown in Eq. (6.29), determine the undamped natural frequency, ω_n, and the damping ratio, ζ, in terms of the physical parameters L, C, R_a, and R_b.



6-21. The figure below shows a simple mechanical system consisting of a mass, two dashpots, and a velocity source with friction between the mass and

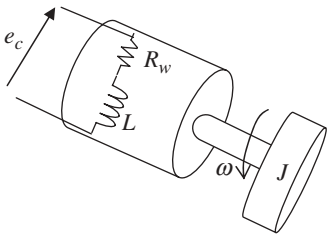
the stationary base. The dashpots and the ground friction are represented as if they are linear friction elements with coefficients B_2, B_3, and B_5. The input forcing function is the velocity $V_6(t)$, which can be represented in general terms as the flow source $f_6(t)$.



(a) First, convince yourself that the bond graph does represent the system. Then apply causality and write the state equation using general effort and flow variables and generalized coefficients such as, I_1, R_2, and so on. Based on the equation, determine the time constant in terms of the generalized coefficients as well as the corresponding physical coefficients, M, B_2, and so forth.

(b) Now suppose that the applied velocity, $V_6(t)$, is replaced by an applied force, $F_6(t)$. (The applied force is represented in general terms as the effort $e_6(t)$. It is too bad that force and flow both start with F. In part a, $f_6 = V_6$, but here e_6 is the force F_6.) Again apply causality and write the state equation. What is the time constant now under the new type of forcing?

6-22. The figure shows a dc motor with winding resistance and inductance attached to rotating inertia.



(a) Construct a bond graph model of this system, assign causality, identify the state variables, and derive the state equations. Put these into standard matrix format $[\dot{x}] = [A][x] + [b]u$.

(b) Put the system into the s-domain and derive the characteristic equation. Solve for and interpret the eigenvalues. Is there a natural frequency? What is the damping ratio?

(c) Derive the transfer function relating the output angular velocity to the input voltage.

(d) Modify this transfer function and show the transfer function relating the output angular position of the rotating inertia to the input voltage.

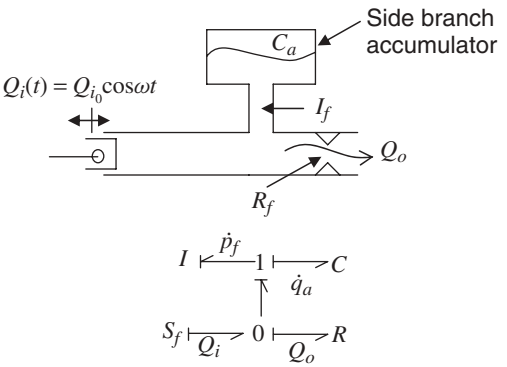
6-23. The figure shows a part of a piping system with an oscillatory flow input Q_i at one end, a side branch accumulator with fluid inertia I_f and fluid compliance C_a, and an orifice resistance with flow output Q_o. The bond graph for the system is also shown with state variables identified.

(a) Derive the state equations and put into matrix form $[\dot{x}] = [A][x] + [b]u$.

(b) Put into the s-domain and derive the characteristic equation and the eigenvalues.

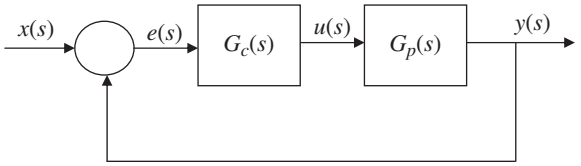
(c) Derive the transfer function $\dfrac{Q_o}{Q_i}(s)$.

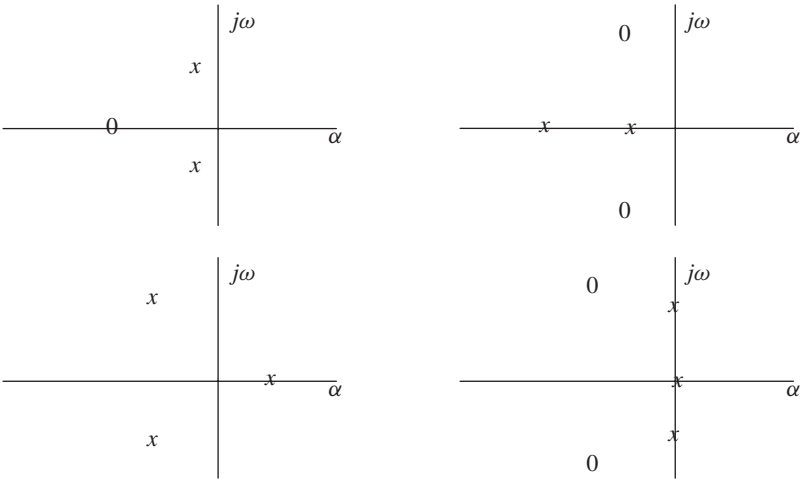
(d) Derive the complex frequency response function and sketch the magnitude response. Comment on any interesting behavior.



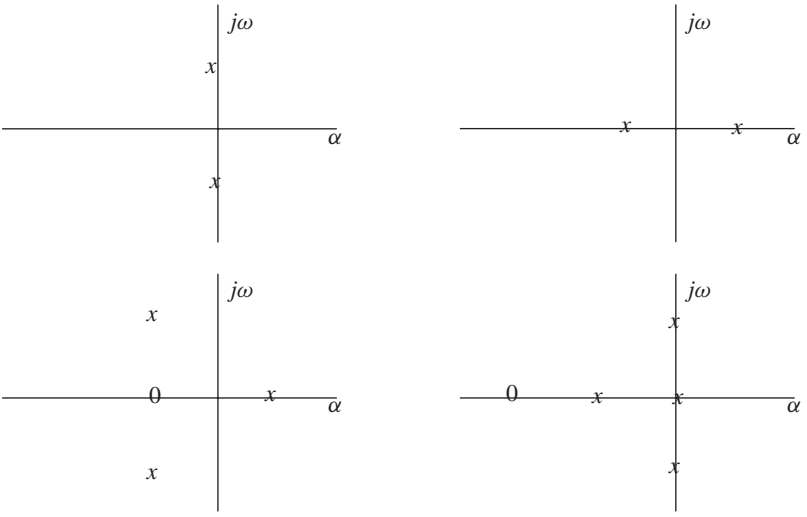
6-24. The classic starting point for single input–single output control is the block diagram shown here where poles and zeros of the forward loop come from the plant and the controller transfer functions. Shown also are several open-loop pole/zero configurations. It is interesting that once poles and zeros are shown on the s-plane, we lose knowledge of which poles and zeros are part of the plant and which come from the controller.

For each of the pole/zero configurations shown, sketch the root locus. Show the asymptotes and where the locus exists on the real axis.





6-25. Some linear physical systems have open-loop pole/zero configurations shown in the figure below. Using approximate root locus, show how P, PD, PID, and/or lead-lag control will affect stability of the closed-loop systems.



6-26. A velocity control system is shown in the figure below. It consists of a voice coil actuator with control voltage input e_c, coupling constant T, winding resistance R_w, and inductance L. The actuator acts through a damper b to ultimately move the load mass m. A disturbance force F_d can also act on the load mass. A bond graph of the system is included in the figure with causality assigned and state variables identified.

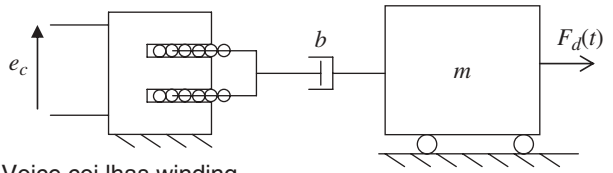
(a) Derive the state equations and put into standard matrix format $[\dot{x}] = [A][x] + [b]u$.

(b) Put your state equations in the s-domain and derive the plant transfer function relating the load mass velocity to the control voltage and the disturbance transfer function relating the load mass velocity to the disturbance force.

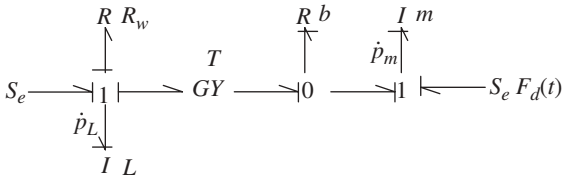
(c) Draw a block diagram that includes both the plant transfer function and the disturbance transfer function. Close the loop and show a general controller operating on the error between the desired and actual velocity.

(d) Show the open-loop poles and zeros of the plant and discuss the control action that potentially produces an acceptable step response.

(e) Discuss how the proposed controller works when the disturbance is acting.



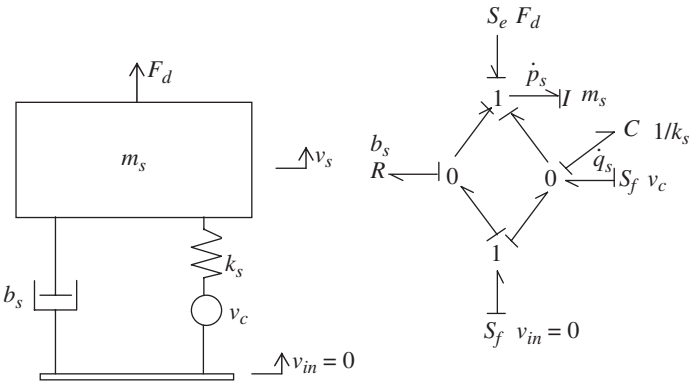
Voice coi lhas winding
resistance, R_w
winding inductance, L
and coupling constant, T



6-27. Problem 6-26 can be turned into a position control system by simply dividing the transfer functions for velocity by s. Derive the position transfer functions and carry out the same procedure as Problem 6-26.

6-28. The figure shows a simple representation of a load-leveling system. This might be a truck suspension where a load change represented by the disturbance force F_d will change the height of the structure mass and we

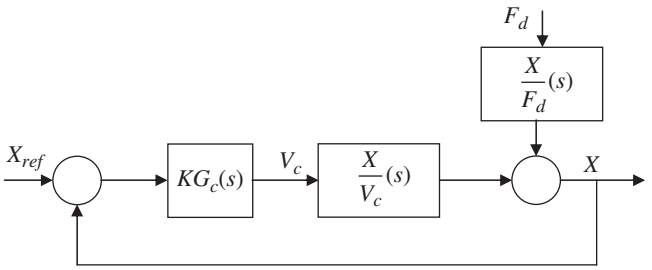
desire the structure mass to return to its original position. On trucks, this control is most often done using pneumatic devices, but here we simply postulate a controllable velocity source v_c.



The figure also shows the system bond graph. The reader should derive the state equations, with the result

$$\frac{d}{dt}\begin{bmatrix} p_s \\ q_s \\ x \end{bmatrix} = \begin{bmatrix} -\dfrac{b_s}{m_s} & k & 0 \\ -\dfrac{1}{m_s} & 0 & 0 \\ \dfrac{1}{m_s} & 0 & 0 \end{bmatrix} \begin{bmatrix} p_s \\ q_s \\ x \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} v_c + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} F_d$$

where the position of the mass x has been added as a state.



(a) Derive the transfer functions $\dfrac{X}{V_c}(s)$ and $\dfrac{X}{F_d}(s)$. The block diagram for the control system is shown above.

(b) Using root locus, demonstrate a control action $KG_c(s)$ that will produce stable closed-loop eigenvalues.

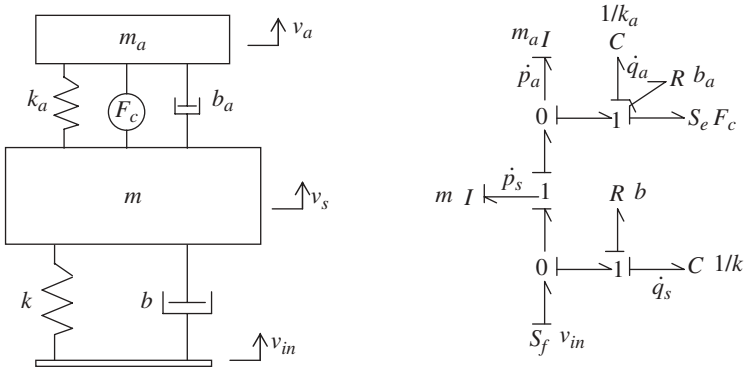
(c) Will this control also produce zero steady-state error to a step change in F_d?

6-29. In Problem 6-28 propose state variable feedback of the form

$$v_c = K\left(x_{\text{ref}} - [k_{p_s}\ k_{q_s}\ k_x]\begin{bmatrix} p_s \\ q_s \\ x \end{bmatrix}\right)$$

and derive the closed-loop transfer function between X and F_d. Comment on possible placement of closed-loop poles and on the steady-state response to step change in F_d.

6-30. A vibration control system similar to that of Figure 6.65 is shown below along with a bond graph model and identified state variables. A base excited spring–mass–damper system has an inertial actuator attached to the structure mass. The force generated by the actuator is simply represented by the control force F_c. In this system the actuator mass is attached to the structure through a spring and damper k_a, b_a.



In vibration control we often are concerned with isolation of a single frequency. Helicopters, propeller aircraft, and power-generating equipment are examples of systems that operate at near constant rotational speed and thus exhibit fixed-frequency vibration. For such systems it can be beneficial for the actuator to have its own natural frequency.

(a) For the system here derive the state equations and put them into the standard matrix representation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}_d v_{\text{in}} + \mathbf{b}_c F_c$.

(b) Put the state equations in the s-domain and derive the open-loop transfer functions relating the structure mass velocity to the disturbance velocity and the structure mass velocity to the control force, i.e., derive

$$\frac{V_m}{V_{\text{in}}}(s) \quad \text{and} \quad \frac{V_m}{F_c}(s).$$

(c) Derive the complex frequency response functions and note particularly the numerator of $\frac{V_m}{V_{\text{in}}}(s)$. You should observe that if the actuator damping b_a is zero, then the structure mass motion is zero at the special frequency, $\omega_0^2 = \frac{k_a}{m_a}$. If we know the single frequency disturbance, then we can "tune" the actuator to produce a small response at this frequency without using any control power.

(d) Propose a control action that simply cancels the effect of actuator damping, derive the closed loop response of $\frac{V_m}{V_{\text{in}}}(s)$, and demonstrate that such a control action allows zero response at the "tuned" frequency. Also demonstrate what happens if you go too far with your control action and overcompensate for the actuator damping.

7

MULTIPORT FIELDS AND JUNCTION STRUCTURES

In the first part of this book, dynamic models for a variety of physical systems were constructed using the 1-port elements —R, —C, —I, S_e—, S_f—; the 2-port elements —TF— and —GY—; and the 3-port 0- and 1-junctions. In this chapter, *fields*, which are multiport generalizations of —R, —C, and —I elements, are introduced We also introduce the concept of a *junction structure*, which is an assemblage of power-conserving elements,

$$\overset{|}{0}, \overset{|}{1}, \text{—}TF\text{—}, \text{ and } GY\text{—}.$$

Using fields and junction structures, one may conveniently study systems containing complex multiport components using bond graphs. In fact, bond graphs with fields and junction structures prove to be a most effective way to handle the modeling of certain multiport systems, combining both structural detail and clarity with ease of visualization. Succeeding chapters demonstrate the application of the elements introduced here to the modeling of systems containing multiport devices.

In this chapter, we will be dealing with *true bond graphs* in which the product of effort and flow variables represents power flowing on a bond. In Chapter 12, *pseudo-bond graphs* will be introduced to represent thermofluid systems. Pseudo-bond graphs use variables that are *not* power variables. In many cases, the pseudo-bond graph yields a simpler model than the equivalent true bond graph for such systems. On the other hand, what appear to be R-fields and C-fields for pseudo-bond graphs do not have all the properties to be discussed here because the power and energy associated with a bond are not related to the effort and flow variables in the same way as they are for true bond graphs.

Copyright © 2012 John Wiley & Sons, Inc.

7.1 ENERGY-STORING FIELDS

In Chapter 3 it was shown that the elements —C and —I can store energy and can give back stored energy without loss. This is true no matter what the constitutive laws for these elements happen to be. Any single-valued functional relationship between effort and displacement defines an energy-conservative capacitance element, and any single-valued relationship between flow and momentum defines an energy-conservative inertia element. The multiport generalizations of —C and —I, which we shall call C-fields and I-fields, respectively, will also be conservative. As we shall see, the conservation of energy in the multiport cases imposes a special constraint on the constitutive laws for C- and I-fields.

7.1.1 C-Fields

The symbol for a C-field is simply the letter C, with as many bonds as the C-field has ports. An n-port C-field is shown in Figure 7.1. Note that the flow at the ith port has been labeled $\dot{q}_i$ and that an inward sign convention is assumed. The energy $\mathbf{E}$ stored in the C-field can then be expressed as

$$\begin{aligned}\mathbf{E} &= \int_{t0}^{t} \sum_{i=1}^{n} (e_i f_i)\, dt = \int_{t0}^{t} \sum_{i=1}^{n} e_i \dot{q}_i\, dt \\ &= \int_{q0}^{\mathbf{q}} \sum_{i=1}^{n} e_i(\mathbf{q})\, dqi = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{e}(\mathbf{q})\, d\mathbf{q} = \mathbf{E}(\mathbf{q}),\end{aligned} \tag{7.1}$$

in which the identity $f_i\, dt = dq_i$ has been used, and column vectors of efforts and flows have been defined thus:

$$\mathbf{q} \equiv \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_n \end{bmatrix}, \qquad \mathbf{e}(\mathbf{q}) \equiv \begin{bmatrix} e_1(\mathbf{q}) \\ e_2(\mathbf{q}) \\ e_3(\mathbf{q}) \\ \vdots \\ e_n(\mathbf{q}) \end{bmatrix}, \tag{7.2}$$

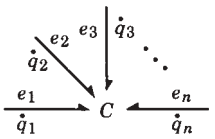


FIGURE 7.1. The symbol of an n-port C-field.

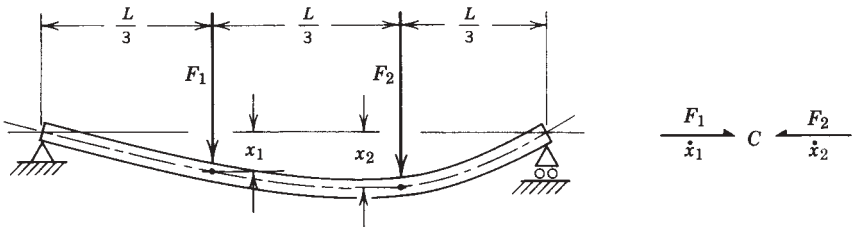


FIGURE 7.2. An elastic beam deformed by the action of two forces and represented as a 2-port C-field. Uniform beam: elastic modulus E; area moment of inertia I.

and $\mathbf{e}\ d\mathbf{q}$ represents the scalar or inner product, which could also be represented as $\mathbf{e}^{\mathrm{t}}\ d\mathbf{q}$ using the transpose of the column matrix $\mathbf{e}$. Note that we have assumed that each effort is a function of all the displacement variables.

At time t_0, $\mathbf{q} = \mathbf{q}_0$, and at time t, $\mathbf{e} = \mathbf{e}(\mathbf{q})$ represents the constitutive laws of the C-field evaluated at $\mathbf{q}(t)$, the instantaneous value of the vector of displacements.

Before going on to study the properties of general C-fields as defined by Eqs. (7.1) and (7.2), it may be worthwhile to exhibit some relatively simple C-fields as they arise in practice. Some components of systems are described naturally by C-fields. The beam of Figure 7.2, for example, might be part of a vibratory system with masses or other elements attached at the locations indicated by the F_1, $\dot{x}_1$ and F_2, $\dot{x}_2$ ports. If one neglects the mass of the beam (which is permissible if the system frequencies are sufficiently low) and if one neglects at least temporarily the possible power losses associated with inelastic material behavior and the support pivots, then the beam is a purely elastic structure and can be represented as a 2-port C-field. The beam then may be represented by constitutive relations among F_1, F_2, x_1, and x_2.

Many engineering undergraduates could find the constitutive laws for this C-field at some point in their careers, although later they may forget just how to do it. A convenient way to represent the beam is through the use of *superposition* to add up the effects on x_1 and x_2 of the forces F_1 and F_2. This procedure (see Reference [1], for example) works only because the beam is represented for a *linear* range of strains and deflections. The result is conveniently expressed in a matrix form:

$$\begin{bmatrix} x_1 \\ \hline x_2 \end{bmatrix} = \frac{L_3}{243\ EI} \left[\begin{array}{c|c} 4 & \frac{7}{2} \\ \hline \frac{7}{2} & 4 \end{array}\right] \begin{bmatrix} F_1 \\ \hline F_2 \end{bmatrix}. \tag{7.3}$$

This form of the constitutive laws, which gives displacements as a function of efforts, is the *compliance form*. As long as the matrix in Eq. (7.3) is invertible, one could solve the equation for F_1 and F_2 in terms of x_1 and x_2. This would be the *stiffness form* for the constitutive laws.

When an element is described from the beginning as a set of effort–displacement relations at n ports, as in the beam example, we call the model of the element an *explicit field*. Generally, it is much more convenient to treat explicit-field elements as fields, although in some cases it would be possible to find an equivalent system of 1-port energy-storing elements bonded together with junction structure elements that would have the same port constitutive laws as the field. On the other hand, when systems are assembled, it often happens that a group of 1-port energy-storing elements, say —C's, are interconnected with

$$\overset{|}{0}, \overset{|}{1}, \quad \text{and} \quad -TF-$$

elements. Such a subsystem often can be usefully treated as a field. Such a field is called an *implicit field*, since the field constitutive laws at the external ports must be deduced from the constitutive laws of the elements constituting the field. An example is shown in Figure 7.3. Except for the sign convention on one internal bond, the electrical and mechanical systems shown in Figure 7.3 have the same bond graph. If one imposes integral causality on the 1-port C-elements as in Figure 7.3a, then it is clear that the state variables are x_1 and δ_2, namely, the deflections of the two springs. Thus, it is not immediately clear that the C-field representation in terms of x_1 and x_2 is possible. Using the mechanical example, let us work out the C-field constitutive laws.

The causal strokes indicate that we can find F_1 and F_2 as output variables in terms of the state variables x_1 and δ_2:

$$F_1 = k_1 x_1 - k_2 \delta 2, \qquad F_2 = 0x_1 + k_2 \delta 2, \tag{7.4}$$

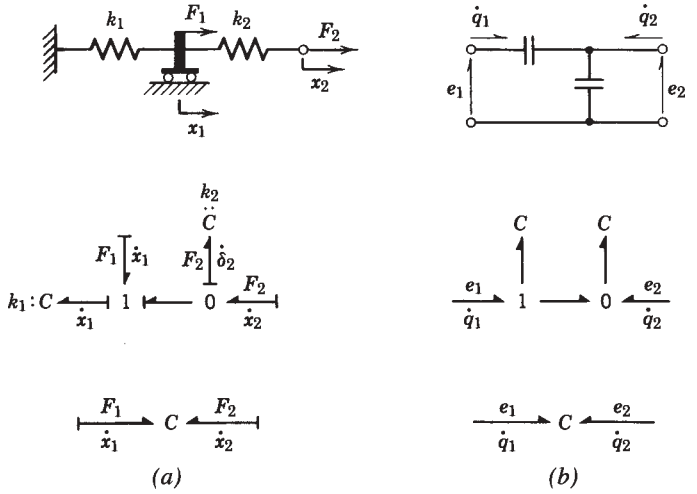


FIGURE 7.3. A C-field composed of 1-port C-elements and junction elements: (a) A mechanical example; (b) an electrical example.

where, for the sake of preserving simplicity in the example, the springs have been assumed to be linear. The state equations are

$$\dot{x}_1 = \dot{x}_1(t) \quad \text{(given)}, \tag{7.5}$$

$$\dot{\delta}_2 = \dot{x}_2(t) - \dot{x}_1(t), \tag{7.6}$$

in which $\dot{x}_1$ and $\dot{x}_2$ are determined by some system external to the C-field and thus function as input variables.

Using Eq. (7.6), we see that δ_2 may be expressed in terms of the C-field variables, x_1 and x_2, by means of a time integration:

$$\delta_2 = x_2 - x_1 + \text{const.} \tag{7.7}$$

Referring to Figure 7.3*a*, it can be seen that x_1 and x_2 can be defined such that when $x_1 = x_2 = 0$, $F_1 = F_2 = \delta_2 = 0$. Thus, the integration constant in Eq. (7.7) can be set to zero when x_1 and x_2 represent deviations from the positions at which the springs are unstretched. Using Eq. (7.7), we can eliminate δ_2 from Eq. (7.4). Then

$$\left[\begin{array}{c} F_1 \\ \hline F_2 \end{array}\right] = \left[\begin{array}{c|c} k_1 + k_2 & -k_2 \\ \hline -k_2 & k_2 \end{array}\right] \left[\begin{array}{c} x_1 \\ \hline x_2 \end{array}\right], \tag{7.8}$$

which represents the constitutive laws of a 2-port C-field in stiffness matrix form. (The reader may benefit from carrying out the steps required to express the constitutive laws of the electrical C-field of Figure 7.3*b* in a form analogous to Eq. (7.8).)

An inspection of the constitutive laws of Eqs. (7.3) and (7.8) shows that for these linear C-fields, the *compliance matrix* (Eq. (7.3)) and the *stiffness matrix* (Eq. (7.8)) are symmetric. In general, compliance matrices and stiffness matrices, which are inverses of each other, are always symmetric. This can be proven in a form valid for even nonlinear C-fields by studying the stored-energy function (7.1).

A change in any component of $\mathbf{q}$, say Δq_i, will produce a change in $\mathbf{E}$, say $\Delta\mathbf{E}$. By a direct inspection of Eq. (7.1), one can deduce that the coefficient relating $\Delta\mathbf{E}$ to Δq_i is, in fact, e_i. Thus, the partial derivative of $\mathbf{E}$ with respect to q_i is just e_i, or

$$\frac{\partial \mathbf{E}}{\partial q_i} = e_i(\mathbf{q}), \qquad i = 1, 2, \ldots, n. \tag{7.9}$$

Because $\mathbf{E} = \mathbf{E}(\mathbf{q})$ is supposed to represent the stored-energy function, which we assume to be a single-valued scalar function of the vector $\mathbf{q}$, if $\mathbf{E}$ is smooth enough to have second derivatives, then

$$\frac{\partial e_i}{\partial q_j} = \frac{\partial^2 E}{\partial q_j \partial q_i} = \frac{\partial e_j}{\partial q_i}, \qquad i, j = 1, 2, 3, \ldots, n, \tag{7.10}$$

which shows how $e_i(\mathbf{q})$ and $e_j(\mathbf{q})$ are constrained by the existence of the stored-energy function, $\mathbf{E}$. In other words, not every set of constitutive laws expressing efforts in terms of displacements could come from an energy-storing, or energy-conservative, C-field. Physically, it is clear that elastic structures, capacitor networks, and the like cannot supply more energy than the amount previously stored, so Eq. (7.10) will apply to such devices.

In the linear case, Eq. (7.10) implies that stiffness matrices are necessarily symmetrical for any conservative C-field. If we use k_{ij} for stiffness matrix coefficients and $\mathbf{k}$ for the matrix itself, then the C-field constitutive laws for an n-port can be written thus:

$$e_i = \sum_{j=1}^{n} k_{ij} q_j$$

or

$$\mathbf{e} = \mathbf{kq},$$

and Eq. (7.8) is a particular example. The stored energy is

$$\mathbf{E}(\mathbf{q}) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} k_{ij} q_i q_j = \frac{1}{2} \mathbf{q}^{\mathrm{t}} \mathbf{kq}. \tag{7.11}$$

Application of Eq. (7.10) to Eq. (7.11) yields

$$k_{ij} = k_{ji}, \qquad \text{or} \qquad \mathbf{k}^{\mathrm{t}} = \mathbf{k}, \tag{7.12}$$

and since compliance matrices such as that in Eq. (7.3) are inverses of stiffness matrices and the inverse of a symmetric matrix is also symmetric, we conclude that compliance matrices must also be symmetric. The equations (7.12), which are often called *Maxwell's reciprocal relations*, are most easily derived from the more general nonlinear relations (7.10), but graphical derivations based on special integration paths are often used for the linear case. See Reference [2], for example.

The reciprocal relations for C-fields and other energy-storing fields are more useful and informative in some cases than it might at first appear. In later chapters dealing with a variety of applications, these relations will prove to be very important. For now, we show only a few examples that are less obvious than our first ones. The system of Figure 7.4 shows how the end of a cantilever beam can interact with a rotational and a translational port. Using beam superposition tables or any other method to find the beam deflections, the compliance matrix may be found and inverted to find the stiffness matrix:

$$\begin{bmatrix} F \\ \tau \end{bmatrix} = EI \left[\begin{array}{c|c} 12/L^3 & -6/L^2 \\ \hline -6/L^3 & 4/L \end{array}\right] \begin{bmatrix} x \\ \theta \end{bmatrix}. \tag{7.13}$$

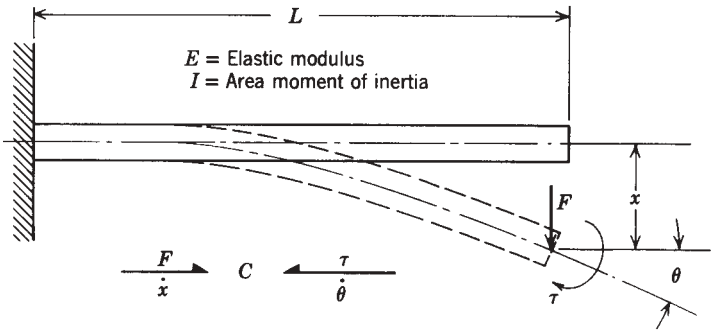


FIGURE 7.4. Cantilever beam represented as a C-field.

The symmetry of this matrix serves as a useful check on the correctness of the computations leading up to it. Furthermore, we can make statements such as the following. The number of newtons (the force) required to keep x at zero when $\theta = 0.1$ rad is equal to the number of newton-meters (the torque) required to keep θ at zero when $x = 0.1$ m. Or, if $F = 1$ N and $\tau = 0$, then θ in radians will be numerically equal to x in meters when $F = 0$ and $\tau = 1$ N-m. (The latter statement is based on the symmetry of the compliance matrix.)

Nonlinear C-fields are common in certain application areas. Here we show how a mechanical element that behaves as a linear 1-port in some restricted cases can be modeled as a nonlinear 2-port C-field when two-dimensional motion is permitted. The element, shown schematically in Figure 7.5, is a simple linear spring attached to ground at one end and deflecting under the action of forces F_x and F_y at the other end. It is possible to express F_x and F_y in

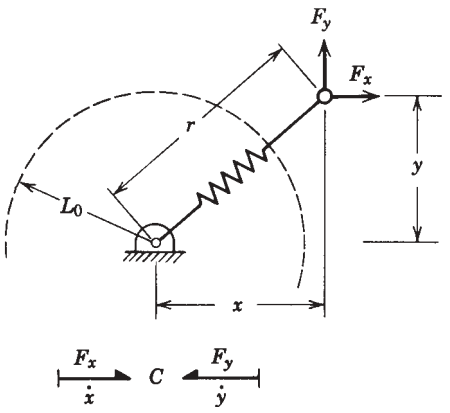


FIGURE 7.5. A nonlinear C-field representation of a linear spring, one end of which can move in a plane.

terms of x and y and thus to consider this device as a C-field. The constitutive relations are

$$F_x = kx - k(x^2 + y^2)^{-1/2} L_0 x,$$
$$F_y = ky - k(x^2 + y^2)^{-1/2} L_0 y, \tag{7.14}$$

where k is the spring constant and L_0 is the free length of the spring.

A check on (7.14) is given by Eq. (7.10):

$$\frac{\partial F_x}{\partial y} = -k(x^2 + y^2)^{-3/2} L_0 x 2y, \tag{7.15}$$

$$\frac{\partial F_y}{\partial x} = -k(x^2 + y^2)^{-3/2} L_0 y 2x. \tag{7.16}$$

Because the expressions on the right-hand sides in Eqs. (7.15) and (7.16) are identical, the constitutive laws (7.14) do represent a nonlinear conservative C-field.

Actually, it is easiest to derive Eqs. (7.14) by using energy methods. The stored energy $\mathbf{E}$ is easily expressed in terms of the radius r:

$$\mathbf{E} = \frac{1}{2} k (r - l_0)^2 = \frac{1}{2} k \left(r^2 - 2 r l_0 + l_0^2 \right),$$

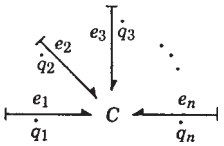
and, using $r^2 = x^2 + y^2$, we have

$$\mathbf{E} = \frac{1}{2} k \left[x^2 + y^2 - 2 l_0 \left(x^2 + y^2 \right)^{1/2} + l_0^2 \right]. \tag{7.17}$$

When Eq. (7.17) is used according to Eq. (7.9) to derive Eq. (7.14), the reciprocal relations, Eqs. (7.15) and (7.16), must follow immediately.

7.1.2 Causal Considerations for C-Fields

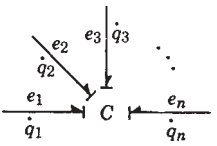
As in the case of 1-port C-elements, we may distinguish between integral and derivative causality for multiport C-fields, but multiport fields also admit mixed integral–derivative causalities. The completely integral causality form is



and the constitutive laws may be stated thus:

$$e_i = \Phi_{c_i}^{-1} (q_1, q_2, \ldots, q_n), \quad i = 1, 2, \ldots, n. \tag{7.18}$$

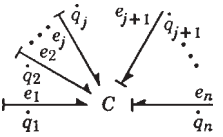
The completely derivative causality form is



and the constitutive laws may be restated in the form

$$q_i = \Phi_{c_i}(e_1, e_2, \ldots, e_n), \quad i = 1, 2, \ldots, n. \tag{7.19}$$

In Eqs. (7.18) and (7.19), the generally nonlinear functions $\Phi_{c_i}^{-1}$ and Φ_{c_i} are named in analogy with their 1-port counterparts from Chapter 3. In mechanical systems the compliance form corresponds to derivative causality and the stiffness form to integral causality. When the ports are numbered so that the first j ports have integral causality and the remainder have derivative causality, then a mixed causal form appears as



and the constitutive laws are

$$e_i = \Phi_i(q_1, q_2, \ldots, q_j, e_{j+1}, \ldots, e_n), \quad i = 1, 2, \ldots, j; \tag{7.20}$$

$$q_k = \Phi_k(q_1, q_2, \ldots, q_j, e_{j+1}, \ldots, e_n), \quad k = j+1, , \ldots, n. \tag{7.21}$$

The question of which causal forms are possible for an energy-storing field is an interesting one. For an explicit field that is specified by a set of constitutive laws in a particular causal form, the question is an algebraic one in which each causal form involves solving the given relationships again in terms of different input and output variables. For nonlinear fields, it can be very hard to decide whether this will be possible, and even for linear fields the problems are not trivial. Of course, to convert from completely integral to completely derivative causality or the reverse, one needs only to invert a matrix in the linear case. Since there are straightforward tests of matrices to determine whether or not the inverse exists, it is not hard to decide, for example, whether given a field in derivative causality the integral causality form is permissible. It is harder, however, to establish a universal rule for testing all the mixed causal forms that an n-port field can exhibit.

For implicit fields, on the other hand, the rules of causality can often be used on the elements of the field to deduce which causal forms of the field are possible. As a very simple example, consider the implicit field formed by a 1-port C-element

and a 0-junction, as shown in Figure 7.6. This system is readily treated as a 2-port C-field in integral causality (see Figure 7.6*a*). Reading the implicit-field causality, we may find e_1 and e_3, assuming that the 1-port C is linear:

$$e_1 = e_3 = \frac{q_3}{C_3}, \tag{7.22}$$

$$e_2 = e_3 = \frac{q_3}{C_3}. \tag{7.23}$$

Now we wish to convert from q_3 as a state variable to the field state variables, q_1 and q_2. Reading the bond graph, we have

$$\dot{q}_3 = \dot{q}_1 + \dot{q}_2, \tag{7.24}$$

or

$$q_3 = q_1 + q_2 + \text{const.} \tag{7.25}$$

If for simplicity we define q_1, q_2, and q_3 so that the integration constant vanishes in Eq. (7.25), then the field equations follow by substitution in Eqs. (7.22) and (7.23):

$$\left[\begin{array}{c} e_1 \\ \hline e_2 \end{array}\right] = \left[\begin{array}{c|c} 1/C_3 & 1/C_3 \\ \hline 1/C_3 & 1/C_3 \end{array}\right] \left[\begin{array}{c} q_1 \\ \hline q_2 \end{array}\right]. \tag{7.26}$$

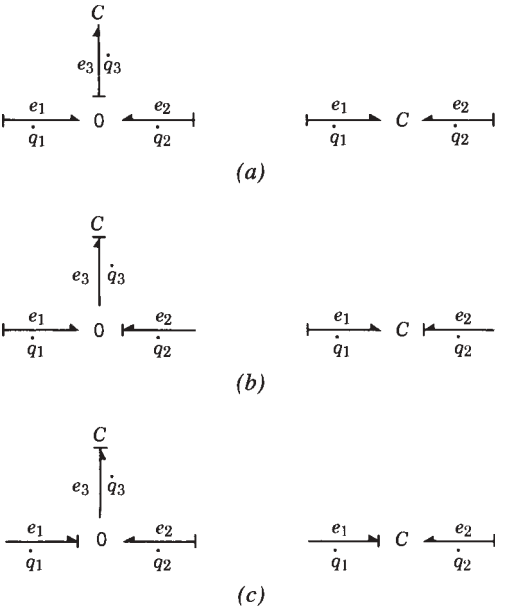


FIGURE 7.6. A simple implicit field: (*a*) integral causality; (*b*,*c*) mixed causality. (Derivative causality for this field is not possible.)

There are two ways to see that the completely derivative causal form is not possible: (1) the rules of causality when applied to the implicit field of Figure 7.6 do not allow the imposition of both e_1 and e_2 as inputs to the 0-junction, and (2) the determinant of the matrix in Eq. (7.26) clearly vanishes, indicating that Eq. (7.26) cannot be solved for q_1, q_2 in terms of e_1 and e_2.

On the other hand, two mixed causal forms are possible, as shown in Figure 7.6*b* and *c*. For the form of Figure 7.6*b*, the relations are derived by reading the implicit-field bond graph:

$$e_1 = e_2, \tag{7.27}$$

$$\dot{q}_2 = -\dot{q}_1 + \dot{q}_3 = -\dot{q}_1 + \frac{d}{dt}(C_3 e_3) = -\dot{q}_1 + \frac{d}{dt}(C_3 e_2). \tag{7.28}$$

Equation (7.28) must be integrated in time, and if q_1 and q_2 are properly defined, the integration constant can be made to vanish:

$$q_2 = -q_1 + C_3 e_2. \tag{7.29}$$

In matrix form, the mixed causal form is

$$\left[\begin{array}{c} e_1 \\ \hline q_2 \end{array}\right] = \left[\begin{array}{c|c} 0 & 1 \\ \hline -1 & C_3 \end{array}\right] \left[\begin{array}{c} q_1 \\ \hline e_2 \end{array}\right]. \tag{7.30}$$

Note that the integral form, Eq. (7.26), is symmetric, as it must be according to the energy argument, but mixed forms such as Eq. (7.30) will generally have antisymmetric terms.

Although almost any nontrivial system may be considered to contain an energy-storing field (as we have just seen, even a single —*C* and a 0-junction can be treated as a 2-port field), it is often not worthwhile to consider the manipulation of a bond graph into a form in which an implicit field is obvious. An exception to this general rule occurs when, in the course of assigning causality, it is found that one or more elements must be assigned derivative causality. Often this occurrence signals the presence of an implicit field that can usefully be converted into explicit form, thus simplifying further equation formulation. Of course, the methods of equation formulation presented in Chapter 5 for the cases involving derivative causality permit equation formulation without identifying fields. Although there is no less labor involved in converting an implicit field with derivative causality into explicit form, as we do here, this process may increase one's insight into the causes and nature of derivative causality. Also, the explicit-field representation can be used repeatedly, thus avoiding the necessity of handling derivative causality in each system of which the field is a part.

As an example, consider the electrical subsystem shown in Figure 7.7*a*. The bond graph of Figure 7.7*b* does not appear to have any unusual characteristics, but a little experimentation with causal assignments will show that the three —*C* elements cannot all have integral causality. What this means is that the three displacements—in this case, three electrical charge variables—cannot be

independent. In Figure 7.7*c*, the *C*-field is singled out for study. In this form, the charges q_2 and q_6, for example, can play the roles of state variables, and the charge q_4 will be statically related to q_2 and q_6. Let us now find the explicit-field representation shown in Figure 7.7*d*, in which the charge variables q_1 and q_7 will play the roles of state variables.

Reading the graph of Figure 7.7*c*, we find that the port output variables, e_1 and e_7, can be expressed in terms of the state variables, q_2 and q_6 (for simplicity, we assume all the —*C* elements are linear, although our operations could be carried out for nonlinear —*C* elements as well):

$$e_1 = e_2 = \frac{q_2}{C_2}, \qquad e_7 = e_6 = \frac{q_6}{C_6}. \tag{7.31}$$

The state equations may be used to relate q_2 and q_6 and q_7:

$$\begin{aligned}\dot{q}_2 &= f_1 - f_3 = f_1 - f_4 = f_1 - \frac{d}{dt}q_4 \\ &= f_1 - \frac{d}{dt}C_4 e_4 = f_1 - \frac{d}{dt}C_4(e_3 + e_5) \\ &= f_1 - \frac{d}{dt}C_4(e_2 + e_6) = f_1 - \frac{d}{dt}C_4\left(\frac{q_2}{C_2} + \frac{q_6}{C_6}\right).\end{aligned} \tag{7.32}$$

Similarly,

$$\dot{q}_6 = f_7 - \frac{d}{dt}C_4\left(\frac{q_2}{C_2} + \frac{q_6}{C_6}\right). \tag{7.33}$$

These equations may be rearranged in matrix form:

$$\left[\begin{array}{c|c} (C_2 + C_4)/C_2 & C_4/C_6 \\ \hline C_4/C_2 & (C_6 + C_4)/C_6 \end{array}\right]\left[\begin{array}{c} \dot{q}_2 \\ \hline \dot{q}_6 \end{array}\right] = \left[\begin{array}{c} f_1 \\ \hline f_7 \end{array}\right]. \tag{7.34}$$

Basically, we wish to integrate Eq. (7.34) in order to relate q_2 and q_6 to q_1 and q_7, but, because of the differential causality, a matrix inversion will be required. The result is

$$\left[\begin{array}{c} \dot{q}_2 \\ \hline \dot{q}_6 \end{array}\right] = \frac{1}{C_2C_4 + C_2C_6 + C_4C_6}\left[\begin{array}{c|c} C_2C_6 + C_2C_4 & C_2C_4 \\ \hline -C_4C_6 & C_2C_6 + C_4C_6 \end{array}\right]\left[\begin{array}{c} \dot{q}_1 \\ \hline \dot{q}_7 \end{array}\right], \tag{7.35}$$

where $\dot{q}_1 = f_1, \dot{q}_7 = f_7$.

Since Eq. (7.35) is a relation between derivatives of charge variables, it can be integrated in time to yield the desired relations between the charges. Again, one must consider the charges at some initial time in order to evaluate integration

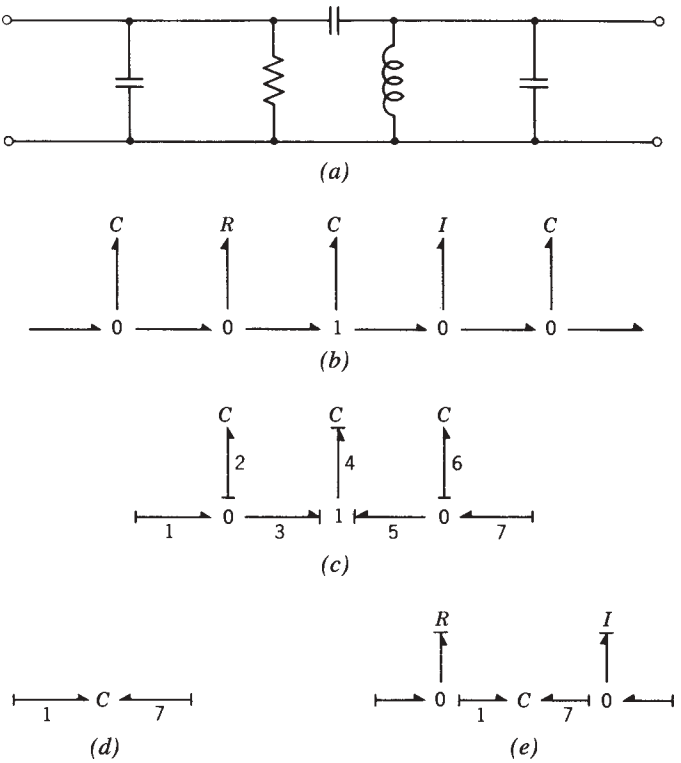


FIGURE 7.7. System containing an implicit C-field: (a) section of an electrical network; (b) bond graph; (c) implicit field identified; (d) explicit-field representation; (e) bond graph of (a) using explicit-field representation.

constants. If the system is assembled out of initially uncharged capacitors, then q_2, q_6, q_1, and q_7 can all vanish initially and the integration constants also vanish. Then Eq. (7.35) can be integrated by simply removing the dots over the qs. When this is done, a substitution into Eq. (7.31) yields the explicit-field equations:

$$\begin{bmatrix} e_1 \\ \hline e_7 \end{bmatrix} = \frac{1}{C_2C_4 + C_2C_6 + C_4C_6} \left[\begin{array}{c|c} C_6 + C_4 & -C_4 \\ \hline -C_4 & C_2 + C_4 \end{array}\right] \begin{bmatrix} q_1 \\ \hline q_7 \end{bmatrix}. \tag{7.36}$$

This relation is the constitutive law for the 2-port C-field of Figure 7.7d in integral causality form. Note that we have arranged an inward sign convention for the C-field. This means that Eq. (7.36) should have a symmetric matrix, and it does. Now, the explicit-field representation, which does not have derivative causality problems, can be used in the original system, as shown in Figure 7.7e. The same field might appear in many systems, and the explicit representation would eliminate the need to perform algebraic manipulations associated with derivative causality in each system.

Finally, it should be mentioned that it is sometimes convenient to reduce implicit fields to an explicit form, regardless of whether or not derivative causality is involved. Figure 7.8 shows two examples in which 1-ports, 2-ports, and 3-ports are connected and form implicit fields. In Figure 7.8*a*, the presence of the field is signaled by the derivative causality required on one of the —*C* elements. In Figure 7.8*b*, no derivative causality is required, but by defining an explicit field, one can reduce the number of state variables required.

Reading the bond graph of Figure 7.8*a*, one finds

$$F_4 = F_2 = \frac{a}{b}F_1 = \frac{a}{b}k_1X_1.$$

$$\dot{X}_1 = \frac{a}{b}V_2 = \frac{a}{b}(V_4 - V_3) = \frac{a}{b}\left(V_4 - \frac{d}{dt}\frac{F_3}{k_3}\right) = \frac{a}{b}\left(V_4 - \frac{d}{dt}\frac{F_2}{k_3}\right)$$

$$= \frac{a}{b}\left[V_4 - \frac{d}{dt}\left(\frac{a}{b}\frac{F_1}{k_3}\right)\right] = \frac{a}{b}\left[V_4 - \frac{d}{dt}\left(\frac{a}{b}\frac{k_1X_1}{k_3}\right)\right] \tag{7.37}$$

or

$$\dot{X}_1 = \frac{b^2k_3}{a^2k_1 + b^2k_3}\frac{a}{b}V_4$$

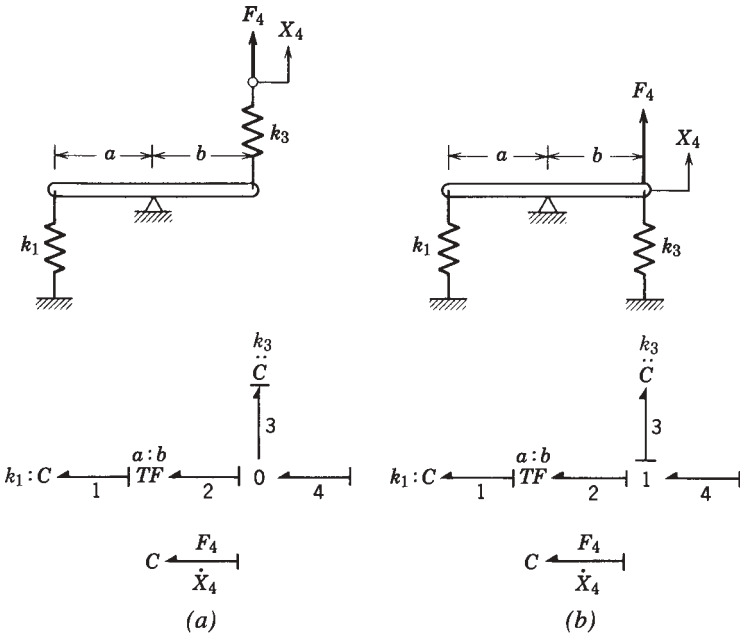


FIGURE 7.8. Two implicit fields: (*a*) a 1-port field with internal derivative causality; (*b*) 1-port field without internal derivative causality.

or

$$X_1 = \frac{b^2 k_3}{a^2 k_1 + b^2 k_3} \frac{a}{b} X_4 \tag{7.38}$$

assuming that X_1 and X_4 represent deflections away from equilibrium. Substituting Eq. (7.38) into Eq. (7.37), the characteristics of an equivalent 1-port compliance field are found:

$$F_4 = \frac{a^2 k_1 k_3}{a^2 k_1 + b^2 k_3} X_4. \tag{7.39}$$

The system of Figure 7.8*b* is easily described, since no derivative causality is involved:

$$F_4 = F_2 + F_3 = (a/b) F_1 + F_3 = (a/b) k_1 X_1 + k_3 X_3. \tag{7.40}$$

Two state equations are involved:

$$\dot{X}_1 = (a/b) V_2 = (a/b) V_4, \tag{7.41}$$

$$\dot{X}_3 = V_4, \tag{7.42}$$

but both equations can be integrated to yield

$$X_1 = (a/b) X_4, \quad X_3 = X_4, \tag{7.43}$$

again assuming that integration constants vanish, that is, that $X_1 = X_3 = X_4 = 0$ represents the condition when all springs are unstretched. Substituting Eqs. (7.43) into Eq. (7.40), the equivalent 1-port C-field constitutive law is found:

$$F_4 = \left((a^2/b^2) k_1 + k_3\right) X_4. \tag{7.44}$$

Both C-field representations (7.39) and (7.44) are practically useful, the first because the algebra associated with different causality is solved once and for all, and the second because an essentially trivial extra state equation has been integrated once and for all.

7.1.3 *I*-Fields

Inertial fields are just the duals of the capacitive fields just discussed. Instead of constitutive laws relating efforts to displacements, inertial elements have constitutive laws relating flows to momenta. All the results for C-fields will hold for I-fields if flows are substituted for efforts and momenta are substituted for displacements.

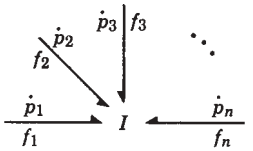


FIGURE 7.9. The symbol for an n-port I-field.

For example, the energy stored in the n-port I-field shown in Figure 7.9 is just

$$\begin{aligned}\mathbf{E} &= \int_{t_0}^{t} \sum_{i=1}^{n} f_i e_i \, dt \\ &= \int_{t_0}^{t} \sum_{i=1}^{n} f_i \dot{p}_i \, dt \\ &= \int_{\mathbf{p}_0}^{\mathbf{p}} \sum_{i=1}^{n} f_i(p) \, dpi \\ &= \int_{\mathbf{p}_0}^{\mathbf{p}} \mathbf{f}(\mathbf{p}) \, d\mathbf{p} = \mathbf{E}(\mathbf{p}),\end{aligned} \tag{7.45}$$

in which, as in Eq. (7.1), column vectors for the flows and momenta have been defined:

$$\mathbf{p} \equiv \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_n \end{bmatrix}, \quad \mathbf{f} \equiv \begin{bmatrix} f_1(\mathbf{p}) \\ f_2(\mathbf{p}) \\ f_3(\mathbf{p}) \\ \vdots \\ f_n(\mathbf{p}) \end{bmatrix}. \tag{7.46}$$

The analogs of Eq. (7.9) are

$$\frac{\partial \mathbf{E}}{\partial p_i} = f_i(\mathbf{p}), \qquad i = 1, 2, \ldots, n, \tag{7.47}$$

and the reciprocity relations are

$$\frac{\partial f_i}{\partial p_j} = \frac{\partial^2 \mathbf{E}}{\partial p_j \partial p_i} = \frac{\partial f_i}{\partial p_i}. \tag{7.48}$$

Equation (7.48), which is valid for nonlinear fields, carries the implication for the linear cases that mass matrices, inductance matrices, and the like must be symmetric because energy must be conserved.

In common with C-fields, I-fields occur in both implicit and explicit forms. Often the form that an I-field takes depends on the system modeler's point of view, and it is important to realize that manipulations of fields may result in practically useful, simplified means of representing system components. In mechanics, for example, the concept of a rigid body implies that elemental masses are constrained to move in ways such that the distances between them do not vary. This means that all rigid bodies are I-fields, and in Chapter 9, the general means of describing mechanical systems containing rigid bodies will be discussed in some detail. For now, a single example may suffice to show how an analyst could construct either an implicit- or an explicit-field representation of a rigid body.

The long, thin, rigid bar of Figure 7.10*a* has cross-sectional area A, mass density ρ, and length L. Thus, its total mass m is

$$m = \rho AL, \tag{7.49}$$

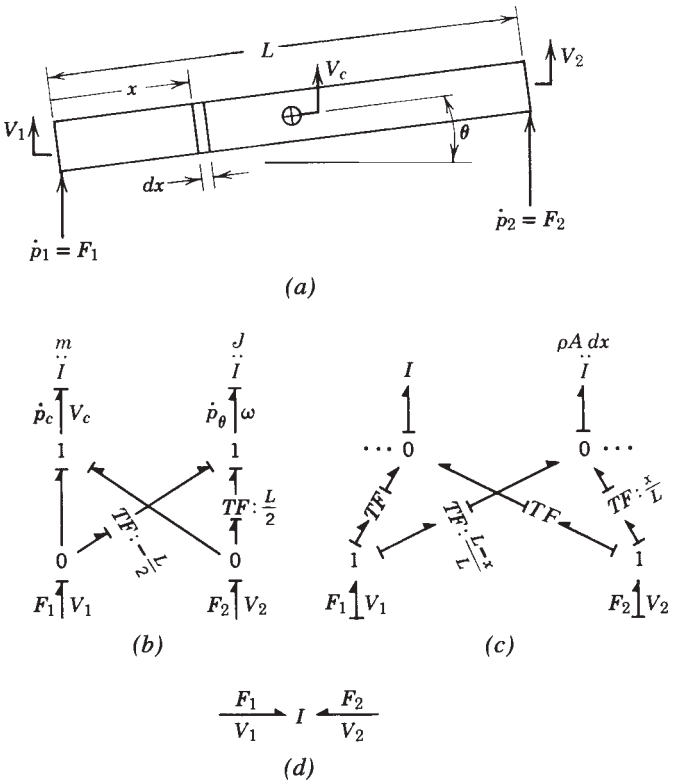


FIGURE 7.10. I-field representation of a rigid body in plane motion: (*a*) schematic diagram; (*b*) implicit field using total mass and centroidal moment of inertia; (*c*) implicit field using differential elements; (*d*) explicit field.

and its centroidal moment of inertia J about an axis perpendicular to the long dimension is

$$J = \frac{mL^2}{12}. \tag{7.50}$$

If we consider plane motion of the bar and allow only vertical motion of the center of mass and a small angular rotation relative to a horizontal axis, and if we consider two ports at the end of the bar with forces F_1 and F_2 and velocities V_1 and V_2, then this rigid body can be described by a linear 2-port I-field.

One way to find the constitutive laws for the I-field is to describe the motion first in terms of the velocity of the center of mass, V_c, and the angular velocity, $\omega = \dot{\theta}$. The net force on the bar is then the rate of change of the linear momentum, and the momentum is related to V_c by the mass $V_c = p_c/m$). Similarly, the net torque about the center of mass is the rate of change of the angular momentum, and the angular momentum is related to $\omega = \dot{\theta}$ by the moment of inertia ($\omega = p_\theta/J$). In Figure 7.10*b*, these constitutive laws are represented by the two 1-port inertial elements. The remainder of the graph serves to relate V_1 and V_2 to V_c and ω, and also to relate the net force and net torque to F_1 and F_2. (The part of the graph involving 0- and 1-junctions and transformers is a special kind of junction structure that appears frequently in mechanics and will be discussed further in Chapter 9.)

From the bond graph (or the schematic diagram) we may now write the output equations:

$$V_1 = V_c - (L/2)\omega = (1/m)p_c - (L/2J)p_\theta, \tag{7.51}$$

$$V_2 = V_c - (L/2)\omega = (1/m)p_c + (L/2J)p_\theta, \tag{7.52}$$

where p_c and p_θ designate the linear and angular momentum variables, respectively. The state equations are

$$\dot{p}_c = F_1 + F_2, \tag{7.53}$$

$$\dot{p}_\theta = -(L/2)F_1 + (L/2)F_2. \tag{7.54}$$

In order to have an explicit field at the external ports, these equations must be integrated in time. If p_c and p_θ (and, hence, V_c and ω) vanish at the initial time in the integration, then the possible integration constants vanish, and the results are

$$p_c = p_1 + p_2, \tag{7.55}$$

$$p_\theta = -(L/2)p_1 + (L/2)p_2. \tag{7.56}$$

Substituting Eqs. (7.55) and (7.56) into Eqs. (7.51) and (7.52), the result is an explicit I-field constitutive law:

$$\left[\begin{array}{c} V_1 \\ \hline V_2 \end{array}\right] = \left[\begin{array}{c|c} (1/m) + (L^2/4J) & (1/m) + (L^2/4J) \\ \hline (1/m) + (L^2/4J) & (1/m) + (L^2/4J) \end{array}\right] \left[\begin{array}{c} p_1 \\ \hline p_2 \end{array}\right]. \tag{7.57}$$

A different, more fundamental, but less convenient approach to this problem will generate an explicit I-field representation directly. We consider the bar to be a rigid massless rod on which are attached an infinite number of masses of value $\rho A\, dx$ at the generic position x. The velocity V_x of such a mass is

$$V_x = \frac{L - x}{L} V_1 + \frac{x}{L} V_2. \tag{7.58}$$

The bond graph of Figure 7.10*c* shows how these elemental masses are related to the external ports. Since the forces F_1 and F_2 are the sums of forces generated by the elemental masses, the derivative causality shown is convenient. Reading the bond graph, we have F_1 and F_2 expressed as the sum (integral) of the components due to the elemental masses:

$$F_1 = \int_0^L \frac{L - x}{L} \frac{d}{dt}[(\rho A\, dx) V x] = \int_0^L \rho A \frac{L - x}{L} \left(\frac{L - x}{L} \dot{V}_1 + \frac{x}{L} \dot{V}_2 \right) dx, \tag{7.59}$$

$$F_2 = \int_0^L \frac{x}{L} \frac{d}{dt}[(\rho A\, dx) V x] = \int_0^L \rho A \frac{x}{L} \left(\frac{L - x}{L} \dot{V}_1 + \frac{x}{L} \dot{V}_2 \right) dx, \tag{7.60}$$

where Eq. (7.58) has been used. The result of the integration in x is

$$\left[\begin{array}{c} F_1 \\ \hline F_2 \end{array}\right] = \left[\begin{array}{c} \dot{p}_1 \\ \hline \dot{p}_2 \end{array}\right] = \frac{\rho A L}{6} \left[\begin{array}{c|c} 2 & 1 \\ \hline 1 & 2 \end{array}\right] \left[\begin{array}{c} \dot{V}_1 \\ \hline \dot{V}_2 \end{array}\right], \tag{7.61}$$

or, if we agree to define velocities such that p_1 and p_2 vanish when V_1 and V_2 vanish, then

$$\left[\begin{array}{c} p_1 \\ \hline p_2 \end{array}\right] = \frac{\rho A L}{6} \left[\begin{array}{c|c} 2 & 1 \\ \hline 1 & 2 \end{array}\right] \left[\begin{array}{c} V_1 \\ \hline V_2 \end{array}\right]. \tag{7.62}$$

This is the constitutive law for the explicit field of Figure 7.10*d* in derivative causality (or mass matrix) form. It is left as an exercise for the reader to show that, using Eqs. (7.49) and (7.50), the constitutive laws of Eqs. (7.57) and (7.62) are the same except for the different causality.

Electrical circuits containing mutually interacting coils can be conveniently represented using explicit I-fields. In order to write the constitutive laws for such fields, one requires not only a sign convention for currents and voltages,

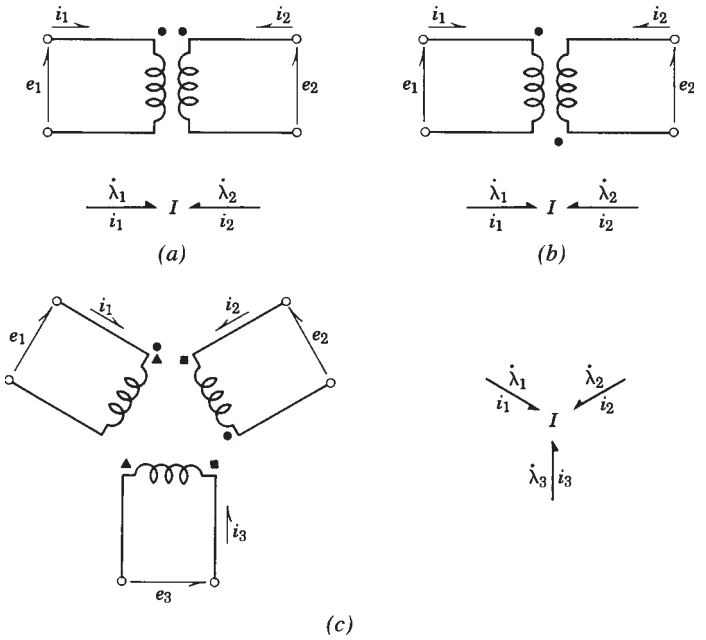


FIGURE 7.11. Mutual inductance in electrical systems: (*a*,*b*) 2-port I-fields with different coil orientation; (*c*) 3-port I-field.

but also a convention dealing with the relative orientation of the coils. One such convention is shown in the circuit diagrams of Figure 7.11*a* and *b* as dots placed near the ends of the coils. The idea is that if the currents, i_1 and i_2, are defined so that when positive, they both enter or both leave dotted ends of their coils, then the mutual-inductance effects will be positive. If, on the other hand, the dots are placed as in Figure 7.11*b*, so that one of the currents enters through a dotted end and the other leaves through a dotted end, then the mutual-inductance effects are negative. The convention is most readily illustrated for the linear case, but it applies equally well for the nonlinear case. For example, if one denotes self-inductance coefficients by the letter L and mutual-inductance coefficients by the letter M, as is the usual convention, then the system of Figure 7.11*a* has the constitutive law

$$\left[\begin{array}{c}\lambda_1\\ \hline \lambda_2\end{array}\right] = \left[\begin{array}{c|c} L_1 & M_{12} \\ \hline M_{12} & L_2 \end{array}\right] \left[\begin{array}{c} i_1 \\ \hline i_2 \end{array}\right]. \tag{7.63}$$

Note that this field representation is in derivative causality form and is symmetric, as it must be. The field of Figure 7.11*b* has the constitutive law

$$\left[\begin{array}{c}\lambda_1\\ \hline \lambda_2\end{array}\right] = \left[\begin{array}{c|c} L_1 & -M_{12} \\ \hline -M_{12} & L_2 \end{array}\right] \left[\begin{array}{c} i_1 \\ \hline i_2 \end{array}\right]. \tag{7.64}$$

For three or more interacting coils, it is rather awkward to show the orientations. In Figure 7.11*c*, the dots have become circles, squares, and triangles to indicate the signs of the three mutual-inductance coefficients. The field laws are readily written:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \left[\begin{array}{c|c|c} +L_1 & -M_{12} & -M_{13} \\ \hline -M_{12} & +L_2 & +M_{23} \\ \hline -M_{13} & +M_{23} & +L_3 \end{array}\right] \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix}. \tag{7.65}$$

In the bond graph I-field representation, the convention indicated by the circles, squares, and triangles is evident in the sign pattern of the mutual-inductance coefficients. Note that with the inward sign convention shown, the self-inductance terms must be positive for physical coils. Also, independently of the coil orientation convention, the inductance matrix and its inverse will be symmetric if the I-field is conservative.

When energy-storing fields are used in systems, it sometimes happens that explicit representations such as Eq. (7.65) must be algebraically manipulated. For example, if the coils of Figure 7.11*c* are interconnected, as shown in Figure 7.12*a*, then the rules of causality indicate that integral causality cannot be applied to all three I-field ports. If there were no mutual inductance, the system would be as shown in Figure 7.12*b*. This system has one element in differential causality, but

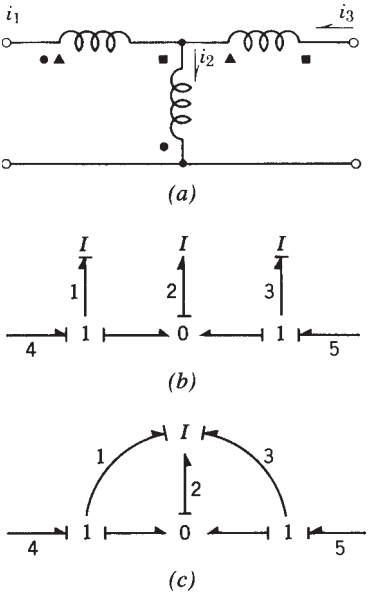


FIGURE 7.12. A network of interacting coils: (*a*) the coils of Figure 7.11 interconnected; (*b*) system bond graph if mutual inductance is absent; (*c*) system bond graph with I-field.

the formulation methods of Chapter 5 would handle the system without trouble. In fact, this system is just the dual of the C-field of Figure 7.7. On the other hand, when mutual-inductance effects are considered, the system appears as shown in Figure 7.12c. If one assigns integral causality on bonds 1 and 3, then bond 2 must have derivative causality.

It is not particularly difficult to switch the constitutive law of Eq. (7.65) into the mixed causal form of Figure 7.12c. Rewriting parts of Eq. (7.65), we have

$$\left[\begin{array}{c} \lambda_1 \\ \hline \lambda_3 \end{array}\right] = \left[\begin{array}{c|c} L_1 & -M_{13} \\ \hline -M_{13} & L_3 \end{array}\right] \left[\begin{array}{c} i_1 \\ \hline i_3 \end{array}\right] + \left[\begin{array}{c} -M_{12} \\ \hline M_{23} \end{array}\right] [i_2]. \tag{7.66}$$

Inverting this relation to yield i_1 and i_3 in terms of λ_1, λ_3, and i_2 gives

$$\left[\begin{array}{c} i_2 \\ \hline i_3 \end{array}\right] = \frac{1}{L_1 L_3 - M_{13}^2} \left[\begin{array}{c|c|c} L_3 & M_{13} & L_3 M_{12} - M_{13} M_{23} \\ \hline M_{13} & L_1 & M_{13} M_{12} - L_1 M_{23} \end{array}\right] \left[\begin{array}{c} \lambda_1 \\ \hline \lambda 3 \\ \hline i_2 \end{array}\right]. \tag{7.67}$$

Substituting Eq. (7.67) into the second equation of (7.65), we can also obtain the third equations for λ_2 in terms of λ_1, λ_3, and i_2:

$$\lambda_2 \equiv \lambda_2 (\lambda_1, \lambda_2, i_2) . \tag{7.68}$$

When mutual-inductance effects are present, the writing of state equations is not quite as straightforward as when there is no mutual inductance. Basically, this is because of the role of i_2 in Eq. (7.67). We start with the equation for λ_1,

$$\dot{\lambda}_1 = e_4 - e_2 = e_4 - \frac{d}{dt} \lambda_2 (\lambda_1, \lambda_3, i_2). \tag{7.69}$$

Now,

$$i_2 = i_1 + i_3, \tag{7.70}$$

and we can use Eq. (7.67) to find i_1 and i_3 (and hence i_2) in terms of λ_1, λ_3, and i_2. But now we are in an algebraic loop reminiscent of those discussed Chapter 5 when the imposition of integral causality did not suffice to determine the causality of all bonds. Using Eqs. (7.70) and (7.67), we can express i_2 in terms of itself and then solve this algebraic equation for i_2 in terms of λ_1, λ_3 in order to compute the state equation (7.69). Failure to do this will result in an endless looping through the equations in an attempt to eliminate zi_2.

Although one could write state equations for the system in the manner outlined above, there is a simpler way to handle the system. In Figure 7.13, it is shown that the imposition of complete differential causality (which corresponds to Eq. (7.65) anyway) allows one to compute the explicit 2-port field

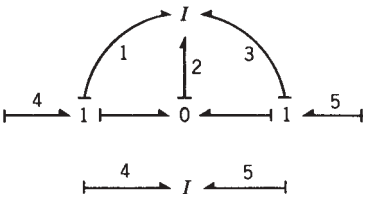


FIGURE 7.13. Reduction of the system of Figure 7.12 to explicit 2-port I-field form.

without solving any algebraic equations. Reading the graph, and using Eq. (7.65), we have

$$\begin{aligned} e_4 = e_1 + e_2 &= \frac{d}{dt}(\lambda_1 + \lambda_2) \\ &= \frac{d}{dt}[(L_1 - M_{12})i_1 + (L_2 - M_{12})i_2 + (M_{23} - M_{13})i_3] \\ &= \frac{d}{dt}[(L_1 - M_{12})i_4 + (L_2 - M_{12})(i_4 + i_5) + (M_{23} - M_{13})i_5], \end{aligned}$$

or

$$\lambda_4 = (L_1 + L_2 - 2M_{12})i_4 + (L_2 - M_{12} - M_{13} + M_{23})i_5 \tag{7.71}$$

upon integration. Similarly,

$$\lambda_5 = (L_2 - M_{12} - M_{13} + M_{23})i_4 + (L_2 + L_3 + 2M_{23})i_5. \tag{7.72}$$

These equations represent the constitutive laws for the 2-port I-field in derivative causality form. The all-integral-causality form follows by inverting the equations. Thus, λ_4 and λ_5 can readily be used as the two independent state variables. Clearly, the 2-port I-field is much more convenient to use in equation formulation than the 3-port field and its associated junction structure. Such studies of fields are, of course, as useful for C-field problems as they are for I-fields.

7.1.4 Mixed Energy-Storing Fields

There are occasions when an energy-storing device cannot be described as a C-field or an I-field but rather acts as a C-field at some ports and an I-field at others. Within a single-energy domain, the need for such an element is far from evident, but when transducers are studied in Chapter 8, it will be seen that many transducers are essentially energy conservative but are not pure I-fields or C-fields. Even in the case of transducers, it has been a common practice to arrange the analogies between variables in the two energy domains linked by the transducer in such a way that the transducer is a pure field. There is no real need to do this,

however, and in bond graphs we prefer to retain our identification of effort, flow, displacement, and momentum quantities for all energy domains. We need, then, to discuss what will be called *IC*-fields. Not only is this policy just as convenient as the policy of switching analogies to suit the problem at hand; it is required in principle for some systems. There is no way, for example, to pick an effort–flow identification for an electromechanical system containing both a movable-plate capacitor and a solenoid transducer such that both transducers would be described as pure *I*- or *C*-fields.

Figure 7.14 shows the general *n*-port *IC*-field. The ports are numbered so that the first *j*-ports ($1 \leq j \leq n$) are inertial in character and the ports from $j+1$ to n are capacitive in character. In integral causality form as shown, the first j state variables are momenta and the remainder are displacements. The stored energy, **E**, is a function of this mixed set of state variables:

$$\mathbf{E} = \int^{t} \sum_{i=1}^{n} e_i f_i \, dt = \int^{\mathbf{p}} \sum_{i=1}^{j} f_i \, dp_i + \int^{\mathbf{q}} \sum_{k=j+1}^{n} e_k \, dq_k, \tag{7.73}$$

where **p** represents the vector of momenta and **q** represents the vector of displacements. From Eq. (7.73), we see that

$$f_i = \frac{\partial \mathbf{E}}{\partial P_i}, \quad i = 1, 2, \ldots, j, \tag{7.74}$$

$$e_k = \frac{\partial \mathbf{E}}{\partial p_k}, \quad k = j+1, j+2, \ldots, n. \tag{7.75}$$

The reciprocity conditions for the constitutive laws can be easily derived by computing second partial derivatives of **E**. For example, the second partial derivatives of **E** with respect to a momentum from the first set of ports and a displacement from the second set of ports yields

$$\frac{\partial f_i}{\partial q_k} = \frac{\partial^2 E}{\partial q_k \partial p_i} = \frac{e_k}{\partial p_i}, \tag{7.76}$$

where $1 \leq i \leq j$, $j+1 \leq k \leq n$. These results should be compared with the results for pure *C*-fields and *I*-fields (Eqs. (7.1), (7.10), (7.45), and (7.46)). Physical examples involving *IC*-fields will be found in Chapter 8.

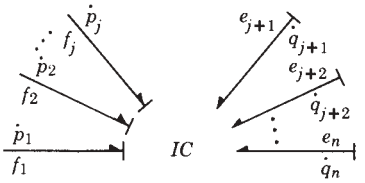


FIGURE 7.14. The general mixed energy-storage field.

7.2 RESISTIVE FIELDS

An R-field is an n-port, the constitutive laws of which relate the n-port efforts and the n-port flows by means of static (or algebraic) functions. This definition includes power-conservative elements such as 0- and 1-junctions and elements containing sources as R-fields, but in practice most R-fields studied dissipate power. Both explicit and implicit R-fields exist. Explicit multiport R-fields arise frequently in modeling nonlinear devices. Implicit fields arising from the interconnection of 1-port resistances, transformers, gyrators, and junction 3-ports may be conveniently represented in R-field form as relations between external port efforts and flows.

The causality for R-fields usually is determined by the source and energy-storing elements in the system. Two fundamental causal patterns are shown in Figure 7.15. The constitutive laws for the *resistance* causality shown in Figure 7.15*a* may be written

$$e_i = \Phi_{R_i}(f_1, f_2, \ldots, f_n), \quad i = 1, 2, \ldots, n. \tag{7.77}$$

The *conductance* causality shown in Figure 7.15*b* has constitutive laws that may be represented by the following:

$$f_i = \Phi_{R_i}^{-1}(e_1, e_2, \ldots, e_n), \quad i = 1, 2, \ldots, n. \tag{7.78}$$

In addition to these two fundamental causal forms, a large number of other forms are possible in which some of the bonds are causally oriented as in Figure 7.15*a* and the remainder are oriented as in Figure 7.15*b*. Although R-fields are basically neutral with respect to causality, we shall encounter some constitutive laws for R-fields that cannot be unique for certain causalities (this happens, of course, only in the nonlinear case).

Although there is no stored energy function for R-fields as there is for energy-storing fields, and hence there is no simple way to show that R-field constitutive laws are constrained, still, for special classes of R-fields, some useful properties of the constitutive laws may be found. We will illustrate some of these properties by example.

First, let us consider linear R-fields which contain no sources and no gyrators. A typical example is shown in Figure 7.16. This R-field will accept resistance

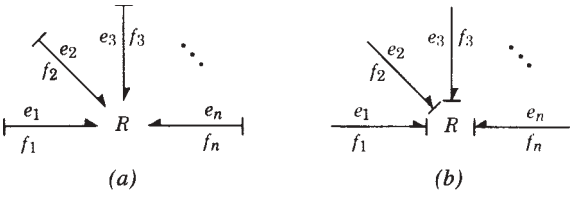


FIGURE 7.15. The general R-field: (a) resistance form; (b) conductance form.

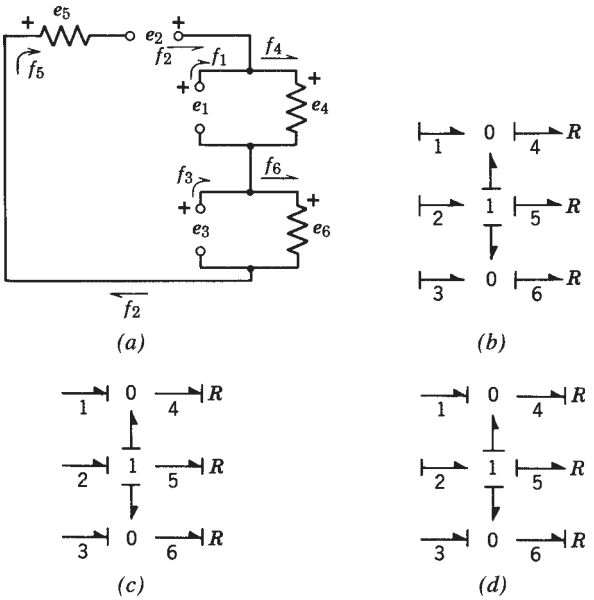


FIGURE 7.16. An implicit R-field: (a) circuit graph; (b) bond graph showing resistance causality; (c) bond graph showing conductance causality; (d) bond graph showing mixed causality.

causality on the three external ports, as shown in Figure 7.16b. Reading the bond graph, the constitutive laws are readily derived:

$$\begin{bmatrix} e_1 \\ \hline e_2 \\ \hline e_3 \end{bmatrix} = \left[\begin{array}{c|c|c} R_4 & R_4 & 0 \\ \hline R_4 & R_4 + R_5 + R_6 & R_6 \\ \hline 0 & R_6 & R_6 \end{array}\right] \begin{bmatrix} f_1 \\ \hline f_2 \\ \hline f_3 \end{bmatrix}. \tag{7.79}$$

As the bond graph of Figure 7.16c demonstrates, this R-field will also accept conductance causality. So, by inverting Eq. (7.79) or, more simply, by reading the bond graph of Figure 7.16c, the constitutive laws in conductance form may be found:

$$\begin{bmatrix} f_1 \\ \hline f_2 \\ \hline f_3 \end{bmatrix} = \left[\begin{array}{c|c|c} 1/R_4 + 1/R_5 & -1/R_5 & 1/R_5 \\ \hline -1/R_5 & 1/R_5 & -1/R_5 \\ \hline 1/R_5 & -1/R_5 & 1/R_5 + 1/R_6 \end{array}\right] \begin{bmatrix} e_1 \\ \hline e_2 \\ \hline e_3 \end{bmatrix}. \tag{7.80}$$

Notice that both Eqs. (7.79) and (7.80) are symmetric. These forms may be called *Onsager forms* in analogy to the well-known Onsager reciprocal relations of irreversible thermodynamics [3]. Onsager proposed his reciprocity conditions for variables called "affinities" and "fluxes" that are analogous to the efforts and

flows of bond graphs. In general, implicit R-fields composed of linear 1-port resistances, 0- and 1-junctions, and transformers will obey Onsager reciprocity; that is, when expressed in resistance or conductance form, the matrix in the constitutive laws will be symmetric. On the other hand, for mixed causality, the matrix will have antisymmetrical terms, and if gyrators are present, the Onsager reciprocity relations do not hold. Thus, the reciprocity of energy-storing fields, which may be called *Maxwell reciprocity*, is more general than *Onsager reciprocity*. A perfectly reasonable explicit R-field characterization expressed in resistance or conductance form can be unsymmetrical. One may think of the unsymmetrical part of the matrices as arising from gyrational effects that introduce antisymmetrical terms, although in an explicit field one cannot identify a gyrator unless one can find some implicit field that has the same port constitutive laws as the explicit field.

Using our example, let us first show that an R-field in mixed causality has antisymmetric terms. A possible mixed causal pattern is shown in Figure 7.16*d*. Reading the bond graph, the constitutive laws are as follows:

$$\left[\begin{array}{c} f_1 \\ \hline e_2 \\ \hline f_3 \end{array}\right] = \left[\begin{array}{c|c|c} 1/R_4 & -1 & 0 \\ \hline 1 & R_5 & 1 \\ \hline 0 & -1 & 1/R_6 \end{array}\right] \left[\begin{array}{c} e_1 \\ \hline f_2 \\ \hline e_3 \end{array}\right], \tag{7.81}$$

in which the antisymmetric terms are evident. Such a form, which results when some of the input and output variables are interchanged for an R-field that obeys the Onsager reciprocal relations in resistance or conductance causality, is sometimes called a *Casimir form* [4].

It may be tempting to conclude that if R-fields are described with either resistance or conductance causality, they will exhibit Onsager reciprocity, and if they are described with a mixed causality, they will show a Casimir form. But it is easy to show that if gyrators are allowed in an implicit R-field, then this conclusion is false. Consider, for example, the R-field of Figure 7.17, which is formed from the field of Figure 7.16 by adding a single gyrator. In the conductance causality shown in Figure 7.17*a*, the field exhibits a Casimir form rather than an Onsager form:

$$\left[\begin{array}{c} f_1 \\ \hline f_2 \\ \hline f_3 \end{array}\right] = \left[\begin{array}{c|c|c} 1/R_4 & -1/r & 0 \\ \hline 1/r & R_5/r^2 & 1/r \\ \hline 0 & -1/r & 1/R_6 \end{array}\right] \left[\begin{array}{c} e_1 \\ \hline e_2 \\ \hline e_3 \end{array}\right], \tag{7.82}$$

in which r is the gyrator parameter. Similarly, for the mixed causality of Figure 7.17*b*, the R-field is symmetric:

$$\left[\begin{array}{c} f_1 \\ \hline e_2 \\ \hline f_3 \end{array}\right] = \left[\begin{array}{c|c|c} 1/R_4 + 1/R_5 & -r/R_5 & 1/R_5 \\ \hline -r/R_5 & r^2/R_5 & -r/R_5 \\ \hline 1/R_5 & -r/R_5 & 1/R_5 + 1/R_6 \end{array}\right] \left[\begin{array}{c} e_1 \\ \hline f_2 \\ \hline e_3 \end{array}\right]. \tag{7.83}$$

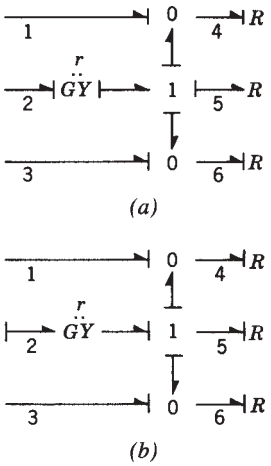


FIGURE 7.17. Modified version of the *R*-field of Figure 7.15: (*a*) conductance causality; (*b*) mixed causality.

A general *R*-field is not necessarily representable in either an Onsager or a Casimir form. A simple example is shown in Figure 7.18. Its constitutive laws in resistance form are

$$\left[\frac{e_1}{e_2}\right] = \left[\begin{array}{c|c} R_3 + R_4 & r + R_4 \\ \hline -r + R_4 & R_4 + R_5 \end{array}\right] \left[\frac{f_1}{f_2}\right], \tag{7.84}$$

in which a symmetrical part due to the elements that obey Onsager reciprocity and an antisymmetrical part due to the gyrator can be recognized. If Eq. (7.84) were presented with numerical values for the parameters, it would have a rather undistinguished appearance, being neither an Onsager nor a Casimir form, i.e., neither symmetric nor having antisymmetric terms.

Note that if positive resistance parameters are used, then all of the example fields will dissipate power for any possible port conditions. This is true because the 1-port resistors can only dissipate power, and the junction elements and the 2-ports, —*TF*— and —*GY*—, conserve power. A way to check for power

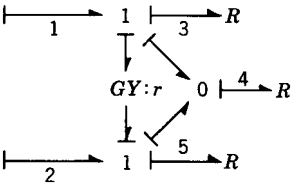


FIGURE 7.18. An *R*-field containing a gyrator.

dissipation given an explicit linear field is to check whether the matrix of coefficients is positive definite or not. In the resistance form, for example, if **e** is the column vector of efforts, **f** the column vector of flows, and **R** the resistance matrix, then the power **P** is

$$\mathbf{P} = \mathbf{f}^{\mathrm{t}}\mathbf{e},$$

and the constitutive law for the field is

$$\mathbf{e} = \mathbf{R}\mathbf{f},$$

so that

$$\mathbf{P} = \mathbf{f}^{\mathrm{t}}\mathbf{R}\mathbf{f}, \tag{7.85}$$

where t stands for the transpose of a vector or matrix.

Equation (7.85) shows that power will be dissipated for any **f** if **R** is positive definite. If one can only say that no power should be created, then **R** must be positive semidefinite; that is, there may be some finite flows for which zero power is dissipated. Using the rules for checking a matrix for positive definiteness, the reader may demonstrate that all the examples of *R*-fields given above are at least positive semidefinite. In this regard, it is useful to remember that the antisymmetric part of any matrix contributes nothing to the definiteness of the matrix. This correlates with the idea that the antisymmetric terms due to the presence of a gyrator represent no power generation or dissipation in a field such as that associated with Eq. (7.84).

7.3 MODULATED 2-PORT ELEMENTS

The 2-port elements —*TF*— and —*GY*— are linear, power-conserving elements, the usefulness of which has been demonstrated in previous chapters. Here we discuss the modulated transformer, —*MTF*—, and modulated gyrator, —*MGY*—, which are nonlinear, power-conserving generalizations of —*TF*— and —*GY*—. Basically, the parameters of the elements —*TF*— and —*GY*— are allowed to be a function of some parameter, say ξ, in the modulated 2-ports. The usual symbols and constitutive laws or these elements are shown below:

$$\begin{array}{cc} m(\xi) & r(\xi) \\ \downarrow & \downarrow \\ \underset{1}{\rightharpoonup} MTF \underset{2}{\rightharpoonup}, & \underset{1}{\rightharpoonup} MGY \underset{2}{\rightharpoonup}, \\ m(\xi)e_1 = e_2, & e_1 = r(\xi)f_2, \\ f_1 = m(\xi)f_2, & r(\xi)f_1 = e_2. \end{array} \tag{7.86}$$

Notice that for both elements the power $e_1 f_1$ is always equal to the power $e_2 f_2$ no matter what the value of $m(\xi)$ or $r(\xi)$ may be. Also, the parameters m

and r are shown as changing by means of a signal or active bond rather than by means of a power bond. Thus, it is characteristic of these modulated-parameter elements that the parameters change value without a directly associated power flow.

The modulated elements will prove useful for modeling certain classes of systems in Chapters 8 and 9. In Figure 7.19, two examples are shown that illustrate typical uses of the modulated elements. The modulated transformer is particularly useful in describing mechanical systems moving though large angles. The rigid, pivoted bar of Figure 7.19a may be described thus:

$$\tau = (l \cos\theta) F, \tag{7.87}$$

$$(l \cos\theta)\omega = V, \tag{7.88}$$

where the variables are defined in the figure.

These constitutive relations may be derived by applying the laws of mechanics to the linkage. Equation (7.87) is a torque equilibrium statement, and

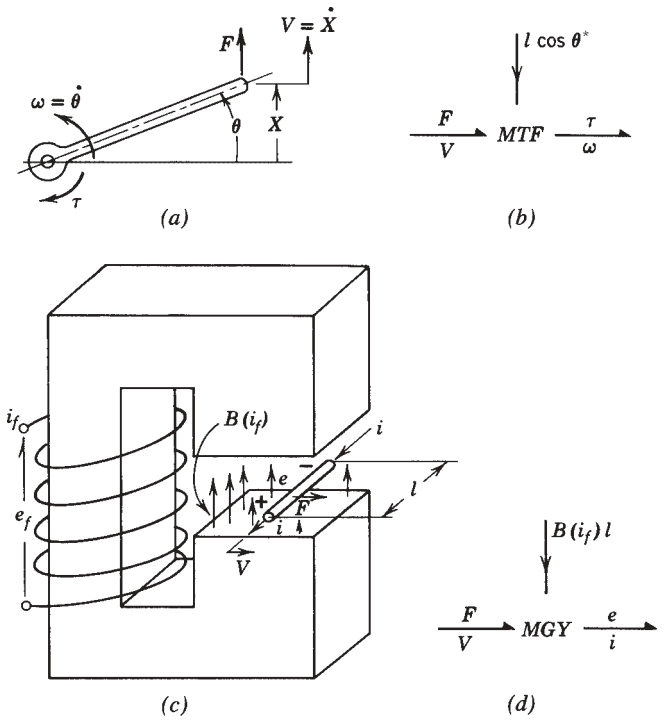


FIGURE 7.19. Examples of modulated 2-port elements: (a) mechanical linkage; (b) modulated transformer for (a); (c) electromechanical system; (d) modulated gyrator for (c). Note that the modulus for (b) is defined unambiguously by Eqs. (7.87) and (7.88).

Eq. (7.88) is an angular velocity–velocity relation which could be derived by time-differentiating the equation

$$l \sin \theta = x. \tag{7.89}$$

The bond graph of Figure 7.19*b* represents Eqs. (7.87) and (7.88), where the transformer modulus is $l \cos \theta$ and θ plays the role of ξ in the general equations (7.86). It is a peculiarity of the *MTF*s of mechanics that the modulus varies with the displacements at the ports. For this reason, these *MTF*s are sometimes called *displacement-modulated transformers*.

Figure 7.19*c* shows an electromechanical system in which a field current i_f is responsible for establishing a magnetic field $B(i_f)$ in a gap. If then a current-carrying conductor of length l moves with velocity V, as shown, it is a consequence of Faraday's law that a voltage e will be induced along the conductor according to the following relation:

$$e = B(i_f) l V. \tag{7.90}$$

The force F required to move the conductor is found from the Lorentz force law to be

$$B(i_f) l i = F. \tag{7.91}$$

The modulated gyrator of Figure 7.19*d* represents these laws, with $B(i_f)l$ representing the modulus r and i_f the parameter ξ in the general form of Eq. (7.86). Note that $ei = FV$ (as long as the electrical and the mechanical power are measured in the same units), and i_f can be changed with no power associated with the *MGY*. On the other hand, i_f is also associated with a self-inductance effect and coil resistance, so that there is power flow and energy storage associated with i_f. The point is that no energy is associated with changes in Bl from the point of view of the *MGY*. The modulated gyrator will be used in modeling voice coils, electrical motors, and similar devices in Chapter 8.

Since the sign conventions and causal restrictions for the modulated elements are the same as for the —*TF*— and —*GY*— elements studied earlier, there is little new to be said here. A few words of caution are in order, however. Since the modulated elements incorporate a pure signal interaction, it is quite possible to construct bond graphs that do not have physical interpretations by imagining that the modulus of a transformer or gyrator can be a function of any variable at all. Just as it is easy to make incorrect block diagrams, signal flow graphs, or other signal descriptions that violate power and energy constraints that exist in real physical systems, so too is it rather easy to assume that an element can be modeled using a modulated 2-port when, in fact, the element may really be a true 3-port or when the physical system may not allow the modulus to be a function of the signal assumed.

To show just one example of the trap into which one may sometimes fall, consider the rack-and-pinion system shown in Figure 7.20. Clearly, if the pinion is small ($r << l$), then the rack is nearly a lever with a lever ratio of

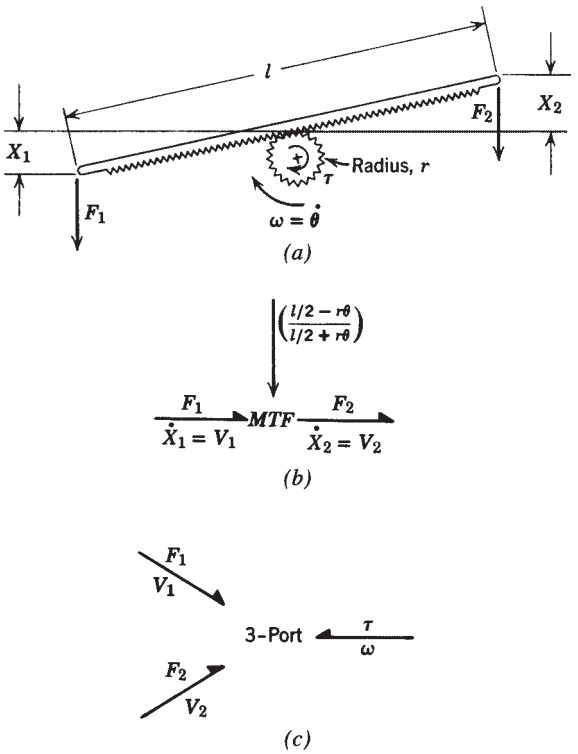


FIGURE 7.20. Rack-and-pinion system: (*a*) schematic diagram; (*b*) incorrect bond graph; (*c*) correct bond graph.

$(l/2 - r\theta)/(l/2 + r\theta)$ if the rack is centered when $\theta = 0$. Furthermore, if we rotate the pinion to some position and then hold θ fixed, the rack will function almost as a —*TF*— with the transformer modulus set by the value of θ. Thus, it seems reasonable to represent the system as in Figure 7.20*b*. However, this representation is false. The reason is that, whenever F_1 and F_2 are not zero, θ cannot be changed with no power. In fact, the torque τ is proportional to F_1 or F_2. Thus, the device is a true 3-port as shown in Figure 7.20*c*. It can be modeled using *MTF* elements and 0- and 1-junctions as a multiport—modulated transformer, as will be shown in Chapter 9, but the representation of Figure 7.20*b* is fundamentally incorrect.

7.4 JUNCTION STRUCTURES

Junction structures, which are assemblages of 0- and 1-junctions, transformers, and gyrators, are the energy switchyards that enforce the constraints among parts of dynamic systems. No power is dissipated or generated in a junction structure,

so the *net* power into a junction structure at its ports always vanishes. Since junction structures do provide relations between efforts and flows at their ports, they are special types of R-fields that never dissipate power. One might expect, for linear junction structures, to find the Onsager and Casimir forms that were found for general R-fields. These forms are indeed found, but junction structures typically cannot accept all possible causal assignments at their ports, so some forms for the constitutive laws simply cannot exist.

Some example junction structures are shown in Figure 7.21. Note that all these structures are shown with an inward sign convention, so that the sum of the port powers must vanish. This means that, in each case, the matrix relating inputs to outputs must be antisymmetric; that is, zeros must appear on the main diagonal, and the ijth component must be the negative of the jith component.

For systems without gyrators and in mixed resistance–conductance causality, such as Figures 7.21*a*, *c*, and *d*, the relations are Casimir forms (but with zeros on the main diagonal). The system of Figure 7.20*b* is in conductance form but, because it contains a gyrator, does not obey Onsager reciprocity and is instead antireciprocal. Figure 7.21*e* is antisymmetric partly because of mixed causality (which contributes two antisymmetric terms) and partly because of the gyrator (which contributes the other two).

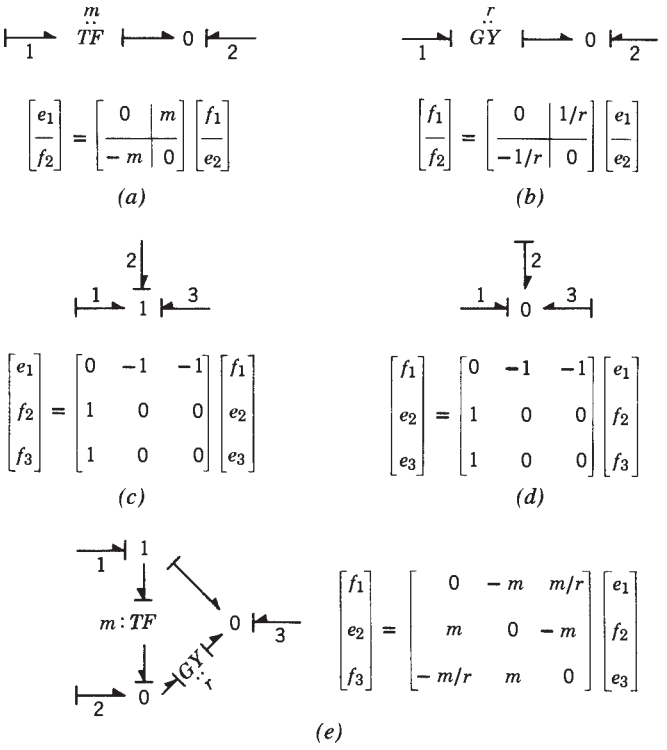


FIGURE 7.21. Constitutive laws for some simple junction structures.

Since an R-field that contains no gyrators must obey Onsager reciprocity in resistance or conductance form, and any junction structure must have antisymmetric constitutive laws, we conclude that junction structures without gyrators cannot accept resistance or conductance causality on all ports.

7.5 MULTIPORT TRANSFORMERS

Another interesting viewpoint about junction structures involves the transformations of variables that they provide. Consider, for example, the junction structure of Figure 7.22. It transforms the variables e_3 and e_4 into the variables e_1 and e_2:

$$\left[\begin{array}{c} e_1 \\ \hline e_2 \end{array}\right] = \left[\begin{array}{c|c} -1 & -1 \\ \hline -m_1 & -m_2 \end{array}\right] \left[\begin{array}{c} e_3 \\ \hline e_4 \end{array}\right]. \tag{7.92}$$

But this transformation of efforts is accompanied by a transformation of flows:

$$\left[\begin{array}{c} f_3 \\ \hline f_4 \end{array}\right] = \left[\begin{array}{c|c} 1 & m_1 \\ \hline 1 & m_2 \end{array}\right] \left[\begin{array}{c} f_1 \\ \hline f_2 \end{array}\right], \tag{7.93}$$

and, since the junction structure is accomplishing the transformations, the transformations must conserve power. Rearranging the equations, the antisymmetric form characteristic of power conservation appears:

$$\left[\begin{array}{c} e_1 \\ \hline e_2 \\ \hline f_3 \\ \hline f_4 \end{array}\right] = \left[\begin{array}{c|c|c|c} 0 & 0 & -1 & -1 \\ \hline 0 & 0 & -m_1 & -m_2 \\ \hline 1 & m_1 & 0 & 0 \\ \hline 1 & m_2 & 0 & 0 \end{array}\right] \left[\begin{array}{c} f_1 \\ \hline f_2 \\ \hline e_3 \\ \hline e_3 \end{array}\right]. \tag{7.94}$$

Such a structure is usefully considered to be a multiport generalization of a 2-port transformer. It is appropriate to change from an "all inward" sign convention which has been used in studying general junction structures to a "through" sign convention for the external ports, as shown in Figure 7.23. With this type of

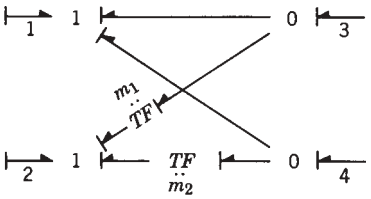


FIGURE 7.22. A junction structure that is a 2 × 2 bilateral transformation of variables.

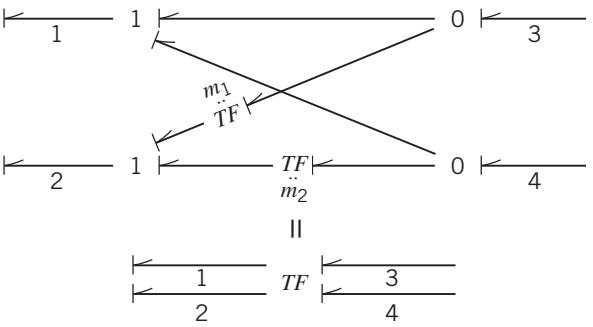


FIGURE 7.23. Structure of Figure 7.22 with modified sign half-arrows at ports 1 and 2.

sign convention, the signs in Eq. (7.92) change and the transformation equations can be written in a form analogous to the form used for 2-port transformers:

$$\left[\frac{e_1}{e_2}\right] = \left[\begin{array}{c|c} 1 & 1 \\ \hline m_1 & m_2 \end{array}\right] \left[\frac{e_3}{e_4}\right], \tag{7.92a}$$

$$\left[\begin{array}{c|c} 1 & m_1 \\ \hline 1 & m_2 \end{array}\right] \left[\frac{f_1}{f_2}\right] = \left[\frac{f_3}{f_4}\right]. \tag{7.93a}$$

In this example, it may be seen that the matrix that relates the efforts is simply transposed to form the matrix that transforms the flows. In general, a multiport transformer is characterized by a matrix (and its transpose), just as a 2-port transformer is characterized by a modulus that appears in both the effort and the flow relations. (One can even think of the modulus of a 2-port *TF* as a 1×1 matrix which is equal to its transpose.)

Although this example involves a 2×2 matrix since there are two ports on either side of the transformer, there can be n ports on one side of the transformer and m ports on the other side. Then the matrix will be $n \times m$ and its transpose will be $m \times n$.

It is easy to prove that transformations involving a matrix and its transpose imply that power flowing into one side of a multiport transformer must equal the power flowing out of the other side [7]. Let $[e_1]$, $[f_1]$ be column vectors of efforts and flows on one side of a multiport transformer and $[e_2]$, $[f_2]$ be vectors of efforts and flows on the other side. Then let $[M]$ be a transformation matrix of appropriate dimensions so that the equations analogous to Eqs. (7.92a) and (7.93a) are

$$[e_1] = [M]\,[e_2]\,, \tag{7.95}$$

$$\left[M^t\right]\left[f_1\right] = \left[f_2\right]. \tag{7.96}$$

By transposing Eq. (7.96) and postmultiplying both sides by $[e_2]$, the result is

$$\left[f_1\right]^t [M]\,[e_2] = \left[f_2\right]^t [e_2]\,, \tag{7.97}$$

or, upon using Eq. (7.95),

$$\left[f_1\right]^t [e_1] = \left[f_2\right]^t [e_2], \tag{7.98}$$

which simply states that the sum of the port powers on the number 1 side equals the sum of the port powers on the number 2 side. With the new sign convention this means the net power flowing in on one side equals the power flowing out of the other side.

It is interesting that in the above proof, not only can there be any number of ports on the two sides of the transformer, but there is no requirement that the elements of the matrix be constant. Thus, the considerations apply not only to multiport transformers with constant matrices, which occur in many types of systems, but also for multiport-modulated transformers in which the elements of the matrix change with time. An important application of multiport-modulated transformers is in mechanics, where geometric nonlinearities often arise. This topic is explored in depth in Chapter 9, but here we merely point out that when geometric constraints or transformations among displacements are converted to velocity relations, they always can be represented by multiport-modulated transformers or transformer junction structures. Furthermore, when the velocity relationships have been determined and incorporated in a matrix of functions, the transpose of the matrix will automatically relate the forces.

Consider, for example, the transformation between rectangular coordinates and polar coordinates, as shown in Figure 7.24. (This will be involved in an introductory example in Chapter 9.) The relationships between r and θ and x and y are rather complicated nonlinear functions which can be expressed in several forms. For example, to convert from r and θ to x and y we can write

$$x = r\cos\theta, \quad y = r\sin\theta. \tag{7.99}$$

The inverse relationship is

$$r = \left(x^2 + y^2\right)^{1/2}, \quad \theta = \tan^{-1}(y/x). \tag{7.100}$$

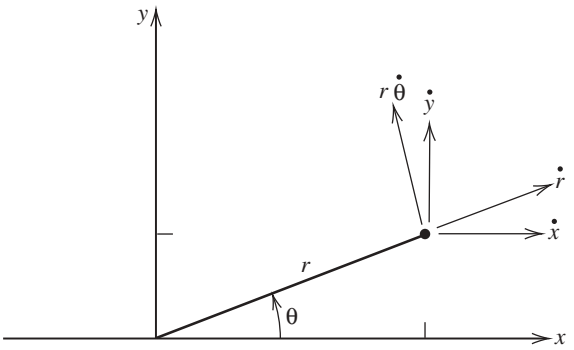


FIGURE 7.24. Rectangular and polar coordinates.

These displacement relationships can be arbitrarily complicated, but the corresponding velocity relationships have a particular structure. Starting with Eq. (7.99), for example, we find

$$\frac{dx}{dt} = \frac{\partial x}{\partial r}\frac{dr}{dt} + \frac{\partial x}{\partial \theta}\frac{d\theta}{dt},$$

and

$$\frac{dy}{dt} = \frac{\partial y}{\partial r}\frac{dr}{dt} + \frac{\partial y}{\partial \theta}\frac{d\theta}{dt}. \tag{7.101}$$

Working this out, we find a relation among flow variables,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \left[\begin{array}{c|c} \cos\theta & -r\sin\theta \\ \hline \sin\theta & r\cos\theta \end{array}\right] \begin{bmatrix} \dot{r} \\ \dot{\theta} \end{bmatrix}. \tag{7.102}$$

If this transformation is represented by a modulated multiport transformer, then we know that the transformation of the corresponding efforts (forces or torques) must involve the transpose of the matrix in Eq. (7.102):

$$\left[\begin{array}{c|c} \cos\theta & \sin\theta \\ \hline -r\sin\theta & r\cos\theta \end{array}\right] \begin{bmatrix} F_x \\ F_y \end{bmatrix} = \begin{bmatrix} F_r \\ \tau \end{bmatrix}. \tag{7.103}$$

The validity of Eq. (7.103) is easily checked by imagining forces F_x and F_y acting on the point in Figure 7.24 and then computing the radial force and the torque exerted by these forces. Note that instead of a representation as a multiport transformer, as in Figure 7.25, one could create a junction structure version involving four 2-port *MTF*s in the manner of Figure 7.23. Because of the implied causality, the 0- and 1-junctions in Figure 7.23 would interchange places.

It is not always necessary to derive velocity relationships by differentiating displacement relations. In this example, one could find $\dot{r}$ and $\dot{\theta}$ by direct consideration of the velocity components sketched in Figure 7.24. The result is

$$\begin{bmatrix} \dot{r} \\ \dot{\theta} \end{bmatrix} = \left[\begin{array}{c|c} \cos\theta & \sin\theta \\ \hline -\sin\theta/r & \cos\theta/r \end{array}\right] \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}, \tag{7.104}$$

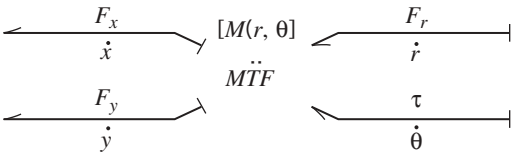


FIGURE 7.25. An *MTF* representation of the polar-to-rectangular coordinate transformation.

which implies

$$\left[\begin{array}{c|c} \cos\theta & -\sin\theta/r \\ \hline \sin\theta & \cos\theta/r \end{array}\right]\left[\begin{array}{c} F_r \\ \hline \tau \end{array}\right] = \left[\begin{array}{c} F_x \\ \hline F_y \end{array}\right]. \tag{7.105}$$

This representation reverses the causality shown in Figure 7.25 and involves the inverse of the matrices used in Eqs. (7.102) and (7.103). Of course, if this were not a transformer with equal numbers of ports on both sides, the matrix would not be square, it would not be possible to reverse causality, and there would be no inverse matrix.

A final example will make some points regarding geometric constraints in mechanical systems. The rod in Figure 7.26 is rigid and acted upon by three forces. The force F_3 is perpendicular to the rod and F_1 and F_2 are horizontal and vertical, respectively. The corresponding velocities are V_1, V_2, and V_3. Geometrically, there is only a single degree of freedom, so if one velocity is known, the remaining two can be determined.

The *MTF* shows that V_1 and V_2 are determined by V_3 and that F_3 is determined by F_1 and F_2. By consideration of the kinematics of the system, one can derive the following velocity relations directly:

$$\left[\begin{array}{c} V_1 \\ \hline V_2 \end{array}\right] = \left[\begin{array}{c} -2\sin\theta \\ \hline 2\cos\theta \end{array}\right][V_3]. \tag{7.106}$$

The transpose of the matrix in Eq. (7.106) relates the forces,

$$[-2\sin\theta \,|\, 2\cos\theta]\left[\begin{array}{c} F_1 \\ \hline F_2 \end{array}\right] = [F_3]. \tag{7.107}$$

The fact that the angle θ appears in the transformer matrices while is not one of the flows involved points up an interesting aspect of mechanical systems involving modulated transformers. In principle, the coefficients in the transformer

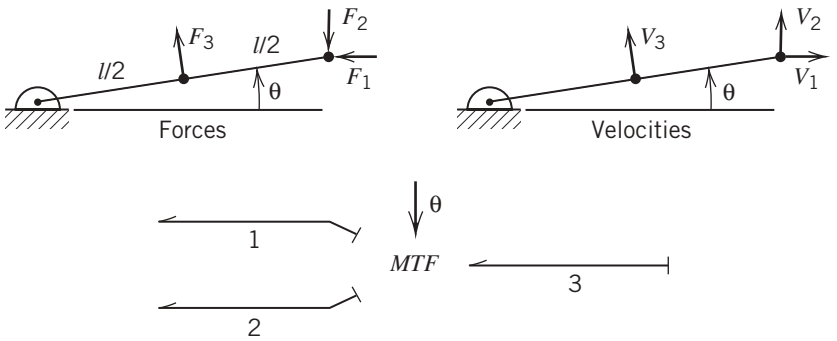


FIGURE 7.26. A rigid rod with three applied forces.

matrices are functions of displacements associated with flows at the transformer ports. This is obvious if displacement relations such as Eq. (7.99) are differentiated as shown in Eq. (7.101) to produce equations such as (7.102) and (7.103). In the present case, θ has been used even though was not a flow of interest. In this case, θ must be considered to be a state variable just like the p's and q's associated with the I's and C's in the system. This means that the *MTF* forces us to write an equation for $\dot{\theta}$ even though there may be no C-element for which θ would be a state variable. Since V_3 determined all the velocities of the rod, it also determines $\dot{\theta}$, so we must only write a separate state equation,

$$\dot{\theta} = \frac{V_3}{l/2}, \tag{7.108}$$

to complete the system formulation.

Often, C-elements attached to modulated transformers automatically make the integrals of some port flows into displacement state variables. These displacements are then available for use in computing the variable coefficients in the transformation matrices. In this example, however, does not appear on any bond so the extra state equation (7.108) is necessary.

As will become clear in Chapter 9, there are often many ways to describe the geometrical constraints in complex mechanical systems. Some choices may lead to straightforward formulations, while others may be nearly intractable. These difficulties are inherent in mechanics and arise whether bond graph methods or other techniques are used.

REFERENCES

[1] S. H. Crandall and N. C. Dahl, eds., *An Introduction to the Mechanics of Solids*, New York: McGraw-Hill, 1959, p. 378.

[2] S. H. Crandall, D. C. Karnopp, E. F. Kurtz, and D. C. Pridmore-Brown, *Dynamics of Mechanical and Electro-mechanical Systems*, New York: McGraw-Hill, 1968, pp. 220, 294, 296.

[3] I. Prigogine, *Introduction to the Thermodynamics of Irreversible Processes*, Springfield, IL: C. C. Thomas, 1955.

[4] J. Meixner, "Thermodynamics of electric networks and the Onsager–Casimir reciprocal relations," *J. Math. Phys.*, **4**, 154 (1963).

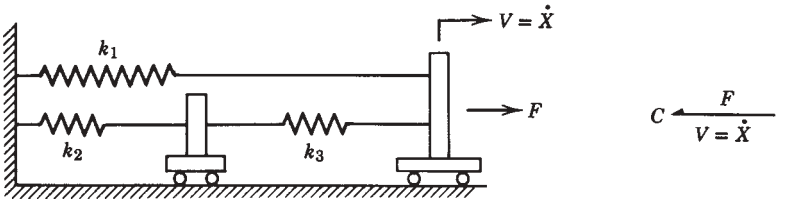
[5] G. Kron, *Tensor Analysis of Networks*, New York: Wiley, 1939.

[6] R. C. Rosenberg, "State space formulation for bond graph models of multiport systems," *Trans. ASME*, *J. Dyn. Syst. Meas. Control*, **93**, Ser. G, No. 1, 35–40 (Mar. 1971).

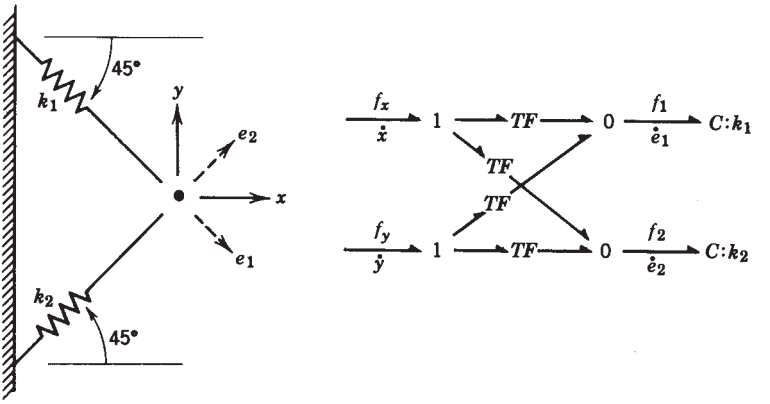
[7] D. C. Karnopp, "Power-conserving transformations: physical interpretations and applications using bond graphs," *J. Franklin Inst.*, **288**, No. 3, 175–201 (Sept. 1969).

PROBLEMS

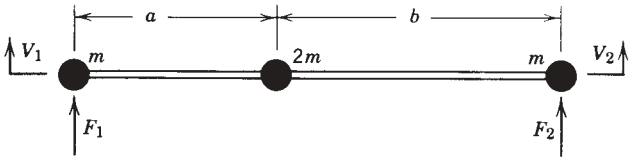
7-1. Three linear springs are attached to massless carts, as shown. Make a bond graph for this system, and manipulate the relations for the implicit C-field into the relation for the explicit 1-port C-field shown. Assume that when $X = 0$, all the springs are relaxed.



7-2. Two linear springs are pin-joined together as shown. Consider *small* motions, x and y, and small spring extensions, e_1, e_2. Show that the implicit field shown represents the system, and find the transformer moduli. Convert to an explicit-field form at the x and y ports.

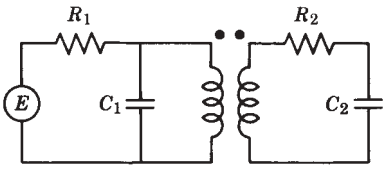


7-3. Three concentrated mass points are mounted on a massless rigid bar. Consider plane vertical motion such that the bar moves through only small angles. By expressing the motion of each mass in terms of $V_1 = \dot{X}_1$ and $V_2 = \dot{X}_2$, show that this system may be represented as a 2-port, implicit I-field. Find the constitutive relations for the I-field as seen at the F_1, F_2 ports.

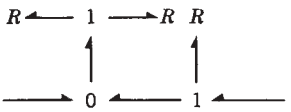


7-4. Three equal pipes carrying an incompressible fluid join at a tee junction. If each pipe has fluid inertia $\rho l/A$, show an implicit I-field representation. Demonstrate that all three 1-port inertial elements cannot simultaneously have integral causality. Find a 3-port, explicit I-field representation.

7-5. Write the equations of motion for the system shown, assuming the constitutive laws for the transformer are given by Eq. (7.63).



7-6. For the R-field shown, show causalities that would result in (a) Onsager forms, and (b) Casimir forms.



7-7. Construct an implicit R-field that is expressible in neither an Onsager nor a Casimir form, and write out its constitutive laws as was done in Eq. (7.84).

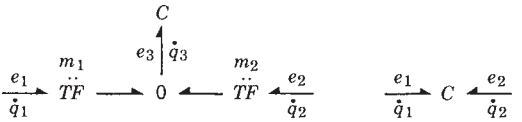
7-8. Prove that if a resistance matrix were antisymmetric (with zeros on the main diagonal), then the net power dissipated would always be zero, assuming an inward sign convention on all external bonds.

7-9. Two bond graph representations are shown in (a) and (b) below:

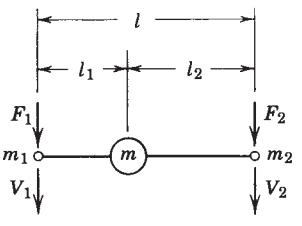
In (a) assume $q_3 = m_1 q_1 + m_2 q_2$ and $e_3 = C_3^{-1}(q_3)$, a nonlinear capacitor. If the representation of (b) is to be valid, and the 2-port C-field is energy conservative, then

$$\frac{\partial e_1}{\partial q_2} = \frac{\partial e_2}{\partial q_1}.$$

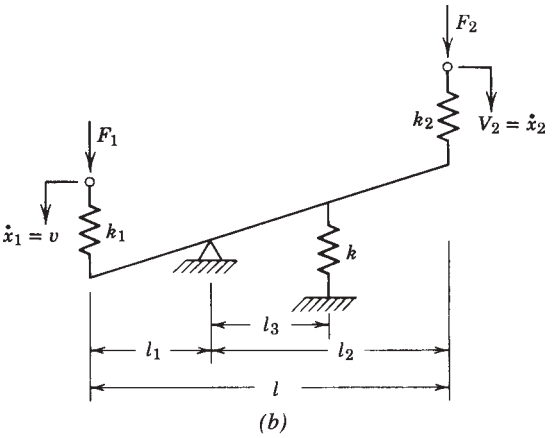
Prove that this expression is valid for the system of (a). Extend the proof to the case in which the TFs in (a) become MTFs and q_3 is a general function of q_1 and q_2, $q_3 = q_3(q_1, q_2)$.



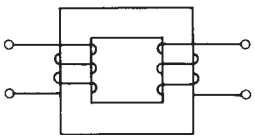
7-10. Study of some 2-port fields:



(a)



(b)



(c)

$$[L] = \left[\begin{array}{c|c} L_{11} + rM & M \\ \hline M & L_{22} + M/r \end{array}\right] = \text{Inductance Matrix}$$

(i) Draw (a), (b), and (c) as 2-port fields, and for (a) and (b) show the fields in terms of 0, 1, TF, I, C elements.

(ii) Show that the mass matrix for (a) is

$$\left[\begin{array}{c|c} m_1 + \left(ml_2^2/l^2\right) & m\left(l_1 l_2/l^2\right) \\ \hline m(l_1 l_2/l^2) & m_2 + \left(ml_1^2/l^2\right) \end{array}\right].$$

Show that the compliance matrix for (b) is

$$\left[\begin{array}{c|c} 1/k_1 + (1/k)(l_2^2/l_3^2) & -(1/k)(l_1 l_2/l_3^2) \\ \hline -(1/k)(l_1 l_2/l_3^2) & 1/k_2 + (1/k)(l_2^2/l_3^2) \end{array}\right].$$

Compare these results with the inductance matrix for case (c).

(iii) Show that

for (a):
if $m_1 \ll m, m_2 \ll m$,
then
$(l_1/l_2)p_1 = p_2$,
$v_1 = (l_1/l_2)v_2 + (l^2/ml_2^2)p_2$.

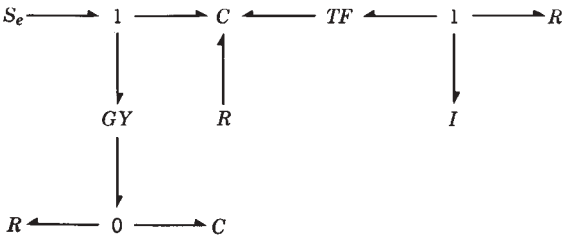
for (b):
if $k_1 \gg k, k_2 \gg k$,
then
$(l_2/l_1)x_1 = x_2$,
$F_1 = (l_2/l_1)F_2 + (kl_3^2/l_1^2)x_1$.

for (c):
if $L_{11} \ll M, L_{22} \ll M$,
then
$\lambda_1 = r\lambda_2$,
$ri_1 = i_2 + \lambda_1/M$.

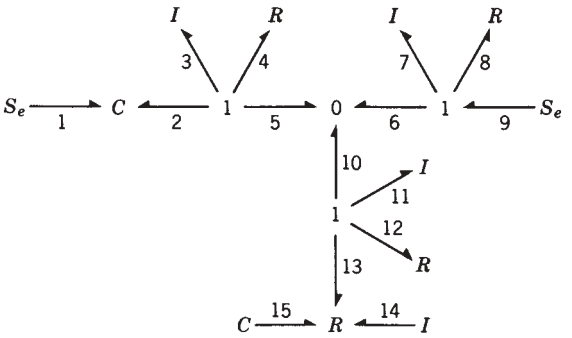
(iv) Show that for (a) if $m \to \infty$, for (b) if $k \to 0$, and for (c); if $M \to \infty$, all systems become —*TF*—.

(v) Show that saturation of the iron in (c) is analogous to the limiting of motion by stops in (b), and thus dc voltages will not pass in (c) and dc velocities will not pass in (b).

7-11. Consider the bond graphs below:
System 1:



System 2:



(a) If the graphs are not in the standard form for sign conventions, insert extra 0- or 1-junctions to bring them into the standard form.

(b) Add causality to the system, and list the state vectors.

7-12. Consider a linear, multiport C-field with stiffness matrix K. The constitutive relations are then

$$e = Kq, \tag{i}$$

where e and q are n-vectors and K is an $n \times n$ matrix. The stored energy **E** is then

$$\mathbf{E} = \tfrac{1}{2} q^{\mathrm{t}} K q. \tag{ii}$$

The text argues that if energy is conserved, then K in Eq. (i) must be symmetrical, that is, $K = K^t$.

(a) Convince yourself that any square matrix may be decomposed into a symmetrical part and an unsymmetrical part. For example,

$$K = K_s^t + K_a; \quad K_s = \frac{K + K^t}{2}, \quad K_a = \frac{K - K^t}{2};$$

$$K_s = K_s^t, \quad K_a = -K_a^t.$$

(b) Using a 2×2 example with

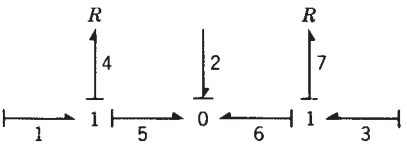
$$K = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}, k_{12} \neq k_{21},$$

show that $\mathbf{E}(q)$ depends on the symmetrical part of K, that is, that

$$\tfrac{1}{2} q^t K q = \tfrac{1}{2} q^t K_s q,$$

and that, hence, $e = K_s q$ if the constitutive laws are derived by differentiation of the stored-energy function.

7-13. The linear R-field is shown in a mixed-causal form, that is, the input and output port variables are related by a matrix K thus:



$$\begin{bmatrix} e_1 \\ f_2 \\ e_3 \end{bmatrix} = K \begin{bmatrix} f_1 \\ e_2 \\ f_3 \end{bmatrix}$$

(a) What can you predict about K before writing out the R-field equations?

(b) Suppose the R-field were forced to accept resistance or conductance causality. What characteristics would the resistance and conductance matrices have?

(c) Using causality, prove that the resistance matrix does not exist.

(d) Write out the matrix K and the conductance matrix to verify your predictions in (a) and (b). Using the equations for mixed-causal form, show that the equations cannot be manipulated into the resistance causality form.

7-14. Find a mechanical system having the implicit C-field shown in Figure 7.7*c*.

7-15. Consider a lightweight horizontal rod on which three masses are fastened, one each at both ends and one at the midpoint. Assume that the end masses move vertically in such a way that the angle of the rod with the horizontal remains small. Show that this system has a bond graph similar to the bond graph shown in Figure 7.12*b* for interconnected self-inductances.

8

TRANSDUCERS, AMPLIFIERS, AND INSTRUMENTS

This chapter deals primarily with models of devices that link two subsystems in two distinct energy domains. In some cases, the efficiency of power transduction is important. Motors, generators, pumps, and transmissions, for example, usually are designed so that they can transduce energy without much loss, at least when operating normally. Instruments and amplifiers, on the other hand, are designed to operate at low power-efficiency. An ideal instrument would extract information from a system without power absorption and could communicate at finite power to another system. An amplifier accepts an input signal at near-zero power level and influences another system in response to the input signal at finite power.

The high-efficiency transducers are usually passive; that is, they contain no sources of power. The instruments and amplifiers are often active in the sense that they need a power supply in order to satisfy the first law of thermodynamics. Most practical transducers are themselves rather complex systems that can be modeled in detail by bond graphs. On the other hand, designers of large systems containing transducers as components cannot afford the luxury of modeling transducers in detail and must use good approximate models. In this chapter, we will show the main features of several types of transducers and demonstrate a philosophy of modeling in which nonideal effects may be progressively added to an idealized basic model. We concentrate first on passive transducers and then briefly study instruments and amplifiers.

Passive transducers that contain loss elements can, on the average, only transmit less power than they receive. Some transducers do, however, have the capability of storing energy so that they can temporarily deliver excess power. We call the two types of passive transducers *power transducers* and *energy-storing transducers*.

Copyright © 2012 John Wiley & Sons, Inc.

8.1 POWER TRANSDUCERS

Ideal power transducers were introduced in Chapter 4 as transformers and gyrators. Devices such as hydraulic rams, positive-displacement pumps and motors, permanent-magnet dc motors and generators, and the like behave roughly as power-conserving elements. Real devices, of course, do exhibit power losses and also contain energy-storing mechanisms and associated inertial and capacitance effects. Although clever design of transducers can result in rather small values for these parasitic effects, they still impose an ultimate limit on the performance of any real device. When a very accurate model of a real transducer is desired, it will almost invariably be necessary to replace linear elements of the model with more accurate nonlinear ones. The process of starting with a highly idealized model, adding parasitic elements, and replacing linear elements with nonlinear ones to achieve increasingly accurate transducer models will be illustrated for some typical power transducers.

In Figure 8.1, two hydraulic ram configurations are shown schematically, and a series of bond graph models are also shown. The model of Figure 8.1*b* for the basic ram has been used in Chapter 4. This transformer model instantaneously and without loss of power transduces the hydraulic power, PQ, into mechanical power, FV. In some cases, this model may be quite adequate for a system analysis. On the other hand, in real pistons the mass effects and frictional losses associated with piston rings, packing, or tight fits to prevent leakage can be appreciable. The bond graph of Figure 8.1*c* essentially provides for a loss in the force F due to the force required to accelerate the piston and the frictional loss.

The mass effect of the piston is quite straightforward, and in many cases, when the piston rod is stiff and connects directly with a load mass, the piston mass and the load mass may simply be lumped together. But the friction force is more complex. A friction force may be small, and therefore negligible for some purposes, but the next simplest representation, the linear friction force as a function of velocity, is virtually never an accurate representation of mechanical friction. In Figure 8.2, several possible friction force laws are shown. The linear law is often used in order to study systems using the analytically convenient linear methods, but usually the friction coefficient is left as a variable so that the model can be "tuned" to reproduce experimental results. Real friction usually includes a component of dry friction, represented by $F_f = F_0 \operatorname{sgn} V$, where sgn V is $+1$ if $V > 0$ and -1 if $V < 0$. This law is not as convenient to work with as the linear law, particularly near $V = 0$. A more complicated phenomenon is sometimes called "stiction." It is easily observed that it takes more force to start a block lying on a table moving than to keep it moving at low speeds. The general friction law shown in Figure 8.2 attempts to model this phenomenon. When this type of friction is present, there is a tendency for the system to chatter. This is fine if the system is a violin string and bow but is distinctly unpleasant if the system is chalk on a blackboard, a machine tool and workpiece, or a computer simulation program with numerical stability problems. The experienced system

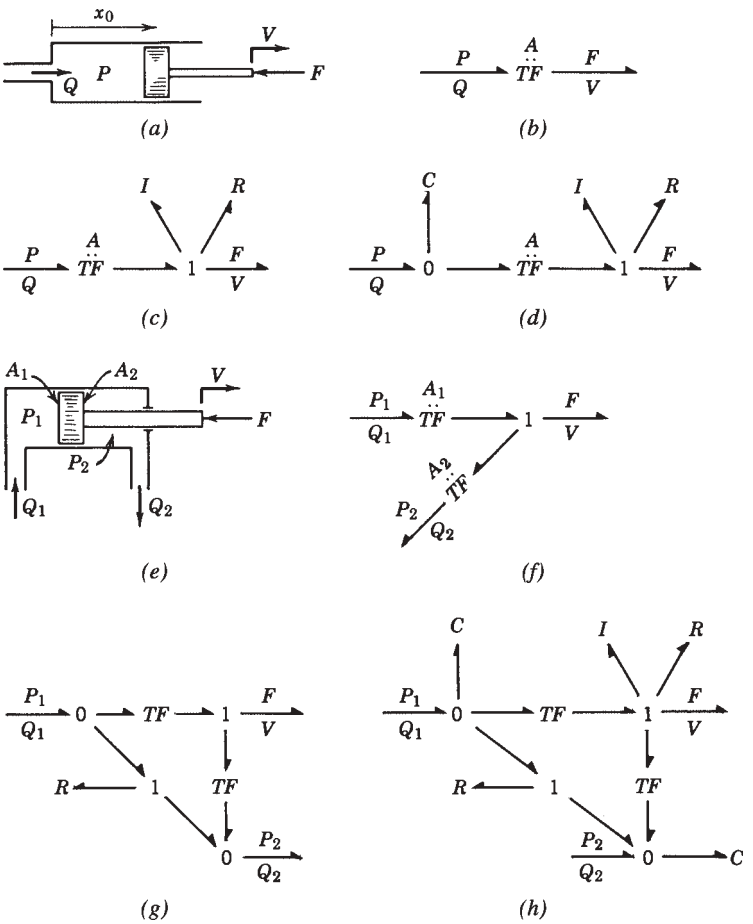


FIGURE 8.1. Hydraulic rams: (*a*) basic ram; (*b*–*d*) models of basic ram; (*e*) hydraulic cylinder; (*f*–*h*) models of hydraulic cylinder.

modeler treads a fine line between realistic but intractable friction models and useful but possibly oversimplified models.

Another phenomenon that is important in high-performance hydraulic systems involves the compliance of the working fluid. Since hydraulic fluids are not usually very compliant, small density changes accompany large pressure changes, and a linear constitutive relation usually suffices. One may define the bulk modulus β as the coefficient relating the pressure P and the decrease in volume, ΔV, of a mass of fluid that occupies volume V_0 when $P = 0$:

$$P = \beta \frac{\Delta V}{V_0}. \tag{8.1}$$

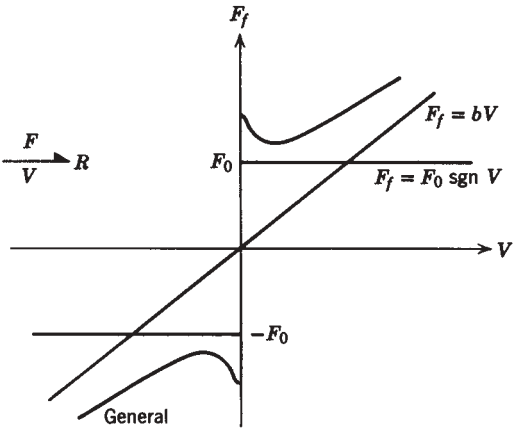


FIGURE 8.2. Several mechanical friction laws.

Actually, β, which is expressed in pressure units, varies somewhat with mean pressure, temperature, amount of air in the working fluid, and so on, so a more useful form of Eq. (8.1) is

$$P = P_0 + \beta \frac{V}{V_0} = P_0 + \frac{\beta}{V_0} \int^t Q\, dt, \tag{8.2}$$

where P_0 is an equilibrium pressure, V_0 is a nominal volume, and $Q = \dot{V}$ is the volume flow rate into the nominal volume. Then β is more obviously the slope of a nonlinear constitutive law relating pressure and the compressed volume of the fluid. The bond graph of Figure 8.1*d* shows a —*C* element that models this compliance effect. Clearly, there is an oscillatory phenomenon that occurs as energy is exchanged between the mass of the piston and the compliance of the hydraulic fluid. Note that for the linear theory of Eq. (8.2), one must pick a nominal volume, $V_0 = Ax_0$, where x_0 is an average position of the piston (see Figure 8.1*a*). Thus, the hydraulic fluid stiffness is high when x_0 is small and gets lower as x_0 is increased.

Figure 8.1*e* shows a 3-port transducer in which, because of the area of the piston rod, the areas that transduce the two pressures, P_1 and P_2, into force components are not equal. Also, the flows, Q_1 and Q_2, are not equal. The bond graph of Figure 8.1*f* represents this system in its most highly idealized form. In Figure 8.1*g* the hydraulic leakage resistance across the piston has been included, and in Figure 8.1*h* mechanical inertia and friction and hydraulic compliance effects have also been modeled. As one may see from this example, even the simplest transducer models can become rather complex when many nonideal effects are accounted for.

Another useful example of power transducers is the dc motor or generator. The basic ideal transducer is shown in Figure 8.3*a*. The power in the armature circuit, $e_a i_a$, is transduced to shaft power, $\tau\omega$, when the device operates as a motor and

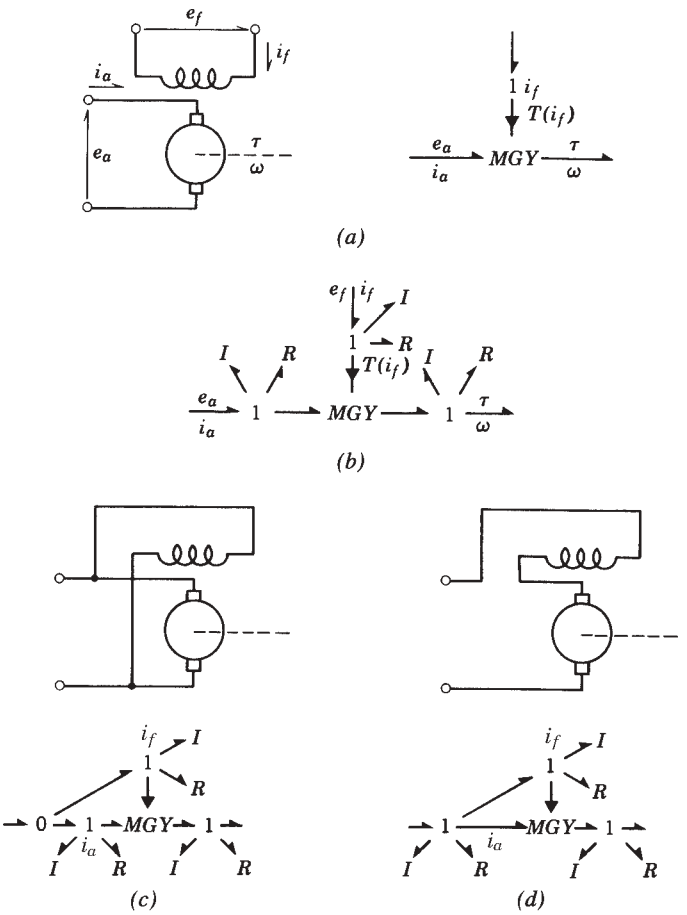


FIGURE 8.3. Direct-current electric motor–generator models: (*a*) basic ideal transducer; (*b*) self-inductances, resistances, inertia, friction, and windage effects added; (*c*) shunt motor; (*d*) series motor.

the power flow reverses during generator operation. The field port establishes a magnetic field, which provides the coupling between electric variables and mechanical variables for the individual conductors on the rotor. For a permanent-magnet motor, this field is constant, but for the separately excited motor shown in Figure 8.3*a*, the field is a function of the field current i_f. The commutator of the motor essentially maintains the field due to the armature current in a direction perpendicular to the field generated by i_f. For this reason, although i_f influences the transduction, there is virtually no back effect on e_f due to e_a or i_a. Thus, the field port influence is shown acting on an activated bond.

The equations corresponding to Figure 8.3*a* are as follows:

$$e_a = T(i_f)\omega, \qquad T(i_f)i_a = \tau, \qquad e_f = 0. \tag{8.3}$$

Note that the transduction coefficient T, which is a gyrator parameter, can assume two different values if electrical and mechanical powers are measured in different units. Thus Eq. (8.3) is sometimes rewritten in the form

$$e_a = T_{\text{em}}\omega, \qquad T_{\text{me}} i_a = \tau, \tag{8.3a}$$

with T_{em} expressed, for example, in volts per radian per second, and T_{me} in foot-pounds per ampere, for example. Power conservation is then

$$e_a i_a = \frac{T_{\text{em}}}{T_{\text{me}}} \tau\omega, \tag{8.4}$$

with $T_{\text{em}}/T_{\text{me}}$ relating volt-amperes to foot-pounds per second. One must not assume, however, that because T_{me} and T_{em} differ numerically, the device is not power conserving. In the SI system, one volt-ampere is identical to one newton-meter per second, so that T_{em} is exactly equal to T_{me}.

The parameter T involves the strength of the field as well as the number and effective lengths of the armature conductors that interact with the field. (See the discussion of the modulated gyrator in Chapter 7 for a discussion of this type of interaction with simplified geometry.) In some cases, it is reasonable to assume that the field is proportional to i_f. Then

$$T(i_f) \cong A i_f, \tag{8.5}$$

where A is a constant. Thus, if i_f is a constant, torque is proportional to i_a, and if i_a is a constant, torque is proportional to i_f. In such special cases, the device behaves linearly, even though it is basically nonlinear in a multiplicative sense. In real devices, saturation of the magnetic material limits the field, so that the relation between T and i_f does not remain linear for large currents. Also, magnetic hysteresis effects, when present, mean that the field is not a single-valued function of i_f but rather depends on the previous history of magnetization. (The permanent-magnet motor is an extreme example of this—even though there is no i_f, a field exists, and hence, $T = \text{const.}$)

In most cases, the ideal transducer must be supplemented with loss and energy storage elements to model practically important effects. In Figure 8.3*b* a common motor model is shown. The self-inductance and resistance of the field and armature coils are included, as well as the moment of inertia of the rotor and a mechanical resistor that models bearing losses and windage from the rotor and any built-in cooling fans. The electrical *I*- and *R*-elements are typically assumed to be linear unless a very accurate model is required. The mechanical resistor is rarely linear in reality, but sometimes an effective linear resistor will give good results if the mechanical friction is small enough that its detailed nature is not critical in determining system performance.

Two common 2-port motors can be constructed from the separately excited motor model. These are the shunt-wound and series-wound motors of Figure 8.3*c*

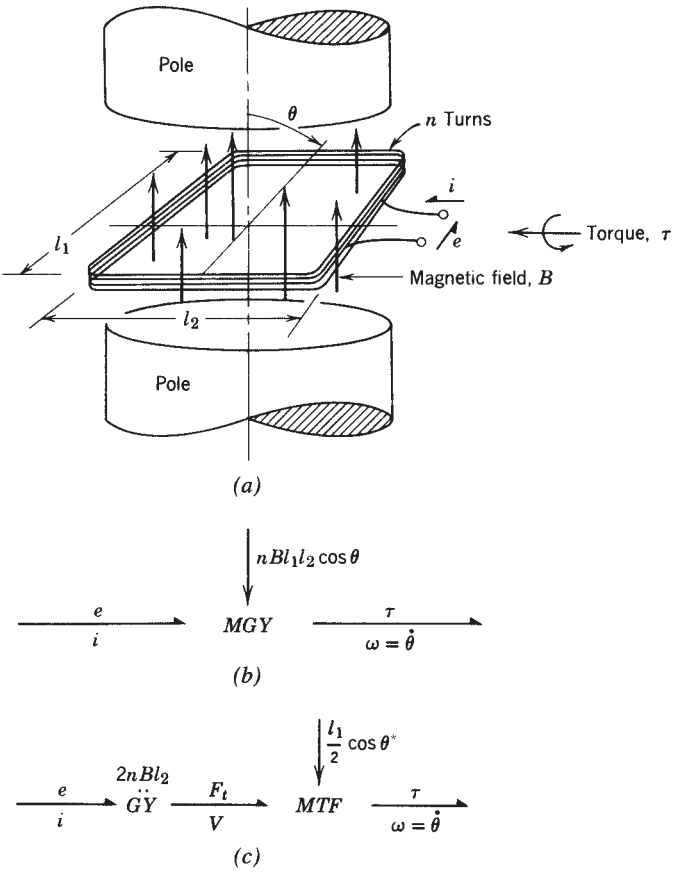


FIGURE 8.4. Elementary ac generator: (*a*) schematic diagram; (*b*) bond graph using *MGY*; (*c*) bond graph using *GY* and *MTF*. Note: The modulus will be defined by Eqs. (8.13) and (8.14).

and *d*. Note that, in Figure 8.3*d*, i_a and i_f are identical, so that one could simplify the bond graph by combining the field and armature inductances and resistances.

Before leaving the subject of power transducers, let us consider two examples in which the ideal transducer is fairly complicated even when loss and parasitic-energy-storage effects are neglected. In Figure 8.4*a* an alternating-current (ac) generator employing a permanent magnet is shown in highly schematic form. As the square coil rotates in the uniform *B*-field, there is an interaction relating the electric power variables, e and i, and the mechanical variables, τ and ω.

One of the constitutive laws of the device may be found by using the basic definition of the flux linkage λ. The amount of flux linked by the current path is B times the projected area of the coil times the number of turns in the coil, n:

$$\lambda = Bl_2l_1(\sin\theta)n. \tag{8.6}$$

Using the fact that $\dot{\lambda} = e$, one may differentiate Eq. (8.6) to find a relation between e and $\omega = \dot{\theta}$:

$$e = (nBl_1l_2 \cos\theta)\omega. \tag{8.7}$$

Then, noting that power must be conserved according to $ei = \tau\omega$, the remaining constitutive law is

$$(nBl_1l_2 \cos\theta)i = \tau, \tag{8.8}$$

where the variables are expressed in metric units so that ei and $\tau\omega$ are both in the same units of power. The constitutive equations of the device, Eqs. (8.7) and (8.8), are embodied in the bond graph of Figure 8.4*b*. Note that the modulating function involves θ, the integral of the local flow ω.

Another way to derive the laws for this device involves computing the force F on the lengths of wire that cut the flux lines. Note that the lengths of wire associated with l_2 cut flux lines, while the lengths associated with l_1 do not. If F is the *force* perpendicular to the direction of B, and V is the velocity in this direction, then for a typical conductor of length l_2, the basic equations as discussed in Chapter 7 are

$$e_1 = Bl_2V, \tag{8.9}$$

$$Bl_2i_1 = F, \tag{8.10}$$

where e_1 and i_1 are the voltage and current associated with a single length of conductor. There are, however, $2n$ such lengths of conductor, so that the terminal voltage is $2ne_1$:

$$e = 2nBl_2V. \tag{8.11}$$

Also, each length of conductor has the same current, and each produces a force that ultimately will add to produce the torque τ. Calling the total torque-producing force F_t, we have

$$2nBl_2i = F_t. \tag{8.12}$$

Now, the cutting velocity V is

$$V = \left(\frac{l_1}{2}\cos\theta\right)\omega, \tag{8.13}$$

and the relation between F_t and τ is

$$\left(\frac{l_1}{2}\cos\theta\right)F_t = \tau. \tag{8.14}$$

Thus, the entire device can be represented by a gyrator with modulus $2nBl_2$ and a modulated transformer with modulus $(l_1 \cos\theta)/2$, as shown in Figure 8.4*c*. Clearly, the two representations of Figure 8.4*b* and *c*, that is, those of Eqs. (8.7) and (8.8) and Eqs. (8.11)–(8.14), are equivalent. Loss and energy-storing elements could be added to this power-conserving model.

There are many analogies between rotary electromechanical devices and hydromechanical devices. Just as a dc motor with multiple windings and a commutator functions as a gyrator, a pump with several pistons and a porting arrangement functions essentially as a transformer. The field port of a dc motor allows modulation of the gyrator parameter, and a stroke control on a pump, if it exists, allows modulation of the transformer ratio. The ac machines are often physically simpler but functionally more complex. The ac generator of Figure 8.4 is similar to the single piston and crank arrangement of Figure 8.5, which could form part of a pump if suitable valving were added. Although the relation between the piston force and velocity and the pressure and flow of the hydraulic fluid is readily represented by a transformer, the relation between the torque τ, the angular speed ω, and the other power variables is more complex.

One way to find a relation between ω and the piston speed V is to start with a relation between x and θ (see Figure 8.5*a*). Working out the geometry, one finds

$$x = a\cos\theta + (b^2 - a^2\sin^2\theta)^{1/2}. \tag{8.15}$$

If $\dot{x} = V$ and $\dot{\theta} = \omega$, then the result of differentiating Eq. (8.15) is

$$V = [-a\sin\theta - (b^2 - a^2\sin^2\theta)^{-1/2}a^2\sin\theta\cos\theta]\omega. \tag{8.16}$$

Since $FV = \tau\omega$ if F is the force on the piston, the remaining constitutive relation must be

$$[-a\sin\theta - (b^2 - a^2\sin^2\theta)^{-1/2}a^2\sin\theta\cos\theta]F = \tau. \tag{8.17}$$

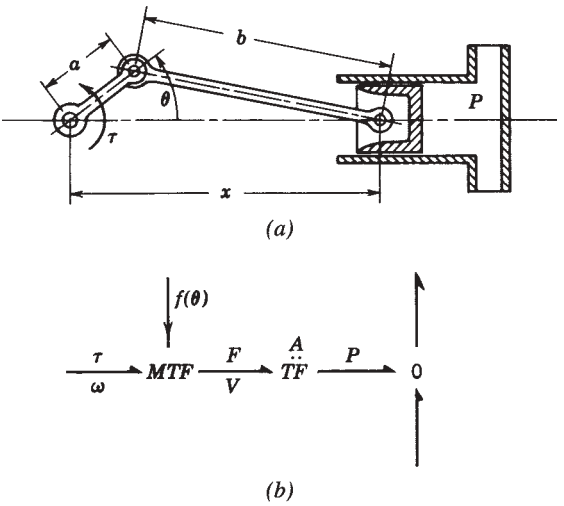


FIGURE 8.5. Crank and piston transducer: (*a*) schematic diagram; (*b*) bond graph.

Recognizing the complicated function of θ as a transformer modulus, and calling it $f(\theta)$ for convenience, one may represent the device by the bond graph of Figure 8.5*b*. Note the similarity between this bond graph and that of Figure 8.4*c*. It should be evident from these examples that considerations of power conservation can greatly aid in modeling power transducers.

8.2 ENERGY-STORING TRANSDUCERS

The transducers of the previous section were modeled, in ideal form, by transformer and gyrator elements. They were energy conservative, but beyond that, power in one domain was instantaneously transduced into another domain. In this section, we study transducers that are also ideally energy conservative but in which energy storage plays an indispensable role. Thus, for these energy-storing transducers, energy from one domain may be stored and released in another domain at a later time.

The transducer models in this section are based on C-fields, I-fields, and mixed IC-fields, which were discussed in Chapter 7. Here, we merely discuss some example systems and use the results of Chapter 7 without much discussion. As in the previous section, models of real transducers can be assembled from ideal models supplemented with loss and dynamic elements to account for effects that are present in real devices but not accounted for in the ideal transducer.

A typical energy-storing transducer of practical interest is the condenser microphone or electrostatic loudspeaker, which can be described as a two-port C-field. The sketch of Figure 8.6*a* shows roughly how these devices may be constructed. A capacitor is formed by mounting a movable plate near a rigidly mounted plate and providing electrical connections as one would in a conventional parallel-plate capacitor. In practice, the movable plate might be a thin diaphragm or a tightly stretched membrane on which a thin layer of conducting material is fixed. Such distributed-parameter "plates" could actually move in complicated ways, but for simplicity we will consider a system with only one mechanical coordinate, X, as shown in Figure 8.6*b*. (One may think of X as the displacement of the first normal-mode shape of a diaphragm for those cases in which the contribution of higher modes may be neglected.) The force F, which may be due to acoustical pressure in practical cases, is defined so that the mechanical power is $F\dot{X} = FV$, where V is the (generalized) velocity of the moving plate. The electrical power is, of course, ei.

When X is fixed, the device is an ordinary electrical capacitor, and when the charge q is fixed, F depends on X as in a mechanical spring. In general, we may hypothesize that both e and F depend on q and X:

$$e = e(q, X), \tag{8.18}$$

$$F = F(q, X). \tag{8.19}$$

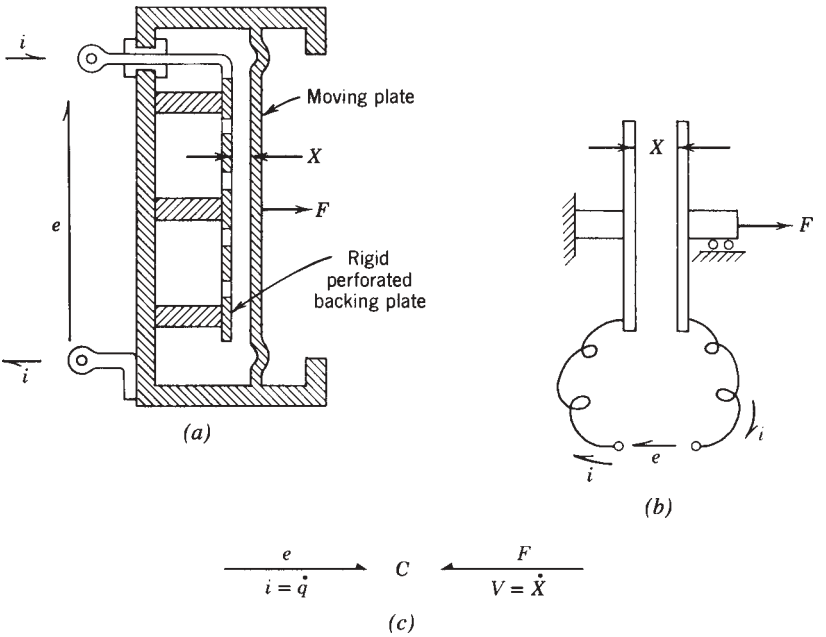


FIGURE 8.6. The movable-plate capacitor: (*a*) sketch of microphone or speaker; (*b*) schematic diagram; (*c*) bond graph representation.

But the constitutive relations above are certainly not arbitrary, since the device can at best conserve energy if we temporarily ignore loss effects. The stored energy, $\mathbf{E}$, is readily computed:

$$\begin{aligned}\mathbf{E}(t) &= \mathbf{E}_0 + \int_0^t (ei + FV)\, dt \\ &= \mathbf{E}_0 + \int_{0,0}^{q,X} e(q, X)\, dq + F(q, X)\, dX = \mathbf{E}(q, X),\end{aligned} \tag{8.20}$$

where $\mathbf{E}_0$, which will generally be assumed to vanish, represents an initial energy at $t = 0$ or when $q = X = 0$.

A close inspection of Eq. (8.20) shows that the constitutive laws of Eqs. (8.18) and (8.19) can be recovered from $\mathbf{E}(q, X)$,

$$e = \frac{\partial \mathbf{E}}{\partial q}, \qquad F = \frac{\partial \mathbf{E}}{\partial X}, \tag{8.21}$$

and the required relation between the two constitutive laws is

$$\frac{\partial e}{\partial X} = \frac{\partial^2 \mathbf{E}}{\partial X \partial q} = \frac{\partial F}{\partial q}. \tag{8.22}$$

This result is the integrability or Maxwell reciprocity conditions which the laws of the device must satisfy in order that energy be conserved. These considerations mirror those discussed in Chapter 7 for general C-fields, so the bond graph representation is simply that shown in Figure 8.6*c*. The constitutive laws for the C-field are Eqs. (8.18) and (8.19), which must obey Eq. (8.22).

A useful approximation to the constitutive laws may be found by assuming that the device is *electrically linear*, that is, that a capacitance C may be defined for every X and that

$$e = \frac{q}{C(X)} \tag{8.23}$$

is the form of Eq. (8.18). Noting that from physical reasoning $F = 0$ when $q = 0$, we may evaluate $\mathbf{E}$ using Eq. (8.20) by letting $q = 0$, taking X to any particular value, and then charging the capacitor with $X = \text{const}$ (or $dX \equiv 0$). Then the integral in Eq. (8.20) is just

$$\mathbf{E}(q, X) = \int_0^q \frac{q}{C(X)} dq = \frac{q^2}{2C(X)}. \tag{8.24}$$

The force law corresponding to Eq. (8.23) is then

$$F = \frac{\partial \mathbf{E}}{\partial X} = \frac{q^2}{2} \frac{d[C(X)]^{-1}}{dX}, \tag{8.25}$$

so that F may be determined from measurements of the variation of C with X. (For an ideal parallel-plate capacitance as sketched in Figure 8.6*b*, the law would be $C(X) = \varepsilon A/X$, where ε is the dielectric constant of the medium between the plates and A is the area of the plates.) The use of the stored energy $\mathbf{E}$ to find the force law is far more convenient than a direct calculation or an experimental measurement.

Note that even when one assumes that this device is electrically linear, the force law is decidedly nonlinear. In use, the device is normally subjected to both a high polarizing voltage together with a fluctuating signal voltage. Under these conditions, the plates tend to move together and to short the electrical circuit. This is prevented by the mechanical spring of the diaphragm, which supplies a force in the direction of F in Figure 8.6*b*. In addition, all real diaphragms have mass and exhibit some energy loss when in motion. Such effects are readily modeled by adding elements to the C-field of Figure 8.6*c*. This device is discussed in more detail in Reference [1], Chapter 6.

Another common electromechanical transducer, the electrical solenoid (shown in Figure 8.7), can serve as a prototype of a mixed IC-field transducer. The device consists simply of a coil of wire in which a soft-iron slug can freely slide. When current flows in the coil, the slug is pulled into the coil. In analyzing this device, it is clear that from the electrical port, the coil will certainly exhibit the characteristic self-inductance effects of any coil, although the position of the coil,

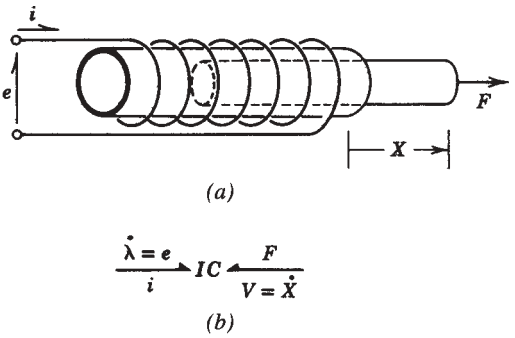


FIGURE 8.7. The solenoid. (*a*) sketch of device; (*b*) bond graph representation.

X, will presumably affect the electrical behavior. From the point of view of the mechanical port, it seems clear that the force on the slug, F, will depend on X, that is, the device will possess some of the properties of a mechanical spring, although the electrical variables will also affect F. If we assume that the current in the coil depends on the flux linkage λ as well as X, and that F also depends on λ and X,

$$i = i(\lambda, X), \tag{8.26}$$

$$F = F(\lambda, X), \tag{8.27}$$

then the stored energy $\mathbf{E}$ is

$$\mathbf{E} = \mathbf{E}_0 + \int_0^t (ie + FV)\, dt = \mathbf{E}_0 + \int_{0,0}^{\lambda,X} i\, d\lambda + F\, dX, \tag{8.28}$$

where $\dot{\lambda} = e$ and $\dot{X} = V$ and we will assume that $\mathbf{E}_0 = 0$.

By direct examination of Eq. (8.28), one may see that

$$i = \frac{\partial \mathbf{E}}{\partial \lambda}, \qquad F = \frac{\partial \mathbf{E}}{\partial X}, \tag{8.29}$$

and

$$\frac{\partial i}{\partial X} = \frac{\partial^2 \mathbf{E}}{\partial X \partial \lambda} = \frac{\partial F}{\partial \lambda}, \tag{8.30}$$

which is the integrability condition or Maxwell reciprocal condition constraining Eqs. (8.26) and (8.27) as discussed in Chapter 7. The bond graph of Figure 8.7*b* represents this ideal transducer. For mnemonic purposes the electrical port is shown impinging on the I and the mechanical port on the C.

Further insight into the device may be gained by assuming that the device is electrically linear. We assume that an inductance, $L(X)$, exists that relates i and

λ for any position of the slug. Thus, Eq. (8.26) becomes

$$i = \frac{\lambda}{L(X)}. \tag{8.31}$$

Also, on physical grounds, when i and λ vanish, F must also vanish, so that in evaluating $\mathbf{E}$ in Eq. (8.28), we may establish the slug at some particular position, X, with $\lambda = 0$, without doing any work on the device and then hold X fixed while λ is brought to its final value. During the change in λ, $dX = 0$, so only electrical energy is stored. The energy is then

$$\mathbf{E}(\lambda, X) = \int_0^{\lambda} \frac{\lambda}{L(X)} d\lambda = \frac{\lambda^2}{2L(X)}. \tag{8.32}$$

The constitutive law corresponding to Eq. (8.27) can now be found using Eq. (8.29):

$$F = \frac{\lambda^2}{2} \frac{d[L(X)]^{-1}}{dX} = -\frac{\lambda^2}{2} \frac{L'}{L^2}, \tag{8.33}$$

where L' is dL/dX. The general forms for $L(X)$ *and* $L'(X)$ are sketched in Figure 8.8. From these sketches and Eq. (8.33), one can see that the slug will experience a force tending to center it, but even when the simple law (8.31) is assumed, the corresponding force law is rather complex and inherently nonlinear in both the electrical and the mechanical variables.

The *IC*-field concept is particularly useful for electromagnetic devices, which often involve interacting magnetic fields associated with moving parts.

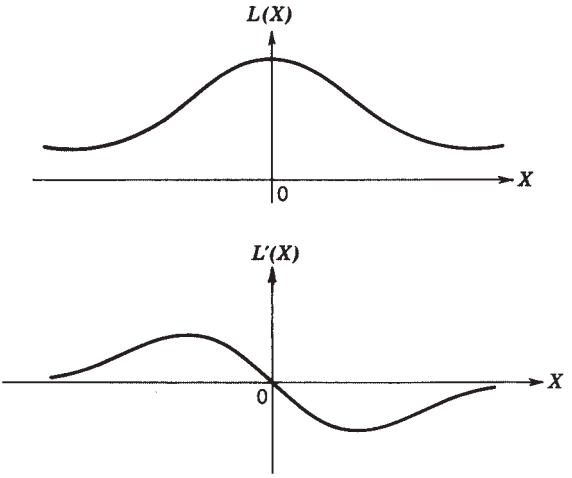


FIGURE 8.8. Inductance functions for the solenoid.

Alternating-current motors and generators, for example, typically involve coils that rotate with respect to each other. The electrical ports of these devices are inertial, and the rotary mechanical port is capacitive when the device is described as an *IC*-field. Examples of such devices appear in the problems and in Reference [1], Chapter 6.

8.3 AMPLIFIERS AND INSTRUMENTS

The central idea behind the words "amplifier" and "instrument" is a low-power or a one-way interaction without significant back effect. The description of ideal amplifiers and instruments is functional rather than physical, and thus the physical, bilateral power interactions of bond graphs must be degenerated into signal interactions by the use of activated bonds in order to represent these devices. One assumes frequently that an ideal amplifier supplies an output power variable such as a voltage or current at finite power in response to an input signal at essentially zero power. Similarly, an instrument is supposed to extract information about some variable without affecting the system in which the variable appears and to transmit the information, often at finite power levels.

Clearly, ideal instruments and amplifiers violate even more physical laws than the ideal power- and energy-storing transducers discussed previously. For example, the usual functional descriptions violate even the first law of thermodynamics, since finite output power is somehow produced from zero input power. In reality, of course, most amplifiers have a readily identifiable power supply that is built into the functional relationships. The bond graphs of Figure 8.9 show the power supply and show that power is not created from a vacuum.

On the other hand, the active bond that indicates a signal flow with no associated power can only be approximate. At the microscopic level, the Heisenberg uncertainty principle is essentially a statement that signal interactions without

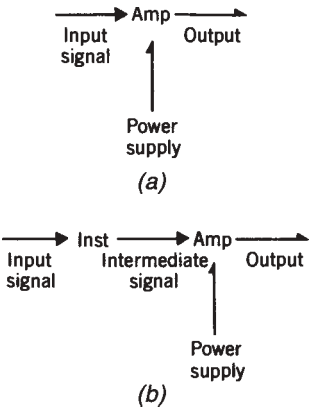


FIGURE 8.9. Amplifiers and instruments: (*a*) basic bond graph representation of an amplifier; (*b*) instrumentation system with signal level transducer and associated amplifier.

back effect are impossible, but at a macroscopic level we know that amplifiers with extremely high power gain can be built and that instruments that have virtually no effect on the observed system are available in many cases. But no absolute statements about the appropriateness of an active bond representation can be made.

Probably every engineering student has had the experience of thinking of some real instrument as an ideal signal transducer and amplifier only to find that, in some cases, the attachment of the instrument to some system greatly distorted the behavior of the system to be measured. An oscilloscope, for example, has a high but finite input impedance. Therefore, it cannot be expected to measure voltages in a system having impedances of the same order of magnitude as its own without a large effect on those voltages due to the current flowing to the instrument. Similarly, the final stage of a hydraulic amplifier of the size, say, of a ship steering engine requires sizable power levels at its input. The input can be considered to be an active bond only under restricted circumstances, for example, if the next to last stage of the amplifier is sized so that the load of the input to the last stage does not cause a significant effect on the response of the next to last stage.

In low-power applications, power efficiency is generally not very important, and amplifiers are designed to provide a drastic decoupling of input and output back effects. In high-power applications, on the other hand, amplifiers must usually be treated more physically, since it is not feasible to build such high-power-gain components that the input to an amplifier stage can reasonably be considered an active signal. Thus, it is a modeling decision whether or not to represent an amplifier or instrument with an active bond, and such a representation must be justified for each system of which the real device is to be a part.

With these caveats in mind, let us consider some useful bond graph representations of ideal elements, incorporating active bonds. The simplest type of amplifier model consists of a signal-controlled effort or flow source. In Figure 8.10*a*, the ubiquitous electrical voltage amplifier is shown in bond graph form. The output voltage e_0 is assumed to be a static or dynamic function of the input voltage e_i. In the static, linear case, there is a voltage gain G and

$$e_0 = Ge_i(t). \tag{8.34}$$

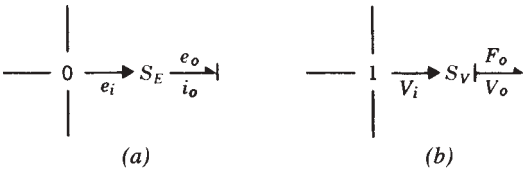


FIGURE 8.10. Amplifier models using controlled sources: (*a*) voltage-controlled voltage source; (*b*) vibration shaker shown as a velocity source.

In the dynamic case, the output voltage may be related to the input voltage by a differential equation, or in the common linear case by means of a transfer function. In the latter case, one may find internal state variables for the amplifier that yield the desired transfer function. Generally, these state variables are nonphysical and merely serve to provide state equations equivalent to the specified frequency-domain representation of the amplifier. Such state equations may be used in time-domain analysis of the complete system.

The input voltage e_i in Figure 8.10*a* is shown as coming from a 0-junction as a signal on an active bond. In writing the current sum relation for all the bonds incident on the 0-junction, one assumes that there is no current associated with the active bond. Thus, the ideal amplifier does not affect the 0-junction from which it obtains its input voltage. At the output, a finite current i_0 may exist, but e_0 is not affected by i_0. Thus, the power gain is infinite in this ideal case.

In well-designed systems, amplifiers may indeed function as controlled sources, but in some cases, the controlled-source assumption is made even when a more detailed and physical model would be preferable. In Figure 8.10*b*, for example, a vibration shaker is modeled as a controlled velocity source in which a desired velocity time history is generated as V_i and a complex servomechanism system is supposed to enforce $V_0(t)$ at a port of a test system in response to $V_i(t)$. In many cases, the reaction force does affect the servo, so that V_0 does not faithfully track V_i. In such a case, the simple model of Figure 8.10*b* is clearly inadequate, but the simplicity of controlled-source models often tempts system engineers to use them, particularly when the dynamics of the real devices have not been well explored or documented.

Many physical devices function essentially as amplifiers or transducers but are inadequately modeled as simple controlled sources. For example, a gasoline engine clearly amplifies the power of a human or automatic controller. The torque–speed curves of an engine, as sketched in Figure 8.11*a*, are drastically modified by the position of the throttle valve, which has been indicated by the angle θ. Also, the power required to move the throttle valve is often extremely small, not only in comparison with the engine power, but also in comparison with spring forces, inertial forces, and pivot-friction forces in the throttle linkage. Thus, it may be reasonable to consider that the torque associated with movement of the throttle linkage is negligible, but we cannot merely assume that θ controls a source, since the torque τ is always a function of the angular velocity of the output shaft, ω. Since torque is related to speed (effort is related to flow), one may describe the engine as a resistor, albeit an unusual one in which power is normally supplied rather than dissipated. (The power supply in the gasoline has been absorbed into the torque–speed curves.) Thus, the amplifier representation of the engine can be shown as a controlled resistance, as in Figure 8.11*b*.

The active bond indicating that there is no torque on the throttle linkage corresponding to $\dot{\theta}$ is equivalent to a signal, so one may use block diagram notation to show that $\dot{\theta}$ is integrated into the signal θ that controls the resistor representing the engine. In general, when amplifiers and instruments are part of a

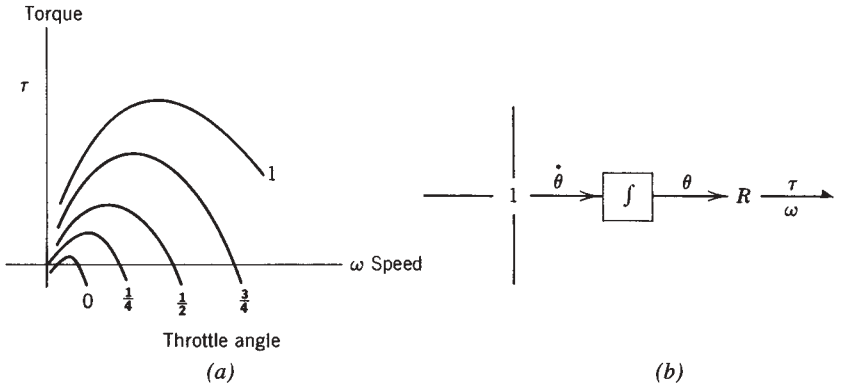


FIGURE 8.11. Static model of gasoline engine as amplifier: (*a*) torque–speed curves as a function of throttle linkage angle θ; (*b*) amplifier model of engine using controlled resistance.

system, the use of block diagrams for showing dynamic relations for active-bond signals in a bond graph can be very useful.

As another example of the use of a controlled resistance, consider the strain-gage instrument shown in Figure 8.12. The idea behind a resistance strain gage is that the electrical resistance of the gage is a function of the mechanical strain of the member to which the gage is attached. In many applications, the member functions as an elastic element that relates strain to stress and, ultimately, to force in the structure. Thus, when the strain gage is put into the bridge circuit of Figure 8.12*a*, and the bridge is supplied with the voltage e_i, then the voltage e_0 reacts to the force F in the structure. A bond graph representation of the basic instrument can be constructed out of a force-controlled resistance and the bond graph version of a bridge circuit, as shown in Figure 8.12*b* and *c*. Of course, one could continue to simplify this instrumentation system. For example, if e_i were supplied by a constant source, then one could reduce the system to a force-modulated resistance at the output port with the voltage supply built into the resistance relation. Furthermore, if an amplifier were connected to the output port, it could be arranged to supply an output voltage as a function of the force. Thus, the entire system might simplify to a force-controlled voltage source.

Although virtually any transducer might be arranged to act as an instrument under certain conditions and, in conjunction with a power supply, could function as an amplifier, variable-resistance elements are particularly important. In this category fall most electronic devices, such as transistors, vacuum tubes, and the like, and the important valve-controlled hydromechanical devices. Perhaps a final example will help indicate how such devices may be modeled in bond graph terms.

Valves are devices in which the position of a mechanical part influences the hydraulic resistance. This change in resistance can be converted to a change in pressure drop across the valve, and thus one may deduce the position of

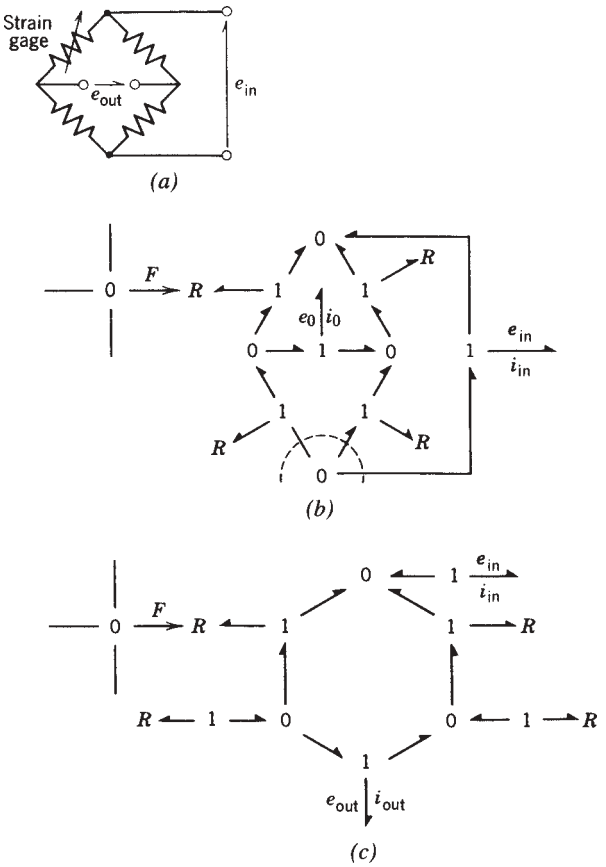


FIGURE 8.12. Strain-gage instrument. (*a*) circuit diagram; (*b*) bond graph; (*c*) simplified bond graph after choice of ground voltage.

the movable member; that is, the valve can be made into a position-indicating instrument. Also, by moving the valve, large amounts of fluid power may be controlled, and the valve may be made into an amplifier. It is this latter application that we now consider.

The so-called four-way valve shown in Figure 8.13 is common in hydraulic power systems. Actually, four resistances in the valve are modulated by the valve spool position z simultaneously. When the valve is connected to a pressure supply and a load (typically a hydraulic ram or a positive-displacement rotary hydraulic motor), a large flow of hydraulic power, $P_m Q_m$, is controlled by a small amount of power associated with the movement of the spool, $\dot{z}$. As shown in Figure 8.13, the hydraulic circuit is a bridge.

Such systems involving hydraulic valves are complicated by the intrinsic nonlinearity of hydraulic resistors. These systems have been extensively studied, however, and rather simple models may often suffice. From Chapter 7 of

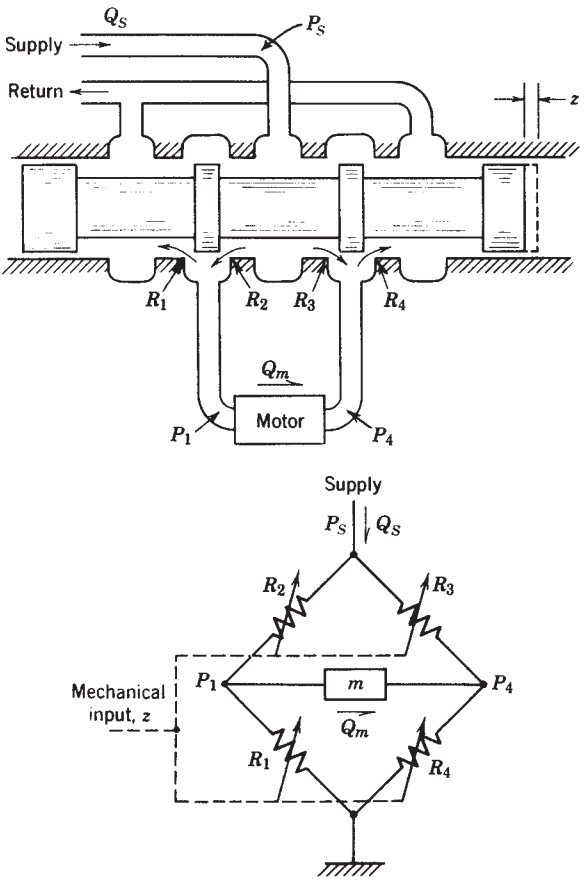


FIGURE 8.13. The general four-way valve and its equivalent circuit.

Reference [2], for example, we find the curves of Figure 8.14, which give the relation between motor flow Q_m and motor pressure P_m ($P_m = P_1 - P_4$ in Figure 8.13) for various values of the valve spool position z. Clearly, this amplifier is represented as a displacement-modulated resistance, and the curves of Figure 8.14 are the constitutive laws of the device. (Various characteristic curves for different valve geometries are given in Reference [2].) The bond graphs of Figure 8.15 show how the amplifier may be represented. In Figure 8.15*a*, the supply pressure is merely absorbed into the constitutive laws of the resistor. In Figure 8.15*b*, the valve is shown in 2-port resistance form. In this case, the effect on the amplifier of operating with various supply pressures can be studied. In either model a complex physical system is represented in a simple functional form that is valid only under restricted conditions. Naturally, more accurate physically based models of the device may be made using standard bond graph methods,

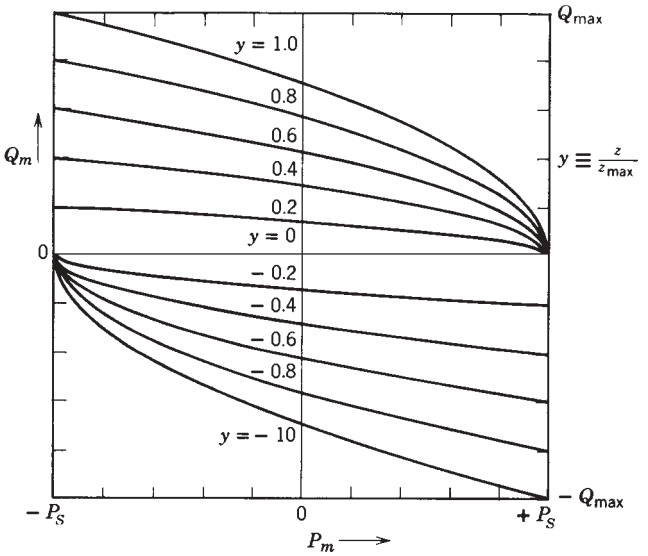


FIGURE 8.14. Pressure flow characteristics for a four-way valve.

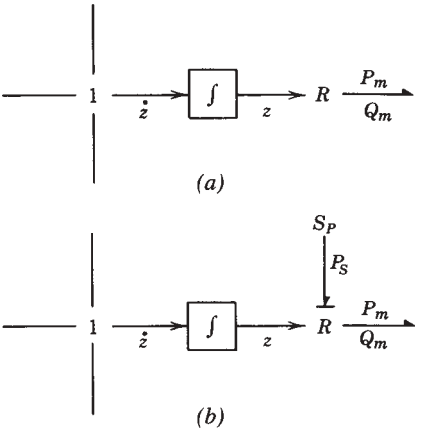


FIGURE 8.15. Bond graph models of the four-way hydraulic valve: (*a*) Displacement-modulated resistance with P_s built into constitutive law; (*b*) representation showing pressure supply.

but such models will be more complex than the functional model, and they may not add much to the usefulness of the overall system model. For this reason, the functional model may at least serve the purposes of a first system analysis. Later, a more refined model may prove desirable.

8.4 BOND GRAPHS AND BLOCK DIAGRAMS FOR CONTROLLED SYSTEMS

As discussed in Chapter 6, feedback control systems are widely used to modify the dynamic behavior of engineering systems. They require *sensors* to measure some aspects of the system response, *signal processors* to realize the control law, and *actuators* to affect the system. The sensors and actuators are usually very similar to instruments and amplifiers discussed in the last section and thus are typically modeled using active bonds. The signal processor, which is usually a special-purpose analog or digital computer, is normally described by its action on certain signals without regard to any physical power required.

Naturally, a sensor or an actuator is really a physical system and could be described by a bond graph—as could, in principle, a computer. For a well-designed system this is neither necessary nor desirable, and it is common practice to use signal flow graphs, block diagrams, or computer programs to describe the information processing that goes on between the sensor output and the actuator input in a controlled system. When the sensors or actuators have dynamics that cannot be neglected, they can be incorporated either as part of the physical system or, more artificially, as part of the dynamic laws of the control system itself.

For controlled systems, then, it is advantageous to use composite representations in which the physical system is represented using bond graphs and the controller is represented as a signal processor. We illustrate the choices of representation through an example. The signal processing will be represented through the use of block diagrams of various degrees of specificity. The combination of bond graphs and block diagrams is particularly useful when a continuous system simulation program that can accept bond graph models directly is available. (See Chapter 13.) In such a program the bond graph equations are generated automatically and the block diagram or control algorithm only needs to be added. Alternatively, one could represent even the physical system in block diagram or equivalent form, since an augmented bond graph is nothing more than a very compact representation of the information in a block diagram or equation set. For simple systems, the resulting combined block diagram may yield insight, but for more complex systems, the resulting block diagram may be too complicated to be of much use.

Figure 8.16 shows a schematic diagram of an electronically controlled automotive suspension featuring a fast-acting load leveler and a semi-active damper. See References [3–5] for details on the systems. For simplicity, the system shown is only a "quarter-car" model with the mass M representing the body and m the wheel. The velocity V_0 represents vertical velocity inputs from roadway unevenness, and k is a linearized tire spring constant. A sensor is supposed to generate an approximation to V, the absolute vertical velocity of the car body, and another measures X, the wheel-to-body deflection.

The controller drives one actuator which provides the relative velocity V_c of the main suspension spring attachment point as a conventional load leveler does. (The sketch shows the mechanical equivalent of systems which might actually

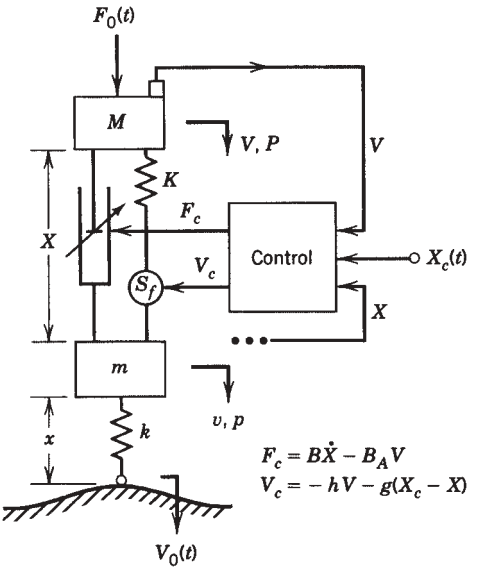


FIGURE 8.16. Quarter car model with electronically controlled suspension system incorporating a semi-active damper and a fast load leveler.

be realized using pneumatic or hydro-pneumatic means.) The other controller output is F_c, a command force that is to be realized by a semi-active damper. The damper is a variable resistor—for example, a hydraulic shock absorber with electromechanical valves. Obviously such a device cannot provide an arbitrary F_c, since its force times its relative velocity $\dot{X}$ must always represent power dissipation. Thus, in the model, a nonlinear function Φ will be used to relate the actual damper force to F_c and $\dot{X}$. The details of Φ depend on the philosophy of the particular semi-active damper law and the physical construction of the device.

In this simple case, the schematic diagram is explicit enough that from it alone equations for the physical system could be written easily; but often this is not the case. A bond graph model combined with a schematic bock diagram is shown in Figure 8.17. In this representation, it is clear that the physical system model is completely defined. Active bonds show that it is assumed that V and X can be measured by the sensors with no significant dynamics. The $1/s$ block merely indicates that the $\dot{X}$-signal on the active bond is integrated to obtain X. (The symbol s represents the Laplace transform generalized frequency variable, so $1/s$ is the transfer function of an integrator. This is discussed in Chapter 6.) This is an instance in which an active bond is treated exactly as a block diagram signal. The controller has a command input $X_c(t)$, which is related to the desired ride height and could come from another supervisory controller. (The ride height might vary with travel speed for aerodynamic reasons or because of driver manual inputs.)

The controller output V_c drives an ideal velocity source. Actuator dynamics could be included if desired by replacing the ideal source with a physical

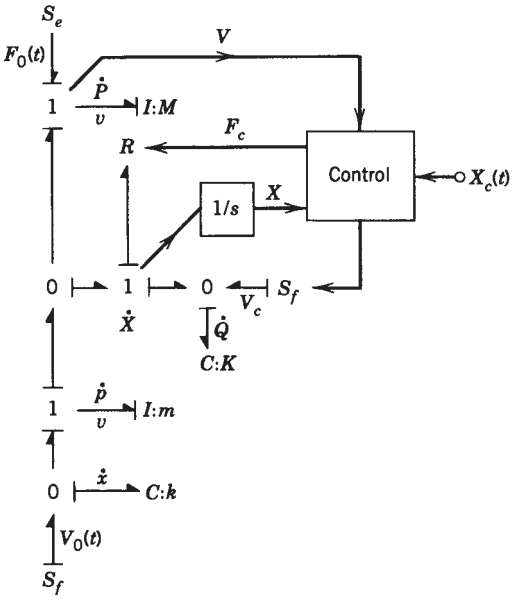


FIGURE 8.17. Bond graph for system of 8.16 using active bonds for controller input and output signals.

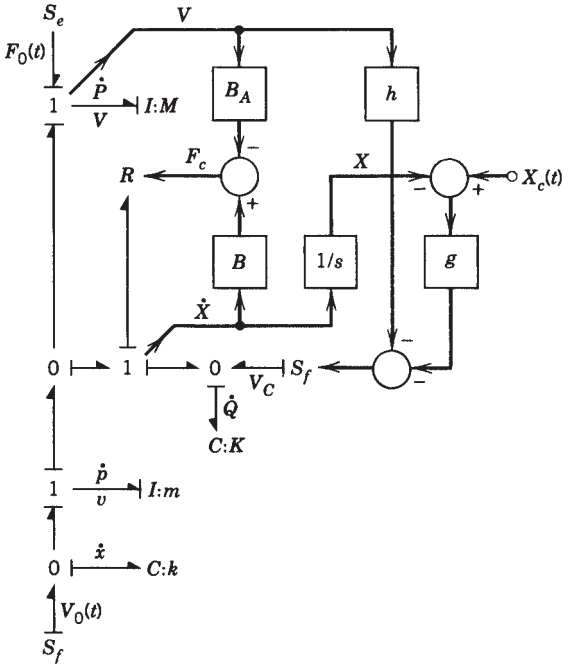


FIGURE 8.18. Bond graph incorporating block diagram for controller.

model of the actuator—perhaps a voltage source acting on a model of an electric motor and gearbox. Alternatively, some lag or delay might be incorporated in the controller block diagram or algorithm to model the actuator dynamics more accurately. The signal F_c modulates the R-element representing the controlled shock absorber. Again, engineering judgment is required to decide whether or not actuator dynamics in the semi-active damper should be included.

To simulate the system of Figure 8.17, one would only need to represent the bond graph without active bonds to a bond graph processor and then add to the program the relationships relating F_c and V_c to V, X, and X_c from a control law such as the one shown in Figure 8.16. In fact, for that rather simple control law, an explicit bond graph block diagram can be shown. See Figure 8.18. In this combined diagram, virtually all relationships are clear except for the semi-active

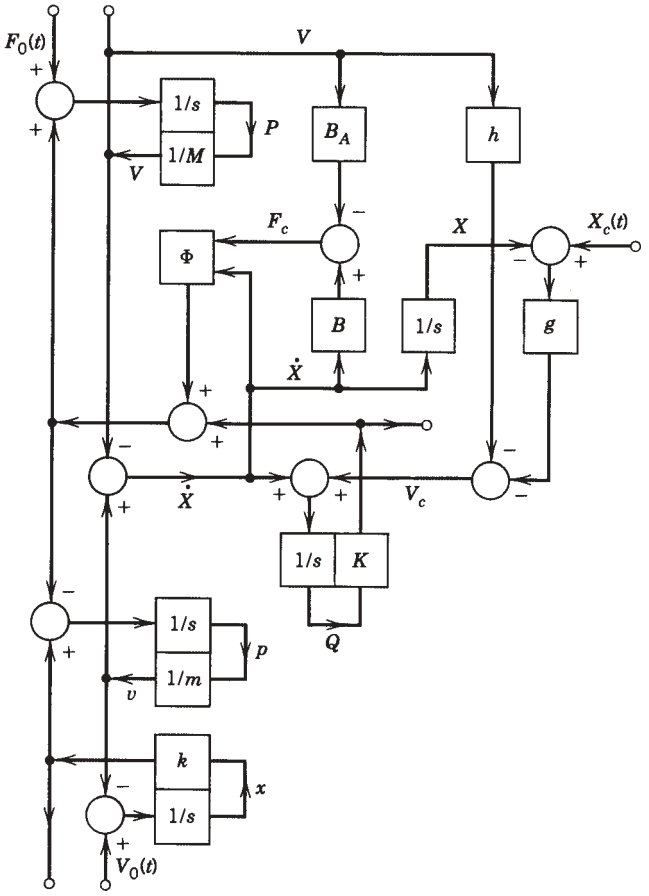


FIGURE 8.19. Block diagram derived directly from bond graph and combined with controller diagram.

damper. Here only the command force to the damper is shown, but what the actual force will be must still be specified.

Since the bond graph of Figure 8.18 is equivalent to a block diagram, it is possible to convert to a complete block diagram representation. As shown in Figure 8.19, this results in a very explicit but rather complex representation because all of the internal efforts and flows are shown. Despite some theoretical advantages to such a block diagram, few control engineers would be enthusiastic about it.

By combining some linear operations and rearranging the diagram of Figure 8.19, a diagram more to the liking of control engineers can be developed. It is shown in Figure 8.20. Here the command X_c, response X, and disturbances F_0 and V_0 are shown together with feedback loops from V and X in a conventional layout. Certainly such a diagram helps one to understand the control aspects of the system. However, the distinction between finite-power and zero-power signals has been lost completely.

The lesson is that every system representation from schematic diagrams to mixed bond graph signal diagrams to pure signal graphs or computational diagrams has its place. Success in system dynamics and control is likely to come to those who can use all types of representations. The bond graph representation is

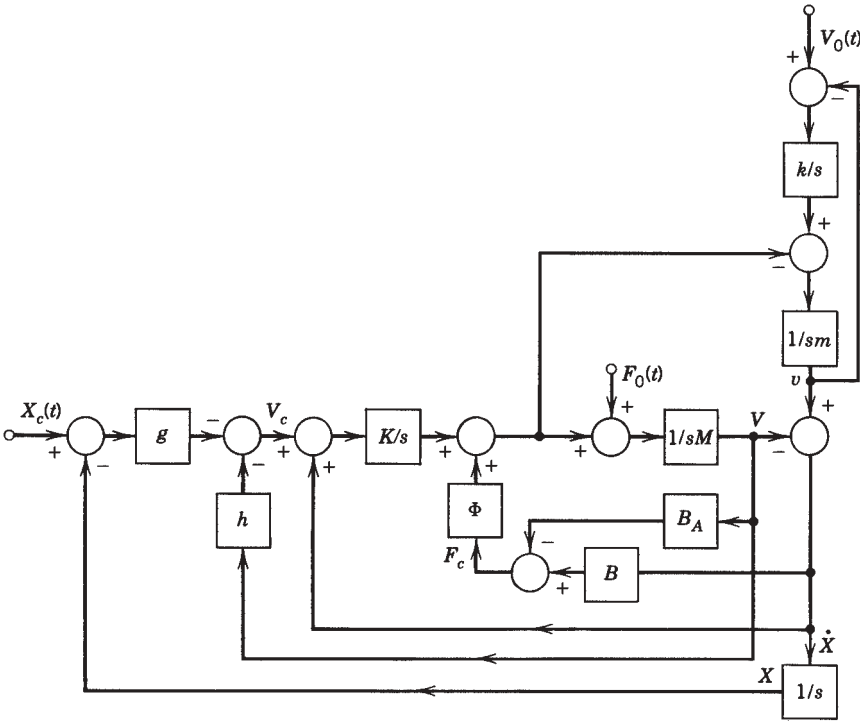


FIGURE 8.20. Block diagram of 8.19 simplified and rearranged into typical feedback-controlled form.

of particular use to make sure that power and energy interaction are consistently represented.

REFERENCES

[1] S. H. Crandall, D. C. Karnopp, E. F. Kurtz, and D. C. Pridmore-Brown, *Dynamics of Mechanical and Electromechanical Systems*, Melbourne, FL: Krieger, 1968.

[2] J. F. Blackburn, G. Reethof, and J. L. Shearer, *Fluid Power Control*, Cambridge, MA: MIT Press, 1960.

[3] D. Karnopp, Active suspensions based on fast load levelers, *Vehicle Syst. Dynam.*, **16**, No. 5–6, 335–380 (1987).

[4] D. Karnopp, Active damping in road vehicle suspension systems, *Vehicle Syst. Dynam.*, **12**, No. 6, 291–311 (1983).

[5] D. Karnopp, M. J. Crosby, and R. A. Harwood, Vibration control using semi-active force generators, *Trans. ASME J. Eng. Ind.*, **96**, Ser. B, No. 2, 619–626 (1974).

PROBLEMS

8-1. The diagram shows a positioning system using a separately excited dc motor. Physical system variables and parameters are identified:

θ_0 = output position angle;

e_{in} = input voltage;

e_a = output voltage of linear amplifier;

i_a = motor armature current;

i_f = motor field current, assumed constant;

K_a = gain of linear amplifier, assumed to have no significant time constants;

R_a = resistance of armature winding;

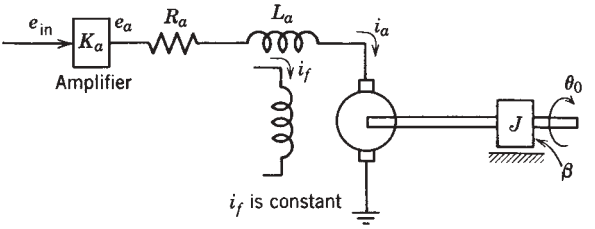
L_a = inductance of armature winding;

J = inertial load;

β = viscous–damping constant;

K_T = torque constant of motor;

K_v = back–emf constant of motor (emf = electromotive force)



The differential equations that govern the dynamics of the system are

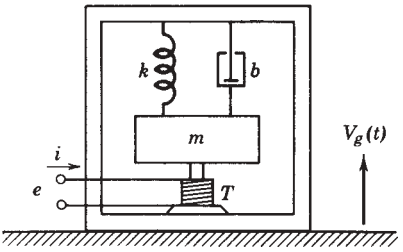
$$J\ddot{\theta}_0 + \beta\dot{\theta}_0 = K_T i_a, \qquad L_a\dot{i}_a + R_a i_a = V_a - K_v\dot{\theta}_0.$$

(a) Construct a bond graph for the system.

(b) Write state-space equations, and verify that your equations are equivalent to those listed above.

(c) Compare the two methods for analyzing this system. For example, is the system third or second order? Are K_T and K_v related in any way?

8-2. Consider the seismometer sketched below:



The input is ground motion, $V_g(t)$, and an electrical transducer using a permanent magnet moving in a coil reacts to the relative motion between the case and the seismic mass m.

(a) Construct a bond graph for the device, leaving the electrical port as a free bond and neglecting coil resistance and inductance.

(b) Assume that the device is connected to a voltage amplifier, so that $i \simeq 0$. Find the transfer function between $V_g(t)$ and e.

(c) It is sometimes preferable to use current rather than voltage as a signal for reasons connected with noise pickup. Suppose that a current amplifier is connected to the seismometer terminals, so that $e \cong 0$. Show that when the coil resistance is neglected, one can still find a transfer function between V_g and i, even though the system state equations degenerate.

(d) Reconsider the case of (*c*) when the coil resistance R_c is not neglected.

8-3. Consider the solenoid shown in Figure 8.7, but include the mass m of the moving element as well as a Coulomb friction force F_f, with the constitutive law

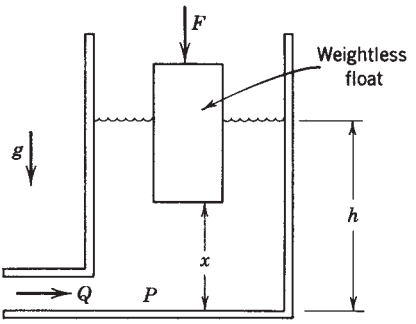
$$F_f = F_0 \operatorname{sgn} V = \frac{F_0 V}{|V|},$$

where F_0 is the magnitude of the friction force. Show a bond graph for this transducer, and write state equations for it assuming electrical linearity and a voltage source input.

8-4. The float shown is cylindrical and is immersed in water in a cylindrical container. The cross-sectional area of the float is not negligible with respect to the tank area. The pressure P at the tank inlet is γh, where γ is the weight density of the water and h is the height of the water. The volume of water, V, is the time integral of the flow rate Q. If we assume that

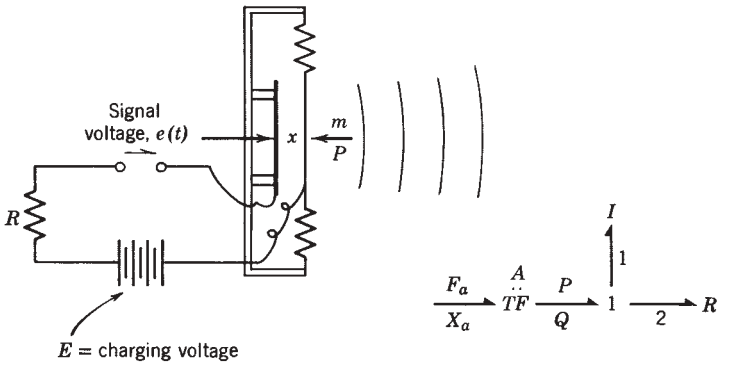
$$P = P(V, x), \qquad F = F(V, x),$$

what type of bond graph element is this device? Are the functions (P and F) related in any way? Can you sketch them?



8-5. An electrostatic loudspeaker system is shown in which a high charging voltage E and a signal voltage $e(t)$ are applied through a current-limiting resistor R. The moving plate or membrane may be assumed to have effective mass m and a mechanical spring constant k. (The mechanical spring represents the combined effect of membrane tension and the air spring formed by the sealed cabinet.) The effect of the acoustic loading is shown by the bond graph fragment, in which A is the effective area of the speaker and P represents the acoustic overpressure. The I and R

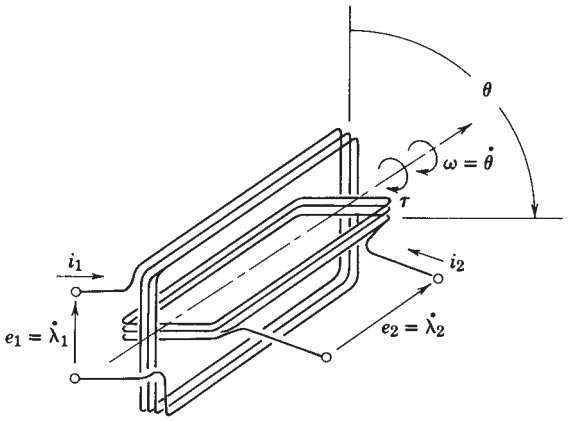
elements may be adjusted at any single frequency to match the impedance of the air:



Assuming the basic transducer is an electrically linear device with capacitance $C(x) = \varepsilon A/x$, where ε is the dielectric constant for the air in the cabinet, construct a bond graph for the system. Let the mechanical spring be relaxed at $x = x_0$ when $E = 0$. Write the state equations for the system, letting the acoustic inertia and resistance be called I_1 and R_2, respectively.

8-6. The basic transduction mechanism for electrical alternators and motors may be understood by studying the ideal energy-storing transducer shown below, consisting of a fixed and a moving coil. Assume the two currents and the torque τ are related to two flux-linkage variables λ_1, λ_2 and the angular position θ:

$$i_1 = i_1(\lambda_1, \lambda_2, \theta), \qquad i_2 = i_2(\lambda_1, \lambda_2, \theta), \qquad \tau = \tau(\lambda_1, \lambda_2, \theta).$$



For the *electrically linear* case, it is conventional to specify self- and mutual-inductance parameters; for example,

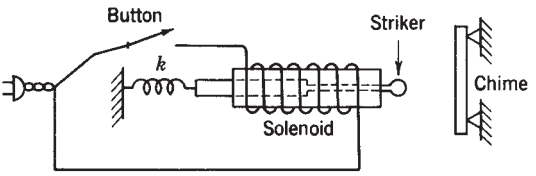
$$\begin{bmatrix} L_1 & L_0\cos\theta \\ L_0\cos\theta & L_2 \end{bmatrix}\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix},$$

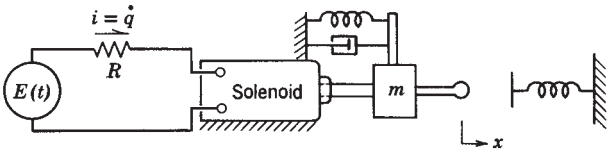
where L_0, L_1, L_2 are constants and $L_1L_2 \geq L_0^2$

(a) What kind of bond graph element describes this device?

(b) Derive the torque relation from the inductance matrix above by computing the stored energy at fixed θ when λ_1 and λ_2 are brought from zero to final values and then differentiating the energy function with respect to θ (see Reference [1], pp. 322–323).

8-7. A model of a doorbell chime is shown that features a solenoid, return spring, frictional force, and a striker and chime. The schematic diagram attempts to depict the physical effects to be modeled for the case in which the button switch is closed. Make a bond graph for the system, including all physical effects. Let the striker–chime interaction be modeled by a nonlinear spring with zero force until the striker contacts the chime. Write the equations of motion for this device, assuming electrical linearity for the solenoid. Leave your results in functional form; that is, let $L(x)$ be the inductance and $F(x)$ be the striker spring force. Sketch the shape of $L(x)$ and $F(x)$.





8-8. A speed control system uses a dc motor with the following parameters:

L_f = field inductance;

R_f = field resistance;

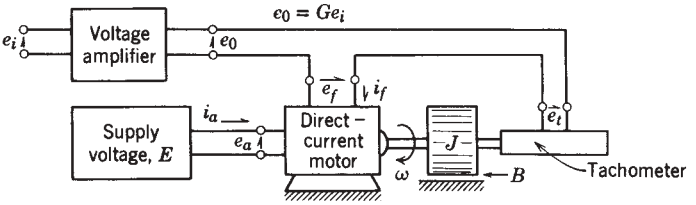
L_a = armature inductance;

R_a = armature resistance;

J = total moment of inertia;

B = rotary dashpot coefficient;

$T(i_f)$ = transduction coefficient (see Figure 8.3).



The control voltage, e_i, drives an amplifier that functions as a voltage-controlled voltage source:

$$\xrightarrow{e_i} S_e \xrightarrow{e_0}, \qquad e_0 = Ge_i(t).$$

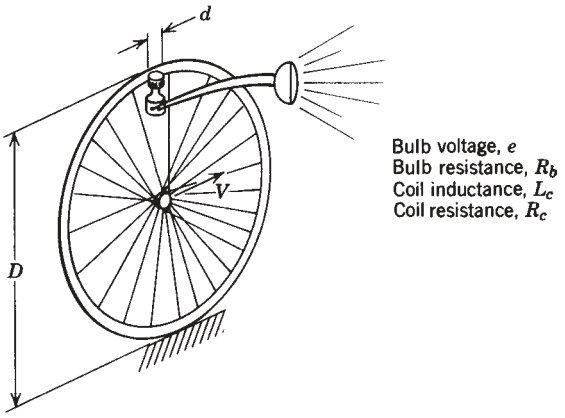
The tachometer is a permanent-magnet device that functions as a instrument, so that its mechanical bond may be activated:

$$1_\omega \rightarrowtail - \overset{K_t}{\ddot{G}Y} \underset{i_t}{\overset{e_t}{\rightarrow}}, \qquad e_t = K_T\omega$$

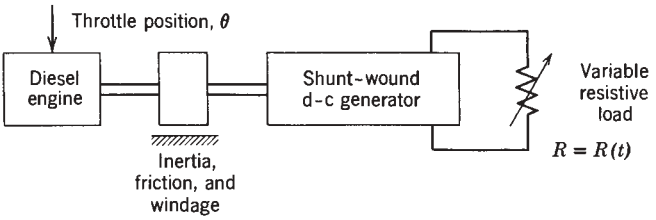
Write a bond graph for the system, write state equations, and find the transfer function between e_i and ω when i_a = const.

8-9. An ac bicycle generator is shown in which a small wheel of the diameter d bears upon the bicycle wheel of diameter D. Assume the generator functions exactly like the one shown in Figure 8.4, although in reality it is more common for a magnet to rotate inside a fixed coil. Set up a bond graph representation that models the coil self-inductance and resistance and the bulb resistance and that would suffice to predict how the bulb voltage would vary with forward speed,

Explain the self-regulation feature of the L–R circuit that allows the bulb voltage to rise less than proportionally to V due to frequency response effects.



8-10. It is desired to study the response of the generating system shown to changes in throttle setting, $\theta(t)$, and load resistance, $R(t)$. Use the bond graph model of Figure 8.11 to represent the engine, and let the load resistance be controlled by an active bond, $-R \dashrightarrow$.



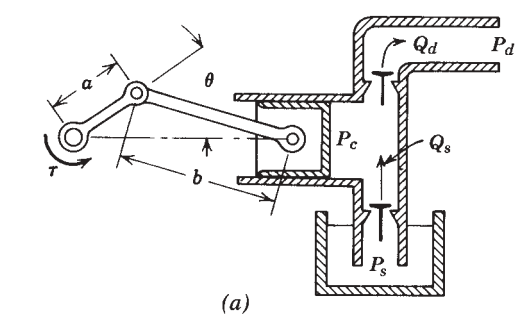
Show a bond graph for the system, and write the equations of motion in functional form. See Figure 8.3.

8-11. A simple pump is sketched in part *a* below. The check valves are represented by nonlinear resistors with characteristics such as that shown in part *b*.

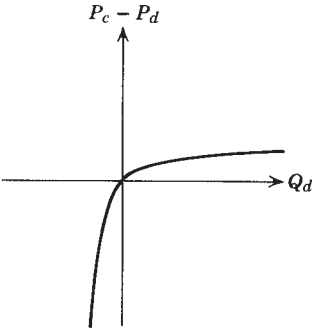
(a) Verify the bond graph shown in part *c*, and compute the modulation function $f(\theta)$.

(b) Suppose the speed, $\omega = \dot{\theta}$, is essentially constant. Sketch how the discharge flow and the suction flow vary as functions of time.

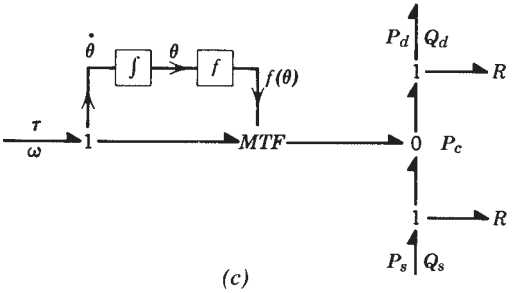
(c) You should be able to see that this pump acts like an electrical half-wave rectifier circuit. Invent a pump that acts like a full-wave rectifier circuit, and show a bond graph for your pump.



(a)



(b)



(c)

8-12. Consider the high-performance speed control system shown below. The idea is to drive a hydraulic pump with a shunt-wound dc motor and to control the speed ω_2 by stroking the bypass valve in a hydrostatic transmission. The valve stroke is $x(t)$, which may be changed manually or, ultimately, automatically by a servocontrol system. The system should have fast response, since if the valve is slammed shut, the hydraulic motor pump will suddenly be directly coupled so that energy stored in the rotating inertia J_1 will be available to accelerate J_2 very rapidly.

(a) Draw a bond graph for the system including all the effects shown in the sketch. Augment the bond graph, and list the state variables required.

(b) Either write out the state equations or construct a block diagram corresponding to the bond graph. Identify the input variables to your dynamic equations.

The basic motor characteristic is

$$T(i_f)i_a = \tau, \qquad e_a = T(i_f)\omega_1.$$

The pressure–flow law for a valve depends on the stroke x, that is,

$$P = A(x)Q\,|Q| = A(x)Q^2 \operatorname{sgn} Q.$$

The model should include armature and field inductance and resistance. The constitutive laws for a pump and hydraulic motor are

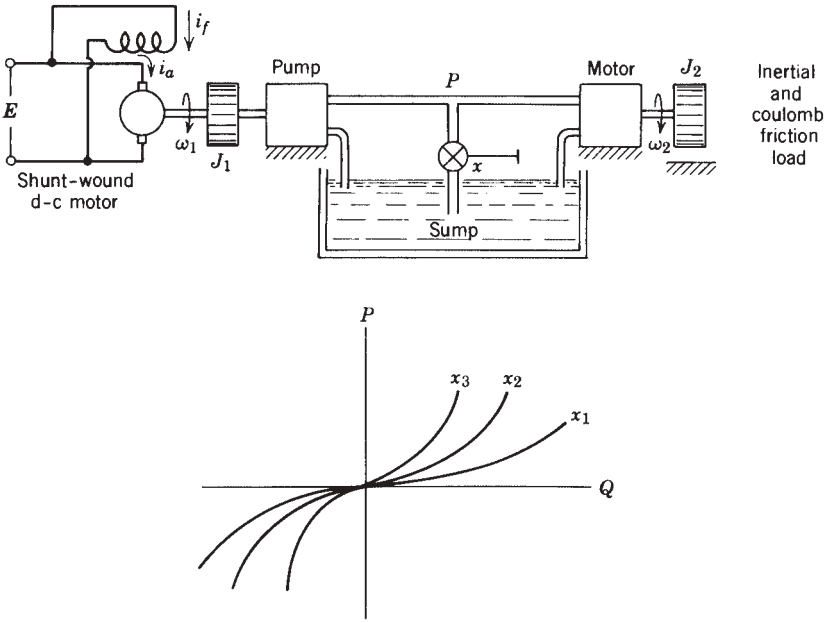
(1) for a pump,

$$\tau_p = \alpha_p P_p, \qquad \alpha_p \omega_p = Q_p;$$

(2) for a motor,

$$\tau_m = \alpha_m P_m, \qquad \alpha_m \omega_m = Q_m.$$

Neglect leakage in the pump and motor and compressibility in the oil lines.



8-13. In the diagram, a conventional hydropneumatic suspension system such as those found on certain automobiles has been made active by the addition of a valve. A simple model of the valve is indicated in which the flow of oil, Q, is simply proportional to the relative displacement of the suspension, x. We assume a very high-pressure source of oil (not shown) so that the valve can act as a source of flow modulated by x,

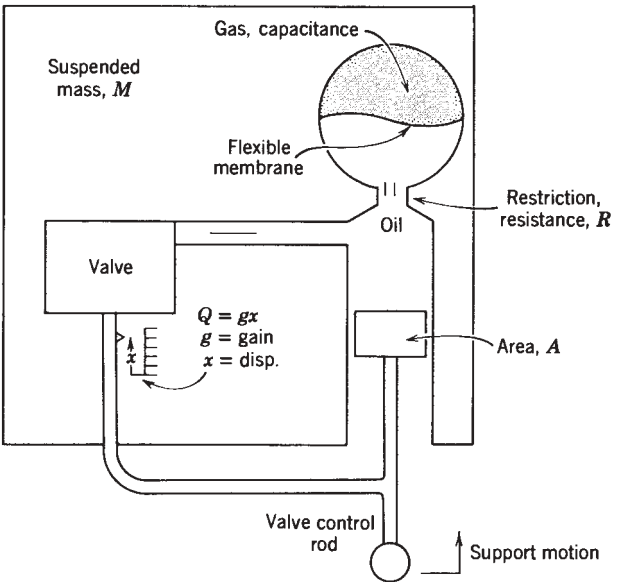
$$\underset{x}{\rightarrow} S_{Q_1} \mapsto,$$

independent of the pressure in the suspension. With this system, the "static deflection" of the suspension is always zero for any gain g, but as g is increased, it is not so clear that the system will remain stable.

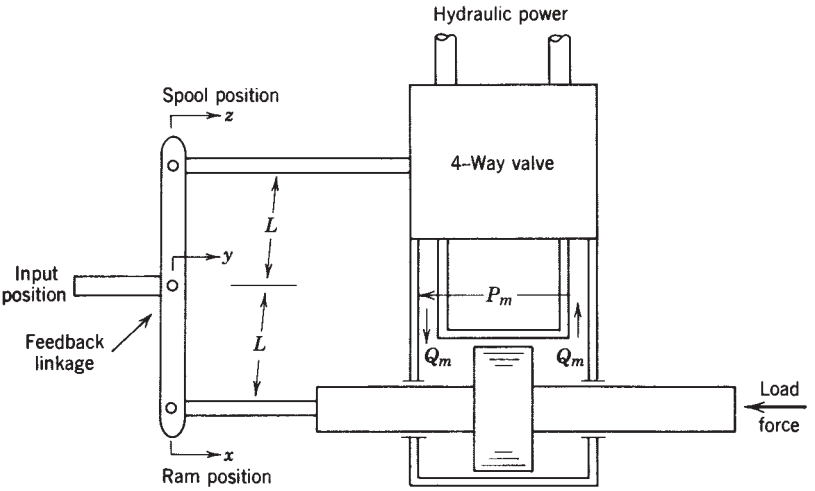
(a) Find a bond graph for the system.

(b) Decide on appropriate system inputs, and write a set of state-space equations for the system.

(c) Assuming linearized characterizations for all elements of the system, set up an expression that will yield system eigenvalues for any particular numerical values of parameters.



8-14. A servomechanism has been constructed by connecting a spool valve and hydraulic ram with feedback linkage:



Using the text discussion and Figures 8.13, 8.14, and 8.15, construct a bond graph for this device. Assume the following:

(a) The force required to move the valve is $F_r = F_0 \operatorname{sgn} \dot{z}$, where $F_0 =$ const.

(b) The load force is inertial and resistive, and oil compressibility is included.

(c) The linkage is light and frictionless and moves only through very small angles.

(d) The working area of the ram is A, and the supply pressure is P_s. Using your bond graph, answer the following questions:

i. What is the force required to move the input, and what is the maximum load force possible?

ii. With zero load force, what is the maximum velocity of the ram?

iii. Write the equations of motion, leaving the valve constitutive laws in a general functional form, for example, $Q_m = Q_m(z, P_m)$.

8-15. Add a control system block diagram to the bond graph for the motor shown in Problem 8-1 for two cases:

(a) A speed control system such that $e_{\text{in}} = g(\dot{\theta}_c - \dot{\theta}_0)$, in which g is a gain and $\dot{\theta}_c$ is a command angular speed.

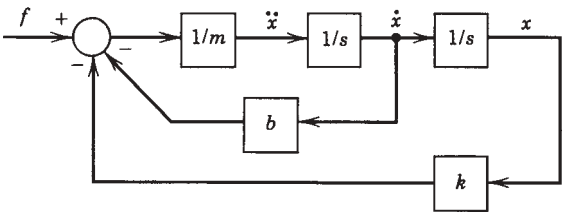
(b) A position control in which $e_{\text{in}} = -g\dot{\theta}_0 + h(\theta_c - \theta_0)$, g and h are gains, and θ_c is an angular position command signal.

In each case, note the order of the controlled system. Write the state equations for (b).

8-16. A mass–spring–damper system has an equation of motion

$$m\ddot{x} + b\dot{x} + kx = f,$$

where f is an applied force. A standard way to make a block diagram for such a system results in the following:



Make a bond graph for this system and convert it directly into a block diagram in the fashion of Figure 8.19. Then rearrange the block diagram and show that it is equivalent to the one shown above.

8-17. Consider an electrically linear parallel-plate capacitor as described by Eqs. (8.23) and (8.25). Rewrite the force law in the form

$$F = \frac{-q^2 C'}{2C^2}, \qquad C' = \frac{dC(X)}{dX},$$

in analogy to the solenoid force law in Eq. (8.33).

(a) Use Eq. (8.23) to find the force for the case when the voltage is held constant at the value e_0 by eliminating q in favor of e_0.

(b) Using the approximate expression $C(X) = \varepsilon A/X$ mentioned below Eq. (8.25), show that when q is constant, the force does not vary with X under this assumption, but if the voltage is held constant, the force does vary with X.

8-18. A butterfly condenser has a capacitance that varies with the angular position of a rotor, θ, rather than with a linear position X as in Problem 8-17. This requires a torque, τ, in analogy with the force F for a parallel-plate capacitor. Assuming the capacitance varies with θ according to

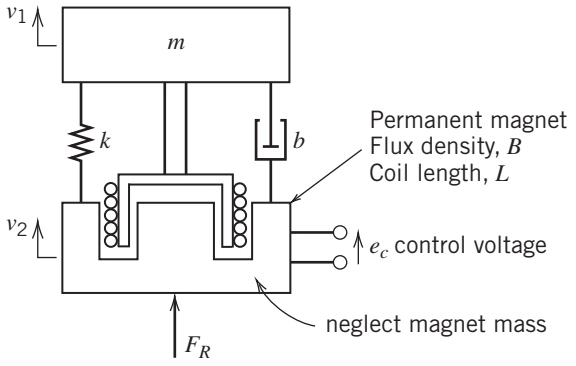
$$C(\theta) = C_0 + C_1 \cos 2\theta$$

with

$$C_0 = 15 \times 10^{-12} F, \qquad C_1 = 10 \times 10^{-12} F,$$

and assuming the condenser is attached to a 1000-V source, compute the magnitude of the torque resulting at an angle when it is a large as possible.

8-19. An inertial actuator is a device in which, by appropriate acceleration of an internal mass, a prescribed reaction force can be generated. By attaching the actuator to a structure, the structure motion can be controlled.



The device shown is voice coil driven, where only the winding resistance is important on the electrical side. The mass m is attached to the base with spring k and damper b.

Here, F_R is the output force. Assume the base has prescribed motion, $v_2(t)$. Construct a bond graph model and derive state equations. Derive the output equation for F_R.

8-20. From the state equations of Problem 8-19, derive the transfer functions relating F_r to v_2 and F_r to e_c. Derive

$$\frac{F_r}{v_2} = G_{Fv}(s) \quad \text{and} \quad \frac{F_r}{e_c} = G_{Fc}(s).$$

8-21. Since Problem 8-19 is treated as a linear system, we can express the output force in terms of the transfer functions from Problem 8-20 as

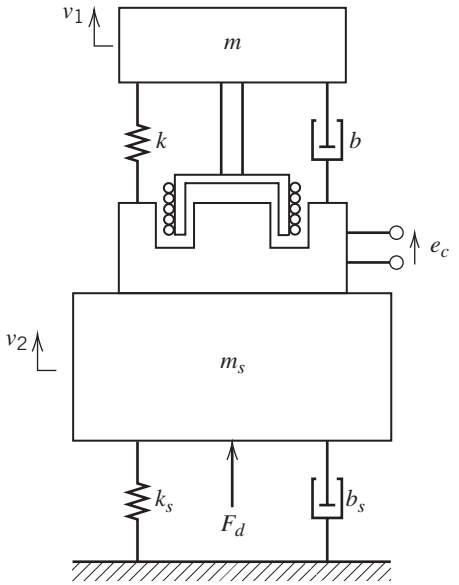
$$F_R(s) = G_{Fv}v_2 + G_{Fc}e_c.$$

Sometimes we desire the reaction force to mimic the effect of a damper attached between the base and inertial ground. In other words, we desire

$$F_R = b_c v_2,$$

where b_c is a controller gain. Derive the ideal control filter that will yield the desired reaction force from the inherent dynamics of the device.

8-22. Shown below is the device from Problem 8-19 attached to a structure consisting of a mass, m_s, spring, k_s, and damper, b_s. A force, F_d, acts upon the structure mass. Construct a bond graph model for this total system and derive state equations.



Derive the open-loop ($e_c = 0$) transfer function relating output, v_2, to input, F_d. Using the control transfer function from Problem 8-21, derive the closed-loop transfer function between v_2 and F_d. Does the result have a component similar to a damper to ground?

8-23. Repeat Problems 8-19 through 8-22, but assume that the voice coil is current driven, i_c, rather than voltage driven.

9

MECHANICAL SYSTEMS WITH NONLINEAR GEOMETRY

Bond graphs straightforwardly indicate, through causality assignment, the state variables appropriate for the system model. But we know that algebraic loops and derivative causality may exist. Methods of dealing with these problems were presented in Chapter 5. However, some algebraic problems may be so complicated, involving so many variables in nonlinear functions, that a computable version of the model is virtually impossible to formulate unless the model is carefully constructed.

This chapter deals with multidimensional rigid-body mechanics and the inclusion of such complex mechanical elements in overall system models. Geometric nonlinearities are common in such systems and are notorious for causing equation formulation problems. It turns out that it is possible to cast rigid-body mechanics into a bond graph format such that the resulting bond graph fragment can readily be used in the construction of overall models.

This chapter starts out with a general exposition of a bond graph representation of three-dimensional rigid-body dynamics in which the complexities of the geometric nonlinearities of mechanical systems is made evident. There follow several methods for representing mechanical subsystems in bond graph terms. The desired form for these models is presented, although sometimes algebraic difficulties make the achievement of the this form difficult. Next, the problems that arise when several rigid bodies are rigidly connected (the *multibody* problem) is discussed and ways to circumvent formulations problems using inserted compliance elements is discussed. Then, the use of IC-fields to represent system described with Lagrange equations is presented. This technique is particularly suited to complex multibody subsystems. Finally, the application to bond graph modeling in the study of vehicle dynamics is presented. Vehicle systems are an example of the engineering use of nonlinear mechanics, although there are many others. Starting with a rigid body vehicle model, one can add suspension

Copyright © 2012 John Wiley & Sons, Inc.

components, tires, engines and engine mounts, brakes, and so on in a manner that is the computational equivalent to bolting the hardware together.

The theory presented in this chapter is not simple. However, with some persistence the reader will be rewarded with the ability to construct reasonably low-order system models that include three-dimensional motion of rigid bodies and yet are straightforwardly formulated into state equations, ready for computation.

9.1 MULTIDIMENSIONAL DYNAMICS

Figure 9.1 shows a general rigid body both translating and rotating in space. Inertial axes X, Y, Z are shown as well as the x, y, z axes, which are attached to the body, at its center of mass, and aligned with the principal axes of the body. With respect to these body-fixed coordinates, the rotational inertial properties remain invariant and the products of inertia are all zero. While these body-fixed coordinates are not the best from which to view the body motion, they are practical coordinates for the computation of body motion.

At the instant shown, the body has absolute velocity $\mathbf{v}$ and absolute angular velocity $\boldsymbol{\omega}$. These vectors have been cast into three mutually perpendicular

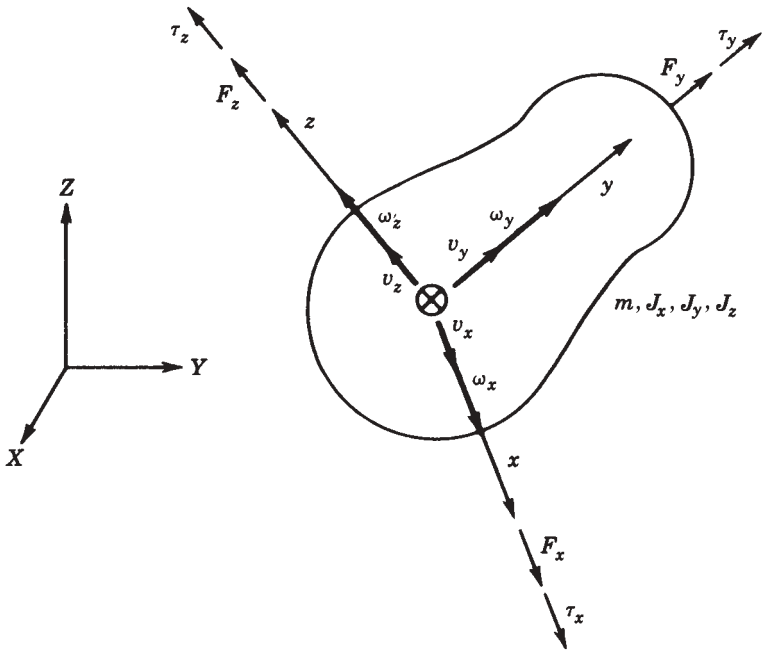


FIGURE 9.1. Body in general 3-D motion.

components: v_x, v_y, v_z and $\omega_x, \omega_y, \omega_z$. According to Newton's law, the net force **F** acting on the body changes its momentum:

$$\mathbf{F} = \frac{d}{dt}\mathbf{p}, \tag{9.1}$$

where the momentum is related to the velocity of the center of mass,

$$\mathbf{p} = m\mathbf{v}. \tag{9.2}$$

If **v** is expressed with respect to the rotating body centered frame, see Ref [1], then

$$\mathbf{F} = \left.\frac{\partial \mathbf{p}}{\partial t}\right|_{\text{rel}} + \omega \times \mathbf{p}, \tag{9.3}$$

where $\left.\frac{\partial \mathbf{p}}{\partial t}\right|_{\text{rel}}$ indicates the rate of change of momentum relative to the moving frame.

An angular momentum law analogous to Eq. (9.1) can be derived if the net torque acting on the body $\boldsymbol{\tau}$ and the angular momentum **h** are evaluated with respect to either a fixed point, if one exists for the body, or the center of mass,

$$\tau = \frac{d}{dt}\mathbf{h}. \tag{9.4}$$

If the $x-y-z$ axis system is assumed to be aligned with the principal axes of the body, then the angular momentum is related to the angular velocity by

$$\mathbf{h} = \mathbf{J}\omega, \tag{9.5}$$

in which **J** is a diagonal matrix of the principal moments of inertia, J_x, J_y, J_z. For **h** expressed with respect to the rotating body centered frame,

$$\tau = \left.\frac{\partial \mathbf{h}}{\partial t}\right|_{\text{rel}} + \omega \times \mathbf{h}. \tag{9.6}$$

Using the right-hand rule, it is straightforward to write down the component equations for (9.3) and (9.6). They are

$$F_x = m\dot{v}_x + m\omega_y v_z - m\omega_z v_y, \tag{9.7}$$

$$F_y = m\dot{v}_y + m\omega_z v_x - m\omega_x v_z, \tag{9.8}$$

$$F_z = m\dot{v}_z + m\omega_x v_y - m\omega_y v_x \tag{9.9}$$

and

$$\tau_x = J_x\dot{\omega}_x + \omega_y J_z\omega_z - \omega_z J_y\omega_y, \tag{9.10}$$

$$\tau_y = J_y\dot{\omega}_y + \omega_z J_x\omega_x - \omega_x J_z\omega_z, \tag{9.11}$$

$$\tau_z = J_z\dot{\omega}_z + \omega_x J_y\omega_y - \omega_y J_x\omega_x. \tag{9.12}$$

These nonlinear differential equations are known as Euler's equations. They have no general solution and can only be solved analytically for some special cases. If they were solved, then we would know $v_x, v_y, v_z, \omega_x, \omega_y, \omega_z$, with respect to a frame that will be aiming in a different direction at every instant, which makes interpreting the body motion a bit difficult. Furthermore, the force and torque components (coming from attached systems) must be aligned with the body-fixed coordinates in order to use these equations. This rarely occurs naturally, so the utility of these equations is questionable at this stage of development.

If we treat the cross-product terms as forces in Eqs. (9.7), (9.8), and (9.9) and as torques in Eqs. (9.10), (9.11), and (9.12), then an elegant bond graph representation results for translation and rotation, as shown in Figure 9.2. The reader should add the forces and torques at the respective 1-junctions and be convinced that Euler's equations are, in fact, represented by these bond graph fragments. These gyrator ring constructions appropriately introduce the momenta p_x, p_y, p_z and angular momenta p_{Jx}, p_{Jy}, p_{Jz} as state variables. The modulated gyrator

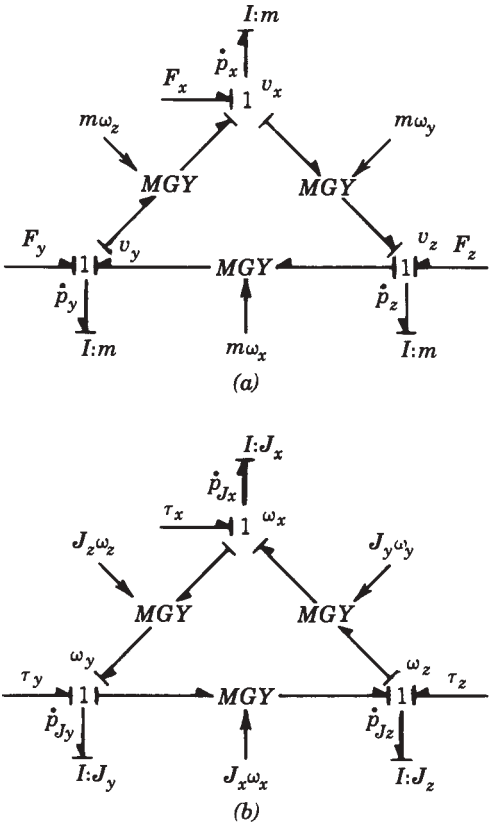


FIGURE 9.2. Bond graphs for 3-D rigid-body motion.

(—*MGY*—) elements represent the cross-product terms in Eqs. (9.7)–(9.12), and the gyrator moduli are time varying and depend on the angular momentum state variables.

As long as external forces F_x, F_y, F_z and torques τ_x, τ_y, τ_z, are causal inputs to the bond graph fragments, integral causality exists for all I-elements. If we agree to always use these fragments with forces and torques as causal inputs and with velocities and angular velocities as causal outputs, then we can write the state equations, once and for all, for these submodels, ready to be appropriately connected to any external systems in the process of constructing an overall model. We make this agreement, and write the state equations as

$$\dot{p}_x = F_x + m\omega_z \frac{p_y}{m} - m\omega_y \frac{p_z}{m}, \tag{9.13}$$

$$\dot{p}_y = F_y + m\omega_x \frac{p_z}{m} - m\omega_z \frac{p_x}{m}, \tag{9.14}$$

$$\dot{p}_z = F_z + m\omega_y \frac{p_x}{m} - m\omega_x \frac{p_y}{m}, \tag{9.15}$$

$$\dot{p}_{J_x} = \tau_x + J_y\omega_y \frac{p_{J_z}}{J_z} - J_z\omega_z \frac{p_{J_y}}{J_y}, \tag{9.16}$$

$$\dot{p}_{J_y} = \tau_y + J_z\omega_z \frac{p_{J_x}}{J_x} - J_x\omega_x \frac{p_{J_z}}{J_z}, \tag{9.17}$$

$$\dot{p}_{J_z} = \tau_z + J_x\omega_x \frac{p_{J_y}}{J_y} - J_y\omega_y \frac{p_{J_x}}{J_x}, \tag{9.18}$$

where

$$\omega_x = p_{J_x}/J_x, \tag{9.19}$$

$$\omega_y = p_{J_y}/J_y, \tag{9.20}$$

$$\omega_z = p_{J_z}/J_z. \tag{9.21}$$

A shorthand notation for the bond graphs of Figure 9.2 is shown in Figure 9.3.

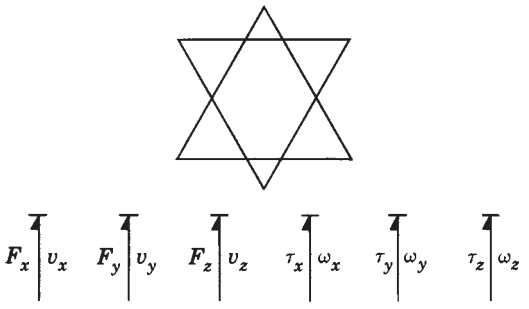


FIGURE 9.3. Shorthand representation of 3-D rigid-body mechanics.

9.1.1 Coordinate Transformations

Since it is unlikely that external forces and torques will be conveniently lined up with the continuously changing principal directions, and since the body motion is difficult to interpret with respect to the body-fixed coordinates, it is necessary to transfer from body-fixed coordinates to other more convenient frames through a series of coordinate transformations. There are many possible coordinate transformations that can take us from body-fixed coordinates to inertial coordinates, the most familiar probably being the transformation involving Euler angles. The transformation used depends largely on the system being analyzed: Spinning bodies of revolution, typical of gyroscopic motion, are most conveniently analyzed using Euler angles, while multibody systems, such as robotic arms, often some use other transformations.

The concept of a coordinate transformation will be developed here using ground vehicles as an example, and we shall introduce the *Cardan* angles, which correspond to the familiar yaw, pitch, and roll angles of an automobile. Figure 9.4 shows an inertial frame (X, Y, Z), a body-fixed frame (x, y, z), and two intermediate frames (x', y', z' and x'', y'', z''). The body-fixed frame is arrived at by first rotating about Z through angle ψ (yaw), yielding the frame x', y', z'. We next rotate about the y'-axis through the angle θ (pitch), yielding the x'', y'', z'' axes. Finally, we rotate about the x''-axis through the angle ϕ (roll), yielding the instantaneous body-fixed frame, x, y, z.

Let us now assume that body-fixed angular velocities $\omega_x, \omega_y, \omega_z$ are known (outputs, by agreement, from the gyrator-ring bond graph fragments), and write down the angular velocity components in the intermediate frames and finally in the inertial frame:

$$\omega_{x''} = \omega_x, \tag{9.22}$$

$$\omega_{y''} = \omega_y \cos\phi - \omega_z \sin\phi, \tag{9.23}$$

$$\omega_{z''} = \omega_y \sin\phi + \omega_z \cos\phi; \tag{9.24}$$

$$\omega_{x'} = \omega_{x''} \cos\theta + \omega_{z''} \sin\theta, \tag{9.25}$$

$$\omega_{y'} = \omega_{y''}, \tag{9.26}$$

$$\omega_{z'} = -\omega_{x''} \sin\theta + \omega_{z''} \cos\theta; \tag{9.27}$$

$$\omega_X = \omega_{x'} \cos\psi - \omega_{y'} \sin\psi, \tag{9.28}$$

$$\omega_Y = \omega_{x'} \sin\psi + \omega_{y'} \cos\psi, \tag{9.29}$$

$$\omega_Z = \omega_{z'}. \tag{9.30}$$

We can write these relationships in matrix form as

$$\begin{bmatrix} \omega_{x''} \\ \omega_{y''} \\ \omega_{z''} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}, \tag{9.31}$$

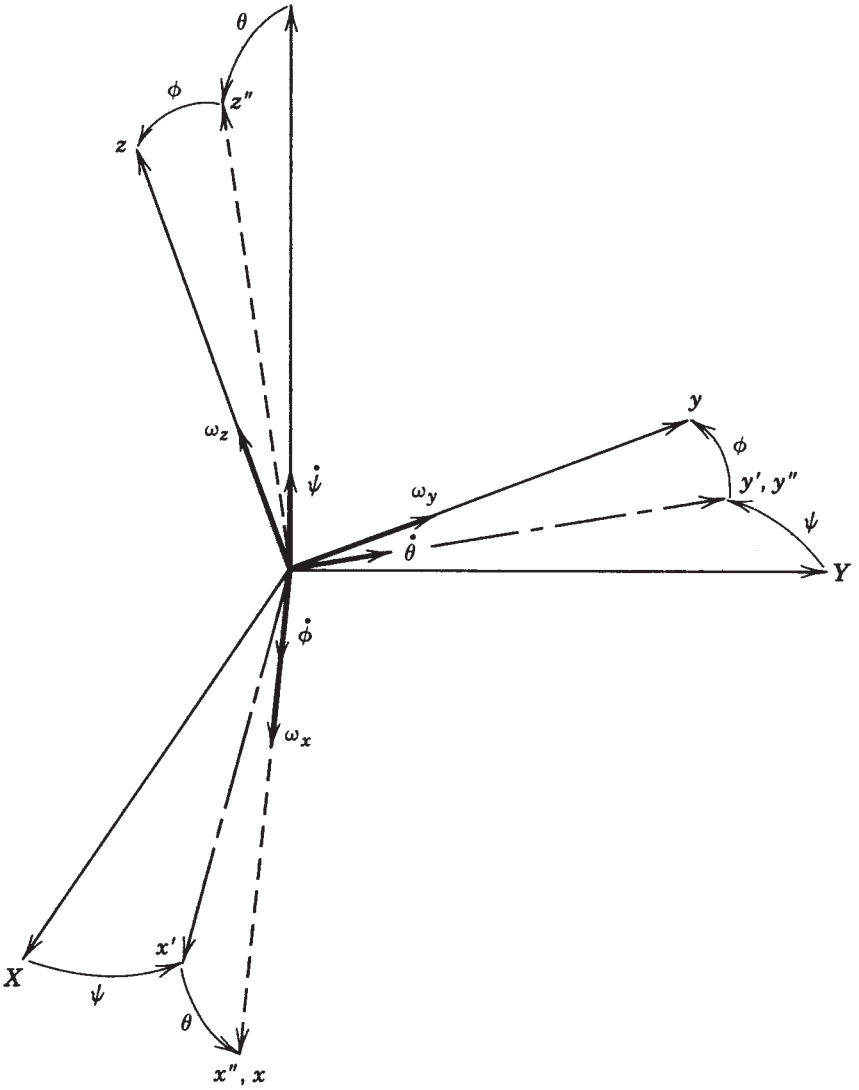


FIGURE 9.4. Cardan angle coordinate transformations.

$$\begin{bmatrix} \omega_{x'} \\ \omega_{y'} \\ \omega_{z'} \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \omega_{x''} \\ \omega_{y''} \\ \omega_{z''} \end{bmatrix}, \tag{9.32}$$

$$\begin{bmatrix} \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_{x'} \\ \omega_{y'} \\ \omega_{z'} \end{bmatrix}. \tag{9.33}$$

Thus, if we know $\omega_x, \omega_y, \omega_z$, then we can determine the angular velocity components in all frames, including the inertial frame. If we denote the transformation matrices as

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix}, \tag{9.34}$$

$$\Theta = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \tag{9.35}$$

$$\Psi = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{9.36}$$

then

$$\begin{bmatrix} \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} = \Psi\Theta\Phi \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}. \tag{9.37}$$

It should be realized that the exact same relationship exists for transforming velocity components in the body-fixed directions to velocity components in the inertial directions. Thus,

$$\begin{bmatrix} v_X \\ v_Y \\ v_Z \end{bmatrix} = \Psi\Theta\Phi \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}. \tag{9.38}$$

This transformation from body-fixed components to inertial components is a power-conserving transformation and is shown as a bond graph in Figure 9.5. The modulated transformers, —*MTF*—, indicate the three transformations $\mathbf{\Phi}$, $\mathbf{\Theta}$, $\mathbf{\Psi}$, which take the body-fixed velocity and angular velocity components through the intermediate frames and produce the inertial velocity and angular velocity components.

We now make use of the power-conserving nature of this transformation of coordinates. Since no energy is stored or dissipated by the transformation, power must be instantaneously conserved. This was shown in Chapter 7 to result in the following relationship among the forces and torques:

$$\begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} = (\Psi\Theta\Phi)^{\mathrm{t}} \begin{bmatrix} F_X \\ F_Y \\ F_Z \end{bmatrix}, \tag{9.39}$$

$$\begin{bmatrix} \tau_x \\ \tau_y \\ \tau_z \end{bmatrix} = (\Psi\Theta\Phi)^{\mathrm{t}} \begin{bmatrix} \tau_X \\ \tau_Y \\ \tau_Z \end{bmatrix}, \tag{9.40}$$

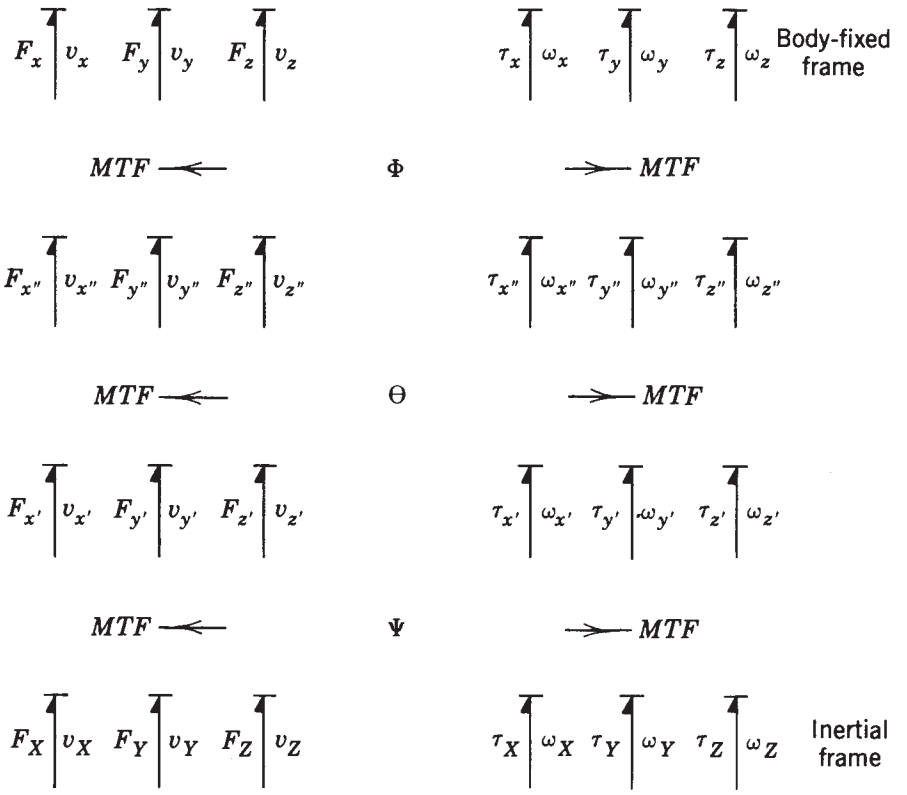


FIGURE 9.5. Bond graph of coordinate transformation.

or

$$\begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} = \Phi^t \Theta^t \Psi^t \begin{bmatrix} F_X \\ F_Y \\ F_Z \end{bmatrix}, \tag{9.41}$$

$$\begin{bmatrix} \tau_x \\ \tau_y \\ \tau_z \end{bmatrix} = \Phi^t \Theta^t \Psi^t \begin{bmatrix} \tau_X \\ \tau_Y \\ \tau_Z \end{bmatrix}. \tag{9.42}$$

Thus, the transformation indicated in Figure 9.5 not only transforms the body-fixed velocities and angular velocities to inertial components; it also transforms inertial forces and torques into the body-fixed directions. This bilateral transformation is indicated by the causality in Figure 9.5.

Let us agree to always use the coordinate transformation of Figure 9.5 with the causality shown. With this agreement we can now connect the bond graph of Figure 9.5 to the rigid-body dynamics of Figure 9.3, and we now have the capability of formulating a complete state representation of 3-D rigid-body dynamics.

This model can accept forces and torques in inertial coordinate directions, properly align these efforts into the instantaneous principal directions, and then, through integration of first-order equations, deliver velocities and angular velocities in principal directions and transform these to inertial directions. This is a very nice package.

We should note that the angles ϕ, θ, and ψ are needed to perform the coordinate transformations. It is relatively simple to relate $\dot{\phi}$, $\dot{\theta}$, and $\dot{\psi}$ to the body-fixed components ω_x, ω_y, and ω_z. From Figure 9.4,

$$\omega_x = \dot{\phi} - \dot{\psi}\sin\theta, \tag{9.43}$$

$$\omega_y = \dot{\theta}\cos\phi + \dot{\psi}\cos\theta\sin\phi, \tag{9.44}$$

$$\omega_z = -\dot{\theta}\sin\phi + \dot{\psi}\cos\theta\cos\phi. \tag{9.45}$$

Equations (9.44) and (9.45) can be solved for $\dot{\theta}$ and $\dot{\psi}$:

$$\dot{\theta} = \cos\phi\,\omega_y - \sin\phi,\omega_z, \tag{9.46}$$

$$\dot{\psi} = \frac{\sin\phi}{\cos\theta}\omega_y + \frac{\cos\phi}{\cos\theta}\omega_z. \tag{9.47}$$

Then, from (9.43),

$$\dot{\phi} = \omega_x + \sin\phi\frac{\sin\theta}{\cos\theta}\omega_y + \cos\phi\frac{\sin\theta}{\cos\theta}\omega_z. \tag{9.48}$$

Equations (9.46), (9.47), and (9.48) are three additional state equations that must be integrated along with the others so that ϕ, θ, and ψ are continuously available as modulating variables.

A very nice shorthand notation for the rigid-body mechanics and coordinate transformation is shown in Figure 9.6. For computational purposes, the building blocks of Figure 9.6 could not be better organized, or more ready for interaction with external dynamic systems. However, remember our agreement. The causality shown in Figures 9.3 and 9.6 must be maintained.

9.2 KINEMATIC NONLINEARITIES IN MECHANICAL DYNAMICS

This section is addressed to the following problem: Given a schematic representation of a mechanical system plus accompanying verbal and mathematical descriptions of the parts, find a convenient way to model the system so as to predict its dynamic behavior.

The importance of the problem hardly needs emphasizing. It has been worked on in essentially modern form at least since the time of Newton and has attracted the attention of many notable scientists, including Hamilton and Lagrange. The development since the time of Newton has relied, not upon pictorial or graphical

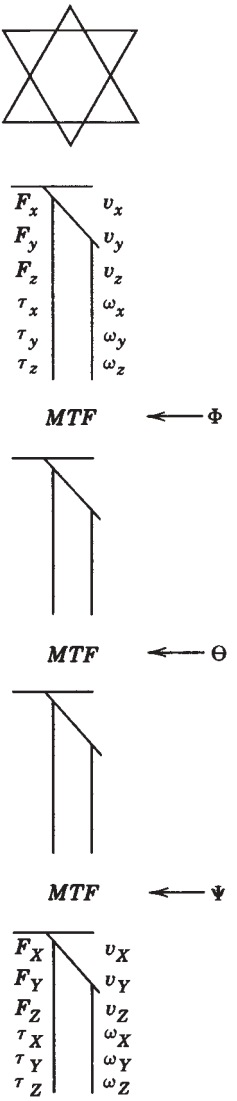


FIGURE 9.6. Shorthand bond graph for the complete 3-D mechanics plus coordinate transformation.

representation, but upon analytically oriented notation in the form of operators and equations of various types. Since we now know how to formulate mechanics problems in a variety of ways, it is fair to say that the basic problem has been solved [1, 2].

The principal purpose of this section is to present and develop the use of certain bond graph forms as standard models in mechanics, thereby bringing the study

of this very important class of generally nonlinear problems into the multiport-systems pattern. We shall study mechanical systems involving the large-scale motion of particles and rigid bodies. A variety of techniques are used, including the selection by the modeler of key variables for formulation, the determination of certain required transformations by simple analytic means, and the combination of all the parts into a unified representation by a bond graph.

The development begins with the simplest type of system, involving straightforward coupling between inertial and compliance elements.

9.2.1 The Basic Modeling Procedure

Let us start by considering the result we would like to achieve. For a given problem in mechanics involving particles, rigid bodies, and elastic elements, we wish to obtain a set of first-order differential equations in terms of variables that yield physical insight into system behavior. We anticipate that the equations will be coupled and, if possible, explicit (i.e., one derivative in each equation), having the following form:

$$C\text{-field:} \qquad \mathbf{f}_C = \phi_C(\mathbf{q}_C), \tag{9.49}$$

$$I\text{-field:} \qquad \mathbf{v}_I = \phi_I(\mathbf{p}_I), \tag{9.50}$$

$$\text{Junction structure:} \qquad \dot{\mathbf{q}}_C = [\mathbf{T}_{CI}(\mathbf{q}_C)]\mathbf{v}_I \tag{9.51}$$

$$\dot{\mathbf{p}}_I = [-\mathbf{T}_{CI}^t(\mathbf{q}_C)]\mathbf{f}_C, \tag{9.52}$$

where $\mathbf{f}_C$ is the set of efforts defining the potential co-energy, $\mathbf{q}_C$ is the set of displacements defining the potential energy, $\mathbf{v}_I$ is the set of (inertial) velocities or flows defining the kinetic co-energy; $\mathbf{p}_I$ is the set of momenta defining the kinetic energy, and $\dot{\mathbf{q}}_C$ and $\dot{\mathbf{p}}_I$ are the time derivatives of $\mathbf{q}_C$ and $\mathbf{p}_I$, respectively.

Inspection of the equations indicates that Eq. (9.49) arises directly from the compliance elements' constitutive laws. The set of relations may be linear or nonlinear and coupled or decoupled according to the nature of the C-field. Equation (9.50) is obtainable directly from the inertia elements in the system. The vector $\mathbf{v}_I$ is defined with respect to an inertial frame and represents translational velocities of the centers of mass of rigid bodies and angular velocities with respect to nonrotating coordinate systems.

Typically, the most difficult aspect of mechanics is the generally nonlinear coupling of C-fields and I-fields introduced by the geometry of large-scale motions. Such coupling is represented in Eqs. (9.51) and (9.52) by the transformation array $\mathbf{T}_{CI}(\mathbf{q}_C)$ and its transpose, and in bond graphs by multiport displacement-modulated transformers as discussed in Section 7.5. For now, let us assume that each element of $\mathbf{T}_{CI}$ is, in principle, a scalar function of all the elements of the vector $\mathbf{q}_C$.

Before we discuss an example, let us observe that Eqs. (9.49)–(9.52) may be combined to eliminate $\mathbf{f}_C$ and $\mathbf{v}_I$, giving

$$\dot{\mathbf{q}}_C = [\mathbf{T}_{CI}(\mathbf{q}_C)]\Phi_I(\mathbf{p}_I), \tag{9.53}$$

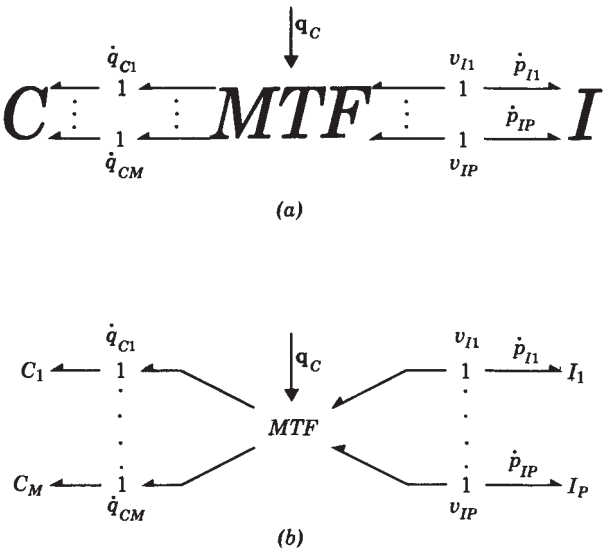


FIGURE 9.7. Symbolic bond graph model of basic nonlinear conservative system: (*a*) coupled C- and I-fields; (*b*) 1-port C- and I-fields.

$$\dot{\mathbf{p}}_I = [-\mathbf{T}^t_{CI}(\mathbf{q}_C)]\Phi_C(q_C). \tag{9.54}$$

These are the equations of a nonlinear conservative system expressed in terms of state variables $\mathbf{q}_C$ and $\mathbf{p}_I$. They can be modified easily to include nonconservative force effects.

Bond graph models representing the type of system discussed thus far are shown in Figure 9.7. The compliances are represented by the coupled C-field in part *a* and the "M" C-elements in part *b*. The inertias are represented by the coupled I-field in part *a* and the "P" I-elements in part *b*. The transformation coupling is represented by the M-port × P-port *MTF* element, and the 1-junctions are introduced for the sake of clarity and to allow a "through" sign convention for the *MTF*. Notice that the moduli of the *MTF* depend on the vector $\mathbf{q}_C$.

As an illustration of the type of formulation in which we are interested, consider the nonlinear oscillator of Figure 9.8*a*. The mass is free to slide on the rod and is constrained by a spring. A second spring constrains the rod as it rotates in the $X-Y$ plane. Key geometric variables are identified in part *b* as r, θ, x, y, v_x, and v_y. Constitutive relations for the linear springs and inertias, in the form of Eqs. (9.49) and (9.50), are given by Eqs. (9.55) and (9.56), respectively:

$$F = k_1 (r - R)\,, \tag{9.55a}$$

$$\tau = k_2\theta, \tag{9.55b}$$

$$v_x = m^{-1} p_x, \tag{9.56a}$$

$$v_y = m^{-1} p_y, \tag{9.56b}$$

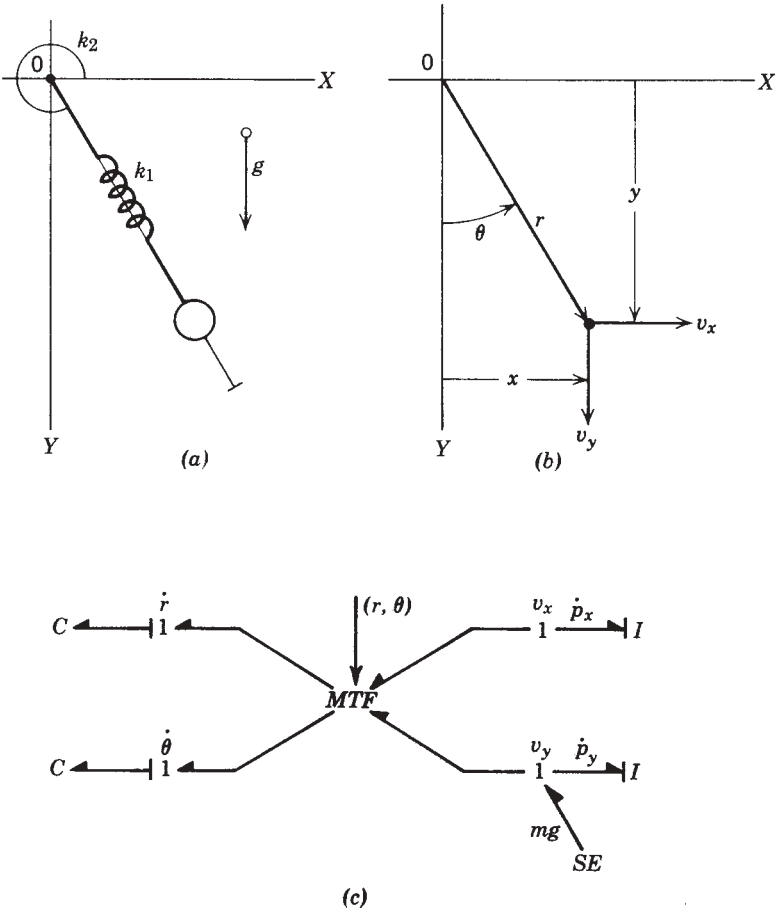


FIGURE 9.8. A nonlinear mechanical oscillator: (*a*) schematic diagram; (*b*) key geometric quantities defined; (*c*) bond graph model.

where R is the free length of the rod spring, F is the rod spring force, τ is the torque of the torsional spring on the rod about an axis through the origin normal to the $X-Y$ plane, and k_1 and k_2 are spring constants. Both the C-field and the I-field are linear and decoupled.

If we identify r and θ as elements of the vector $\mathbf{q}_C$, and v_x and v_y as elements of the vector $\mathbf{v}_I$, the array $\mathbf{T}_{CI}(\mathbf{q}_C)$ may be found by studying the velocity components as shown in Section 7.5,

$$\dot{r} = (\sin\theta)v_x + (\cos\theta)v_y, \tag{9.57a}$$

$$\dot{\theta} = \left(\frac{\cos\theta}{r}\right)v_x + \left(-\frac{\sin\theta}{r}\right)v_y, \tag{9.57b}$$

and

$$\mathbf{T}_{CI}(\mathbf{q}_C) = \begin{bmatrix} \sin\theta & \cos\theta \\ \dfrac{\cos\theta}{r} & -\dfrac{\sin\theta}{r} \end{bmatrix}. \tag{9.58}$$

It remains for us to calculate $\dot{\mathbf{p}}_I$ in terms of the spring forces. By appropriate resolution of the force F and torque effect τ in the directions of X and Y (i.e., p_x and p_y), we obtain

$$\dot{p}_x = (-\sin\theta)F + \left(-\frac{\cos\theta}{r}\right)\tau, \tag{9.59a}$$

$$\dot{p}_y = (-\cos\theta)F + \left(\frac{\sin\theta}{r}\right)\tau + mg. \tag{9.59b}$$

Careful inspection of Eqs. (9.59) shows that indeed the transformation array $\mathbf{T}_{CI}(\mathbf{q}_C)$ is embedded as the negative transpose, namely,

$$-\mathbf{T}^{\mathrm{t}}_{CI}(\mathbf{q}_C) = \begin{bmatrix} -\sin\theta & -\dfrac{\cos\theta}{r} \\ -\cos\theta & \dfrac{\sin\theta}{r} \end{bmatrix}. \tag{9.60}$$

The influence of gravity is merely added in the appropriate way (i.e., to directly influence p_y) and is shown in Figure 9.8*c* as an effort source, *SE*.

At this point, the thoughtful reader might well ask, "If we have obtained $\mathbf{T}_{CI}(\mathbf{q}_C)$ as in Eq. (9.58) once, is it really necessary to obtain it by separate development again, as in Eq. (9.60)?" The answer is no, and a systematic procedure that takes note of this result is the next topic. First we should combine Eqs. (9.57) and (9.59), eliminating F, τ, v_x, and v_y by Eqs. (9.55) and (9.56). The system state equations are

$$\dot{r} = (\sin\theta)m^{-1}p_x + (\cos\theta)m^{-1}p_y, \tag{9.61a}$$

$$\dot{\theta} = \left(\frac{\cos\theta}{r}\right)m^{-1}p_x + \left(-\frac{\sin\theta}{r}\right)m^{-1}p_y, \tag{9.61b}$$

$$\dot{p}_x = (-\sin\theta)k_1(r - R) + \left(-\frac{\cos\theta}{r}\right)k_2\theta, \tag{9.61c}$$

$$\dot{p}_y = (-\cos\theta)k_1(r - R) + \left(\frac{\sin\theta}{r}\right)k_2\theta + mg. \tag{9.61d}$$

A bond graph representation of the complete system is given in Figure 9.8*c*. The fields are shown explicitly, and the junction-structure transformation is implied by the two-by-two *MTF*. The minus signs in Eqs. (9.59) and (9.60) arise because of the sign conventions at the 1-junctions at the right side of Figure 9.8*c*.

Definitions of Key Geometric Quantities. As the first step in specifying a procedure for constructing bond graph models in mechanics, we must identify key variables. The approach of this chapter is geometric, meaning that we shall use displacement and velocity quantities to organize the system, with forces and momenta constrained by *MTF*s.

Two key vectors have already been described. They are $\mathbf{q}_C$, the vector of displacements that define the potential energy, and $\mathbf{v}_I$, the vector of velocities that define the kinetic co-energy. The vector $\mathbf{q}_C$ is directly associated with the C-field, and the vector $\mathbf{v}_I$ with the I-field.

If it were always possible to relate $\mathbf{v}_I$ to $\dot{\mathbf{q}}_C$ in terms of $\mathbf{q}_C$ as nicely as was done in the previous example, then mechanics problems would not be the *bête noire* they typically are. It is often the case that there are more velocities and displacements in the problem than the coupling constraints permit to be independent. To treat this situation, it will be useful to identify another vector, $\mathbf{q}_k$, the kinematic displacement vector or generalized coordinate vector in the language of Lagrangian mechanics. Basically, the elements of $\mathbf{q}_k$ are necessary and sufficient to fix the configuration of the system at any instant. Therefore, the vector $\mathbf{q}_C$ can be found in terms of $\mathbf{q}_k$, and the necessary velocity relations for $\dot{\mathbf{q}}_C$ can be suitably evaluated, as we show next. Furthermore, if the vector $\dot{\mathbf{q}}_k$ represents a necessary and sufficient set of velocities to fix all the motions at any instant, then the system is holonomic.* We shall consider only holonomic systems here, but a slight extension to the development would permit treatment of nonholonomic systems. If $\mathbf{q}_k$ describes all motions, then the vector $\mathbf{v}_I$ must be expressible in terms of $\dot{\mathbf{q}}_k$, given $\mathbf{q}_k$.

Therefore, the modeling process begins with the identification of three key geometric vectors: $\mathbf{q}_k$, the vector of kinematic displacements, or generalized coordinates; $\mathbf{q}_C$, the vector of C-field displacements; and $\mathbf{v}_I$, the vector of I-field velocities.

In Figure 9.8*b* the kinematic displacement vector could have been chosen to contain (r, θ) or (x, y), or even other pairs, such as (θ, x). Notice that if

$$\mathbf{q}_k = \begin{bmatrix} r \\ \theta \end{bmatrix},$$

then

$$\mathbf{q}_k = \mathbf{q}_C = \begin{bmatrix} r \\ \theta \end{bmatrix},$$

that is, $\mathbf{q}_k$ would be identical to $\mathbf{q}_C$. That would be convenient for many reasons. On the other hand, if

$$\mathbf{q}_k = \begin{bmatrix} x \\ y \end{bmatrix},$$

*See, for example, Reference [4].

then

$$\dot{\mathbf{q}}_k = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \mathbf{v}_I ,$$

that is, $\dot{\mathbf{q}}_k$ would be identical to $\mathbf{v}_I$. That would be convenient for many reasons, too. To a significant extent, the wise choice of $\mathbf{q}_k$ to balance the "needs" of $\mathbf{q}_C$ and $\mathbf{v}_I$ can ease the formulation of equations for a complicated system. In this regard, experience is the best teacher. It is an important aspect of the method being presented that a range of choices and some of their implications are made available in clear fashion.

As an example of a problem involving some significant choices, consider the spring–pendulum system shown in Figure 9.9*a*. The pivot point of the pendulum is constrained to slide on the Y-axis. The pendulum swings in the XY-plane. A number of geometric quantities are shown in Figure 9.9*b*.

The configuration of the system can be specified at any instant by the position of the pivot (y_k) and the angle of the pendulum (α). Other choices are possible, such as y_m and α. However, let us use

$$\mathbf{q}_k = \begin{bmatrix} y_k \\ \alpha \end{bmatrix} . \tag{9.62}$$

The potential energy of the system resides in the spring, and y_s is an obvious way to specify it. Hence,

$$\mathbf{q}_C = \begin{bmatrix} y_s \end{bmatrix} . \tag{9.63}$$

Finally, the kinetic co-energy is associated with the particle (the rod being assumed massless), so

$$\mathbf{v}_I = \begin{bmatrix} v_x \\ v_y \end{bmatrix} . \tag{9.64}$$

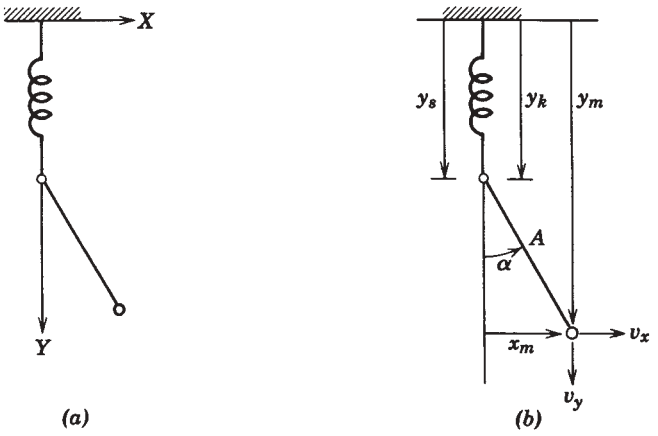


FIGURE 9.9. A mass–spring oscillator example: (*a*) schematic diagram; (*b*) key geometric quantities.

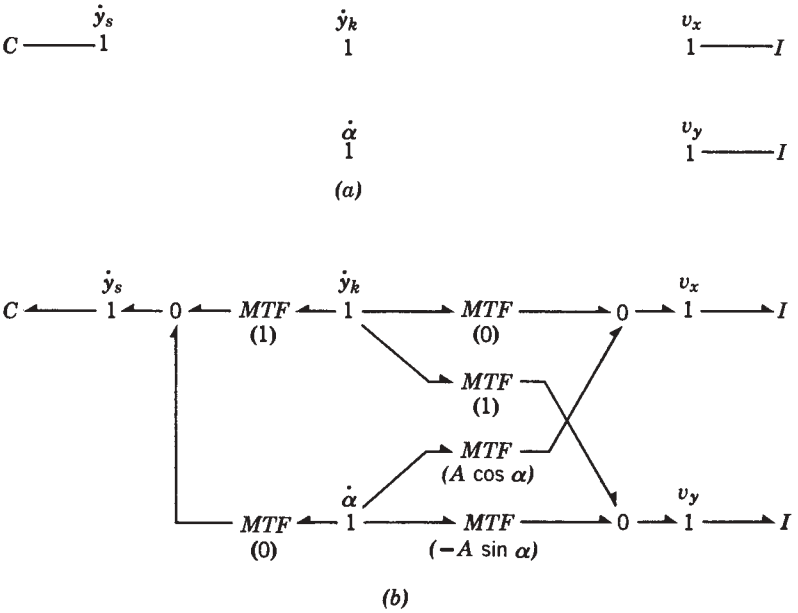


FIGURE 9.10. Explicit bond graph model for the mass–spring oscillator. (*a*) key geometric variable sets; (*b*) insertion of the transformations T_{Ck} and T_{Ik}.

A bond graph identifying the variables and the C- and I-fields is shown in Figure 9.10*a*. It will become standard practice to write a column of 1-junctions for $\dot{\mathbf{q}}_C$, a column for $\dot{\mathbf{q}}_k$, and a column for $\mathbf{v}_I$. The next steps are to develop transformation relations among the vectors and to represent them in the bond graph.

Calculating Velocity Transformations. One way to calculate velocity transformations is first to write displacement relations and then to differentiate them [5]. The relations between $\mathbf{q}_k$ and $\mathbf{q}_C$ may be written as

$$\mathbf{q}_C = \phi_{Ck}(\mathbf{q}_k), \tag{9.65}$$

where ϕ_{Ck} is a set of relations defining each element of $\mathbf{q}_C$ in terms of the elements of $\mathbf{q}_k$. By differentiating each relation in Eq. (9.65) with respect to time, we get

$$\dot{\mathbf{q}}_C = \frac{\partial \phi_{Ck}(\mathbf{q}_k)}{\partial \mathbf{q}_k}\dot{\mathbf{q}}_k = \mathbf{T}_{Ck}(\mathbf{q}_k)\dot{\mathbf{q}}_k. \tag{9.66}$$

The ith relation is

$$\dot{q}_{Ci} = \sum_{j=1}^{N} \frac{\partial \phi_{Cki}}{\partial q_{kj}}\dot{q}_{kj}, \tag{9.67}$$

where there are N displacements in $\mathbf{q}_k$.

For example, in the pendulum problem of Figure 9.9, $\mathbf{q}_C$ is related to $\mathbf{q}_k$ as follows, where the vectors are given by Eqs. (9.62) and (9.63):

$$y_s = 1y_k + 0\alpha; \tag{9.68}$$

then

$$\dot{y}_s = 1\dot{y}_k + 0\dot{\alpha} \tag{9.69}$$

and

$$\mathbf{T}_{Ck}(q_k) = [1 \quad 0]. \tag{9.70}$$

In this case, $\mathbf{T}_{Ck}$ does not depend on $\mathbf{q}_k$ explicitly.

As the next step, we express the inertia–velocity vector $\mathbf{v}_I$ in terms of $\mathbf{q}_k$ and $\dot{\mathbf{q}}_k$, thereby obtaining $\mathbf{T}_{Ik}(\mathbf{q}_k)$. Generally, it is easiest to do this by first creating an inertia–displacement vector $\mathbf{q}_I$, chosen so that

$$\dot{\mathbf{q}}_I = \mathbf{v}_I. \tag{9.71}$$

Then, we may write

$$q_I = \phi_{Ik}(\mathbf{q}_k), \tag{9.72}$$

from which we can calculate $\mathbf{v}_I$ as

$$\mathbf{v}_I = \dot{\mathbf{q}}_I = \left(\frac{\partial \phi_{Ik}}{\partial \mathbf{q}_k}\right)\dot{\mathbf{q}}_k = [\mathbf{T}_{Ik}(\mathbf{q}_k)]\dot{\mathbf{q}}_k. \tag{9.73}$$

The ith relation is

$$v_{Ii} = \sum_{j=1}^{N}\left(\frac{\partial \phi_{Iki}}{\partial q_{kj}}\right)\dot{q}_{kj}, \tag{9.74}$$

where there are N elements in $\mathbf{q}_k$.

In the pendulum example, we may write

$$\mathbf{q}_I = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad \dot{\mathbf{q}}_I = \mathbf{v}_I = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \end{bmatrix},$$

so $\mathbf{q}_I$ is related to $\mathbf{q}_k$ by

$$x = A\sin\alpha, \tag{9.75a}$$

$$y = y_k + A\cos\alpha. \tag{9.75b}$$

From Eqs. (9.75) we obtain T_{Ik} by differentiating, namely,

$$v_x = \dot{x} = (0)\dot{y}_k + (A\cos\alpha)\dot{\alpha}, \tag{9.76a}$$

$$v_y = \dot{y} = (1)\dot{y}_k + (-A\sin\alpha)\dot{\alpha}, \tag{9.76b}$$

and

$$\mathbf{T}_{Ik}(\mathbf{q}_k) = \begin{bmatrix} 0 & A\cos\alpha \\ 1 & -A\sin\alpha \end{bmatrix}. \tag{9.77}$$

In this case, the transformation depends on one element of $\mathbf{q}_k$, that is, α.

Establishing the Junction Structure. Since the basic pattern of both the transformations $\mathbf{T}_{Ck}$ and $\mathbf{T}_{Ik}$ is that the elements of $\dot{\mathbf{q}}_k$ are multiplied by constants or functions of $\mathbf{q}_k$ and added together, we may represent these transformations by 0-junctions and 2-port *MTF*s. The *MTF*s are used to multiply by the appropriate function, and the 0-junctions are used to add up velocity terms, as was discussed in Section 7.5.

Referring to the partial bond graph of Figure 9.10*a*, $\mathbf{T}_{Ck}$ and $\mathbf{T}_{Ik}$ of Eqs. (9.70) and (9.77) are introduced in the graph of Figure 9.10*b* between the appropriate sets of 1-junctions. We now have a basic bond graph model of the spring–pendulum system of Figure 9.9*a*.

Now consider that we wish to add two more dynamic effects—dissipation in the pivot joint and a gravity force. The dissipation torque arises directly from the relative motion $\dot{\alpha}$ and is shown by the *R*-element appended to that junction in Figure 9.11. The gravity force acts on v_y and is appended directly to the v_y-junction as a constant source of force of magnitude mg. We observe that the fields *C* and *R* can be linear or nonlinear. The junction structure is often nonlinear due to the geometric coupling represented by *MTF* moduli. In practice, the difficulty of generating state equations depends somewhat upon the nature of the nonlinearities [6].

The bond graph model of Figure 9.11 has been reproduced in Figure 9.12*a* with *MTF*s having zero moduli eliminated. In part *b*, additional simplifications have been made, and causality has been added. In this case, we have been lucky and there are no algebraic loops or cases of derivative causality. The one peculiarity is that α is not a state variable in the *C*-field and yet is necessary because

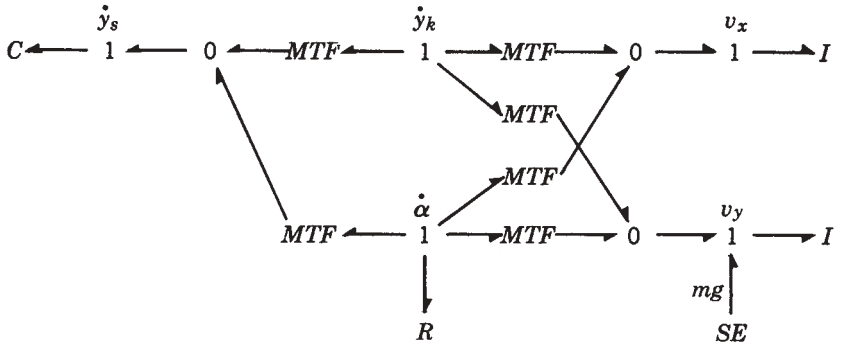


FIGURE 9.11. The mass–spring oscillator of Figure 9.4 with dissipation and gravity effects added.

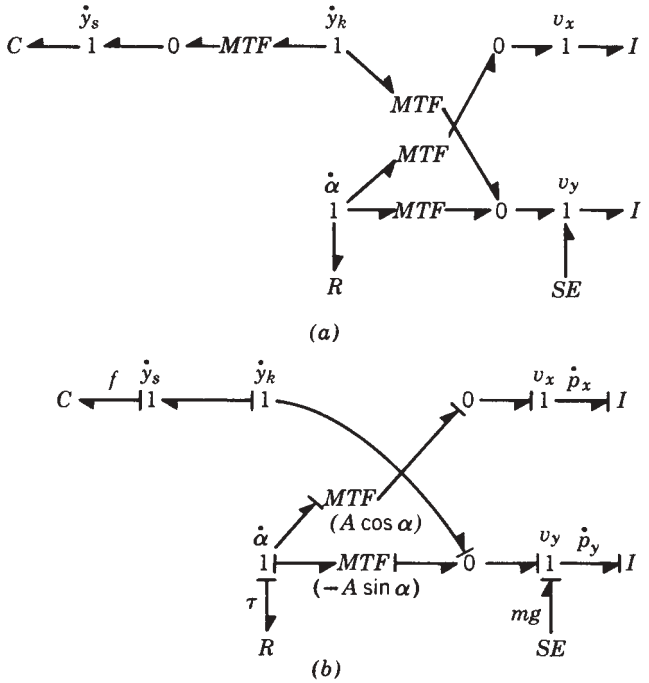


FIGURE 9.12. Simplification of the oscillator bond graph: (*a*) *MTF*s with zero modulus removed; (*b*) further simplification of the graph.

of the *MTF*s. The causal strokes show that $\dot{\alpha}$ can be found from p_x by inverting Eq. (9.76a). The C-, I-, and R-field relations are

$$f = k\,(y_s - Y) \qquad \text{(spring)}, \tag{9.78a}$$

$$v_x = m^{-1} p_x \qquad \text{(mass)}, \tag{9.78b}$$

$$v_y = m^{-1} p_y \qquad \text{(mass)}, \tag{9.78c}$$

$$\tau = R\dot{\alpha} \qquad \text{(dissipation)}, \tag{9.78d}$$

respectively, where Y is the free length of the spring in Eq. (9.78a), m is the mass in Eqs. (9.78b,c), and R of Eq. (9.78d) is the viscous dissipation parameter in the joint.

The connection or junction–structure relations are, in initial form,

$$\dot{p}_x = (-\tan\alpha)\, f - \frac{R}{A\cos\alpha}\dot{\alpha}, \tag{9.79a}$$

$$\dot{p}_y = -f + mg, \tag{9.79b}$$

$$\dot{y}_k = (\tan\alpha)\, v_x + v_y, \tag{9.79c}$$

$$\dot{\alpha} = \frac{1}{A\cos a} v_x. \tag{9.79d}$$

If we use Eqs. (9.78) to eliminate f, τ, v_x, and v_y from Eqs. (9.79), we get a momentum–displacement set of state equations:

$$\dot{p}_x = (-\tan\alpha)k(y_k - Y) - \frac{Rm^{-1}}{(A\cos\alpha)^2}p_x, \tag{9.80a}$$

$$\dot{p}_y = -k(y_k - Y) + mg, \tag{9.80b}$$

$$\dot{y}_k = (\tan\alpha)m^{-1}p_x + m^{-1}p_y, \tag{9.80c}$$

$$\dot{\alpha} = \frac{1}{A\cos\alpha}m^{-1}p_x. \tag{9.80d}$$

One could eliminate p_x in terms of v_x and v_y, to get a velocity–displacement set of equations,

$$\dot{v}_x = (-\tan\alpha)\frac{k}{m}(y_k - Y) - \frac{R}{m(A\cos\alpha)^2}v_x, \tag{9.81a}$$

$$\dot{v}_y = -\frac{k}{m}(y_k - Y) + g, \tag{9.81b}$$

$$\dot{y}_k = (\tan\alpha)v_x + v_y, \tag{9.81c}$$

$$\dot{\alpha} = \left(\frac{1}{A\cos\alpha}\right)v_x. \tag{9.81d}$$

Summary of the Procedure. The procedure for obtaining a bond graph model may be summarized in the following series of steps.

1. Identify the key vectors, $\mathbf{q}_k$, $\mathbf{q}_C$, and $\mathbf{v}_I$ (or $\mathbf{q}_I$). Write the 1-junctions corresponding to $\dot{\mathbf{q}}_k$, $\dot{\mathbf{q}}_C$, and $\mathbf{v}_I$.
2. Obtain the displacement transformation relating $\mathbf{q}_C$ to $\mathbf{q}_k$. Differentiate with respect to time to obtain a velocity transformation relating $\dot{\mathbf{q}}_C$ to $\dot{\mathbf{q}}_k$ in terms of $\mathbf{q}_k$. Then
$$\dot{\mathbf{q}}_C = \mathbf{T}_{Ck}(\mathbf{q}_k)\dot{\mathbf{q}}_k.$$
Write the results into the bond graph using *MTF* and 0-junction elements.
3. Obtain the velocity transformation relating $\mathbf{v}_I$ to $\dot{\mathbf{q}}_k$. Do this directly, or by relating $\mathbf{q}_I$ to $\mathbf{q}_k$ and differentiating with respect to time (being sure that $\dot{\mathbf{q}}_I = \mathbf{v}_I$). Then
$$\mathbf{v}_I = \mathbf{T}_{Ik}(\mathbf{q}_k)\dot{\mathbf{q}}_k.$$
Write the results into the bond graph using *MTF* and 0-junction elements.
4. Append the C-field and I-field elements to the $\dot{\mathbf{q}}_C$ and $\mathbf{v}_I$ junctions as indicated.
5. Append dissipation effects, force sources, and geometric constraints not yet included, as appropriate. If necessary construct additional velocities, using the transformational method.

6. Simplify the bond graph by removing *MTF*s with zero modulus, making *MTF*s with unit modulus into direct bonds, and combining 2-port junctions, when power directions permit.

The procedure outlined above is often successful at deriving a bond graph that will yield state equations fairly easily, but it also can lead to a bond graph with algebraic loops or derivative causality. For example, a little experimentation with the general bond graph for the example shown in Figure 9.11 will reveal that when integral causality is applied using the sequential causal assignment procedure, arbitrary causality must be assigned to some *MTF* bonds. This indicates an algebraic loop problem. This did not happen in Figure 9.12 because two zero-modulus *MTF*s could be eliminated.

A particularly vexing problem occurs whenever two rigid bodies are constrained in a nonflexible way. The constraint might be enforced by a pin joint between the bodies or the bodies might be forced to slide on one another in some way, for example. In any case, the degrees of freedom of the two bodies are reduced by the constraints, and this always leads to derivative causality which can be particularly difficult because geometric nonlinearities are typically involved.

In three dimensions, a rigid body has six degrees of freedom, represented by the six inertial elements in Figure 9.2. When two bodies are connected rigidly, the three linear and three angular velocities for the two bodies cannot all be independent, meaning that not all the I-elements can have integral causality. In the simpler case of plane motion, rigid bodies have only three degrees of freedom, but again whenever two bodies are linked, derivative causality will arise. Thus, it is important to realize that it is not enough to be able to assemble a correct bond graph for a mechanical system, but it is also important to have a scheme for drawing useful results from any bond graph model.

The field of multibody systems is concerned with exactly the problem of describing interconnected rigid bodies, and a variety of formulation and analysis techniques have been developed to treat these systems. Here we will just illustrate a variety of techniques that are particularly applicable to bond graph models of systems that contain constrained rigid bodies as subsystems. For more details on some of these techniques, see References [12–17].

9.2.2 Multibody Systems

The description and analysis of multibody systems have been the subject of research for a long time. When a number of rigid bodies are constrained by different types of joints and yet move through large angles in three dimensions, the kinematic relationships can become very complex. Specialists in multibody systems have developed a number of techniques for handling this type of mechanical dynamic system, and it would be presumptuous to suggest that bond graphs would always be a better choice as a descriptive mechanism than the alternative established techniques. Although bond graphs certainly can be used to study

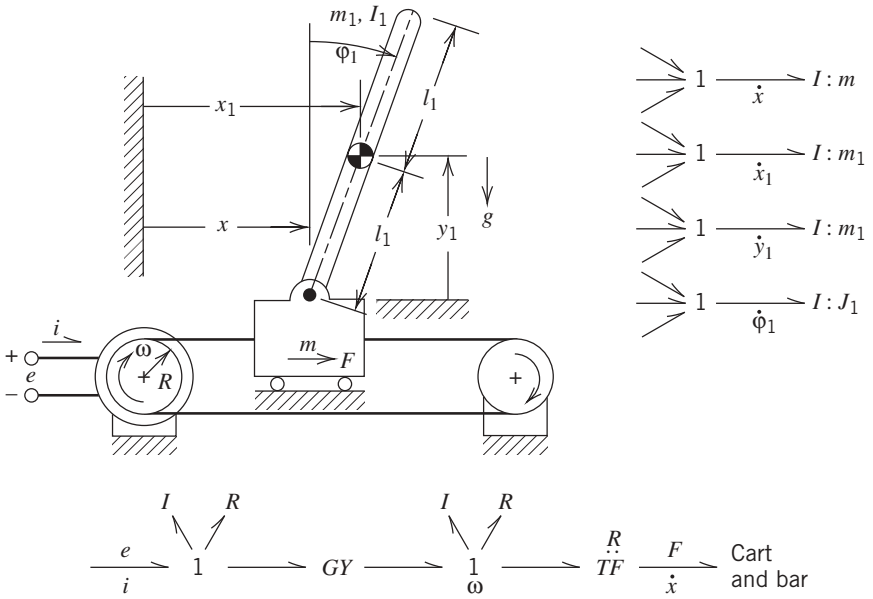


FIGURE 9.13. Control system experiment, including an upside-down pendulum mounted on a massive cart.

purely mechanical multibody systems (see Reference [17], for example), it is probably more logical to use a bond graph approach when the system to be studied contains other components that are particularly well treated using bond graphs.

As an elementary example, consider the system of Figure 9.13, which consists of an electric motor that moves a cart through a pulley and belt arrangement on that an upside-down pendulum is mounted. The object is to design a control system which varies the voltage on the motor in such a way that the pendulum is stabilized in the upright position at any desired cart position x.

As shown in the figure, the bond graph representation of the drive system is straightforward, but the rigid constraint between the cart and the bar will certainly cause derivative causality problems, which are complicated if the angle ϕ_1 is not restricted to be small. This system, neglecting any details of the drive, was one example used in Reference [12] to argue that the so-called descriptor form for multibody equations is particularly good at making the multibody system structure evident.

We will follow the development of Reference [14] to show that one bond graph approach to multibody systems is closely related to the multibody system approach leading to equations in the descriptor form. The first step is to consider bond graph elements representing the rigid-body dynamics without any consideration of the constraints among the bodies. The coordinates x for the cart and

x_1, y_1, and ϕ_1 for the bar are defined with respect to an inertial frame. Thus, the four I-elements at the right of Figure 9.13 properly represent the laws that state that the sums of all applied forces equal the rate of change of the linear momenta and that the velocity components of the centers of mass are the momenta divided by the masses. Also, the sum of the moments about the center of mass of the bar is equal to the rate of change of the angular momentum. For this plane motion case, the angular velocity divided by the centroidal moment of inertia gives the angular velocity of the bar. (As noted previously, the bond graph representing three-dimensional motion of unconstrained rigid bodies is considerably more complicated.)

To constrain the four I-elements in Figure 9.13 and thus to create a bond graph for the bar and cart subsystem to be attached to the drive train bond graph, we first define gaps δ_1 and δ_2 in Figure 9.14. These horizontal and vertical gaps between the bar pivot and the cart pivot should vanish. Furthermore, forces across the gaps λ_1 and λ_2 will be necessary in order to assure that δ_1 and δ_2 do indeed vanish. The use of λ to represent forces associated with constraints is intended to remind one that in the multibody formalism of Reference [12], these gap forces are Lagrange multipliers.

The gaps are easily related to the position coordinates of the rigid bodies in Figure 9.14:

$$\delta_1 = x_1 - x - \ell_1 \sin\phi_1, \qquad \delta_2 = y_1 - \ell_1 \cos\phi_1. \tag{9.82}$$

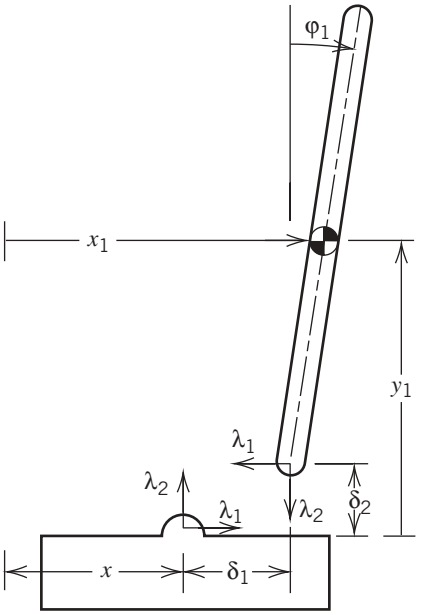


FIGURE 9.14. The bar and cart of Figure 9.13 showing gaps δ_1 and δ_2, which should vanish, and the associated forces λ_1 and λ_2.

By differentiating the gap expressions, one can relate the gap velocities to the body velocities:

$$\begin{bmatrix} \dot{\delta}_1 \\ \hline \dot{\delta}_2 \end{bmatrix} = \left[\begin{array}{c|c|c|c} -1 & 1 & 0 & -\ell_1 \cos\phi_1 \\ \hline 0 & 0 & 1 & \ell_1 \sin\phi_1 \end{array}\right] \begin{bmatrix} \dot{x} \\ \hline \dot{x}_1 \\ \hline \dot{y}_1 \\ \hline \dot{\phi}_1 \end{bmatrix}. \tag{9.83}$$

This relationship can be represented in bond graph terms by a multiport displacement-modulated transformer, as was discussed in Chapter 7. Figure 9.15 shows the transformer attached to the four I-elements representing the two unconstrained bodies.

The transpose of the matrix in Eq. (9.83) then automatically relates the gap forces λ_1 and λ_2 to effective forces at the four I-elements:

$$\begin{bmatrix} F_x \\ \hline F_{x_1} \\ \hline F_{y_1} \\ \hline \tau_1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ \hline 1 & 0 \\ \hline 0 & 1 \\ \hline -\ell_1 \cos\phi_1 & \ell_1 \sin\phi_1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \hline \lambda_2 \end{bmatrix}. \tag{9.84}$$

In Figure 9.15, the drive system force F and the gravity force $m_1 g$ have also been applied. Following the causal marks, some dynamic equations can be written:

$$\begin{aligned} \dot{p}_x &= F - F_x = F + \lambda_1, & \dot{x} &= p_x/m; \\ \dot{p}_{x_1} &= -F_{x_1} = -\lambda_1, & \dot{x}_1 &= p_{x_1}/m_1; \\ \dot{p}_{y_1} &= -m_1 g - F_{y_1} = -m_1 g - \lambda_2, & \dot{y}_1 &= p_{y_1}/m_1; \\ \dot{p}_{\phi_1} &= -\tau_1 = (\ell_1 \cos\phi_1)\lambda_1 - (\ell_1 \sin\phi_1)\lambda_2, & \dot{\phi}_1 &= p_{\phi_1}/J_1. \end{aligned} \tag{9.85}$$

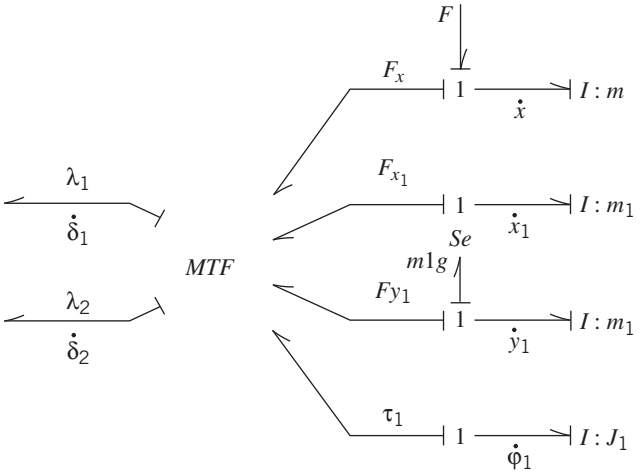


FIGURE 9.15. Bond graph relating gap velocities to body velocities.

The multibody descriptor form for the mechanical subsystem is just Eq. (9.85) slightly rewritten:

$$\begin{aligned} m\ddot{x} &= F + \lambda_1, \\ m_1\ddot{x}_1 &= -\lambda_1, \\ m_1\ddot{y}_1 &= -m_1 g - \lambda_2, \\ J_1\ddot{\phi}_1 &= (\ell_1 \cos\phi_1)\lambda_1 - (\ell_1 \sin\phi_1)\lambda_2. \end{aligned} \tag{9.86}$$

Equations (9.85) or (9.86) are straightforward except that the gap forces λ_1 and λ_2 are unknown. In the conventional multibody formulation, the extra information is supplied by requiring δ_1 and δ_2 to vanish in Eq. (9.82).

Although this formulation is mathematically complete, it represents a differential-algebraic set of equations, and such systems are not particularly easy to solve numerically. The bond graph of Figure 9.15 suggests an alternative approach. See References [16, 17].

Suppose that instead of requiring that δ_1 and δ_2 vanish exactly, we modeled the pivot interaction between the two bodies as if relatively stiff springs existed which would generate the λ forces whenever the gaps were not zero. In one sense, this may seem to be more realistic, since no truly rigid bodies exist. However, if realistic bearing stiffnesses were to be used, one would in all probability create a model with very high vibrational natural frequencies, which would necessitate short time steps in any numerical simulation. It is better philosophically to consider springs and dampers inserted across gaps as artificial devices to enforce constraints approximately and to do this with explicit differential equations instead of differential-algebraic equations. The idea is to experiment with spring stiffness to find the lowest stiffnesses that yield sufficiently small values of δ_1 and δ_2 in actual system operation. This will yield the longest possible time steps in simulation and the shortest simulation times.

The complete system bond graph for the system of Figure 9.13 is shown in Figure 9.16. In essence, the bond graph of the drive system shown in Figure 9.13 has been combined with the multibody bond graph of Figure 9.15 with the addition of *R*- and *C*-elements to generate the constraint forces λ_1 and λ_2 as a function of δ_1, δ_2, $\dot{\delta}_1$, and $\dot{\delta}_2$. Increasing the stiffness of the *C*-elements will have the effect of producing larger forces when the gaps are nonzero and thus reducing the gaps. Since the introduction of springs implies that vibrational motion is possible, it is useful also to introduce artificial *R*-elements to provide damping. Again, some experimentation with damping parameter values will be necessary to achieve a reasonable result.

The observant reader may have noticed that where the drive train bond graph of Figure 9.13 is attached to the multibody bond graph of Figure 9.15, an extra *C*-element has been inserted. The reason is that if the two bond graph fragments are simply joined, the motor rotary inertia and the cart inertia are coupled by a transformer relating rotary to linear motion variables. This is another instance of rigid coupling between inertia elements that always induces derivative causality.

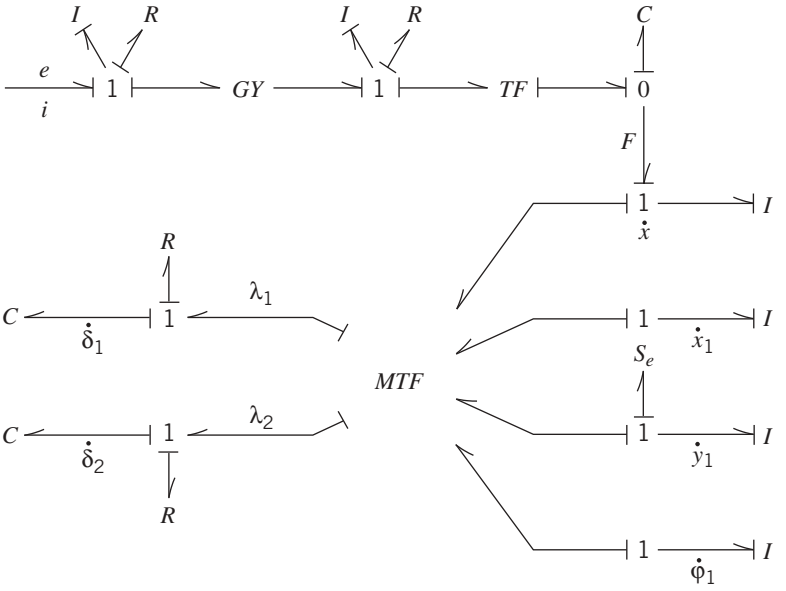


FIGURE 9.16. Bond graph for the system of Figure 9.13.

In this case, it would be simple to eliminate this inconvenience by defining a single equivalent rotary or linear inertia element or to define an I-field using the methods of Chapter 7. Instead, this rigid constraint has been replaced with an approximate one by defining an equivalent C-element. This element could be considered to represent the belt drive flexibility or simply a device for making the entire model capable of having integral causality.

Finally, one may note that it would be easy to write down the state equations for all the I- and C-elements in Figure 9.16. One must remember, however, that in this multibody technique, the displacement-modulated transformer uses the displacements of inertia elements to express the constraint or gap equations. In bond graphs, displacement variables for inertia elements are not normally necessary as state variables, so some extra state equations will be necessary.

The transformer matrices for the example shown in Eqs. (9.83) and (9.84) actually need one displacement, ϕ_1, so strictly speaking only the very last equation in Eq. (9.85) would need to be included. For this type of system, it would probably be useful to include state equations for all displacements x, x_1, y_1, and ϕ_1 in any case, as indicated in Eq. (9.85). These equations do increase the order of the system, but their inclusion has only a minor effect on the simulation time for a computer study of the system since three of the four displacements are not coupled to the remainder of the dynamic system.

To illustrate how this method of enforcing constraints among interconnected bodies can be extended, we briefly consider a second bar attached to the system

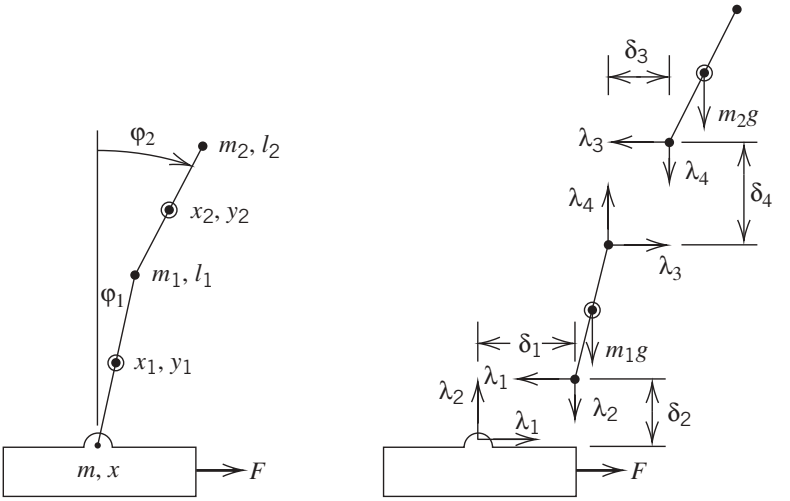


FIGURE 9.17. Double upside-down pendulum system.

already studied, as shown in Figure 9.17. By defining gaps δ_3 and δ_4 and corresponding forces λ_3 and λ_4 pertaining to the second set of constraints relating the position of the top of bar 1 to the bottom of bar 2, we can derive two more relations to supplement Eq. (9.82) which still applies:

$$\delta_3 = x_2 - \ell_2 \sin\phi_2 - x_1 - \ell_1 \sin\phi_1,$$

$$\delta_4 = y_2 - \ell_2 \cos\phi_2 - y_1 - \ell_1 \cos\phi_1. \tag{9.87}$$

A time differentiation of Eq. (9.87) results in equations relating $\dot{\delta}_3$ and $\dot{\delta}_4$ to $\dot{x}_1$, $\dot{y}_1$, $\dot{\phi}_1$, $\dot{x}_2$, $\dot{y}_2$, and $\dot{\phi}_2$, analogous to Eqs. (9.83). The transpose of the matrix relates six force components to λ_3 and λ_4, and both sets of relations are embodied in a 2×6 port modulated transformer, as shown in Figure 9.18. In this case, both bars have three degrees of freedom, so six generalized velocities are involved when the two bodies are constrained.

Again, there is the possibility of relating λ_3 and λ_4 to forces acting on the I-elements and requiring δ_3 and δ_4 to vanish in Eq. (9.87), which completes the multibody formulation in descriptor form. Figure 9.18 shows the alternative of using spring and damping elements to enforce the constraint in an approximate manner.

The idea of using artificial C- and R-elements to enforce constraints and thus to avoid derivative causality or differential-algebraic equations may appear to be a "brute-force" approach. This may be true, but first, a brute-force approach that is effective should not be discounted and, second, it has been argued that this approach is in many cases superior to the alternatives [17].

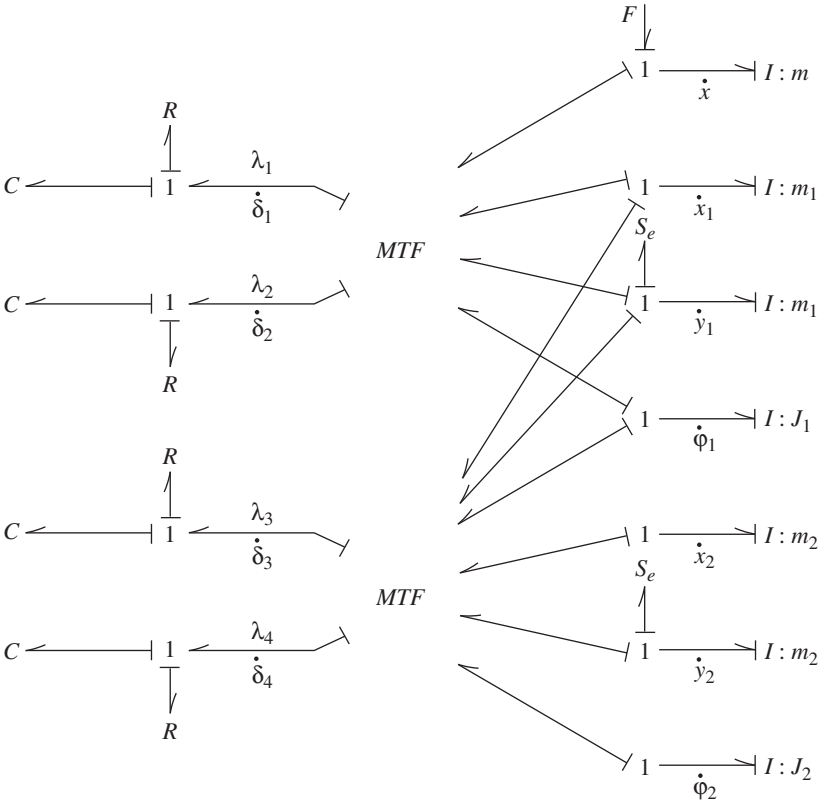


FIGURE 9.18. Bond graph for the double upside-down pendulum system.

9.2.3 Lagrangian or Hamiltonian *IC*-Field Representations

In Section 9.2.1, a procedure for constructing a bond graph for mechanical systems that could be described by generalized coordinates was presented. An example was worked out and illustrated in Figures 9.9–9.11. In this case, using normal bond graph causality, it was found that explicit differential equations could be written fairly easily despite geometric nonlinearities in the kinematic relationships involved. In many cases, however, the resulting bond graph, while formally correct, is intractable due to algebraic loops and derivative causality. Lagrange's equations are particularly useful for such situations, so in this section, it will be shown how a mechanical subsystem can be represented by a special type of *IC*-field, the equations of which are Lagrange's equations in a Hamiltonian form. This often is the simplest way to derive equations for a nonlinear mechanical subsystem.

Consider a holonomic system that can be described by a vector of generalized coordinates, $\mathbf{q}_k$, as described in Section 9.2.1. The kinetic energy T of all the

particles and rigid bodies in the system can then be written as a function of the variables in $\mathbf{q}_k$ and their rates of change, $\dot{\mathbf{q}}_k$. Then the standard form for Lagrange's equation is

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = E_i, \tag{9.88}$$

where q_i is the ith generalized coordinate in $\mathbf{q}_k$ (the subscript k has been dropped for simplicity since will be dealing only with the "kinematic" generalized coordinate displacements in this section); E_i includes all generalized forces for the ith coordinate, including those that could be derived from a potential energy function.

Equation (9.88) leads to coupled second-order equations, but a simple change leads to twice as many first-order equations. This Hamiltonian form not only is more useful for computation but also is more compatible with the bond graph formalism.

First, define generalized momenta p_i corresponding to the displacements q_i by the expression

$$p_i \equiv \frac{\partial T}{\partial \dot{q}_i}. \tag{9.89}$$

Then each equation in Eq. (9.88) may be rewritten in momentum form:

$$\dot{p}_i = \frac{\partial T}{\partial q_i} + E_i$$

or

$$\dot{p}_i = e_i' + E_i, \qquad e_i' \equiv \frac{\partial T}{\partial q_i}. \tag{9.90}$$

State equations for the q_i come by inverting Eq. (9.89) to solve for the $\dot{q}_i$ as functions of the p_i and q_i. As shown in Reference [15], Eq. (9.89) always has a special form. If $\mathbf{p}_k$ is the vector of generalized momenta, Eq. (9.89) can be represented thus:

$$\mathbf{p}_k = \mathbf{M}(\mathbf{q}_k, t)\dot{\mathbf{q}}_k + \mathbf{a}(q_k, t), \tag{9.91}$$

where $\mathbf{M}$ is a symmetric matrix and the vector $\mathbf{a}$ occurs only if the system includes time-varying velocity sources. Then the state equations for the q_i are formally

$$\dot{\mathbf{q}}_k = \mathbf{M}^{-1}(\mathbf{q}_k, t)[\mathbf{p}_k - \mathbf{a}(\mathbf{q}_k, t)]. \tag{9.92}$$

In the worst case, the inversion of the "mass" matrix $\mathbf{M}$ may have to be done repeatedly as a simulation proceeds, but in some cases it can be done only once—if $\mathbf{M}$ is a constant—or analytically if $\mathbf{M}$ is not completely coupled.

Note that Eqs. (9.90) and (9.92) are just generalizations of the type of equations we normally obtain for systems including I- and C-elements when bond graphs are used.

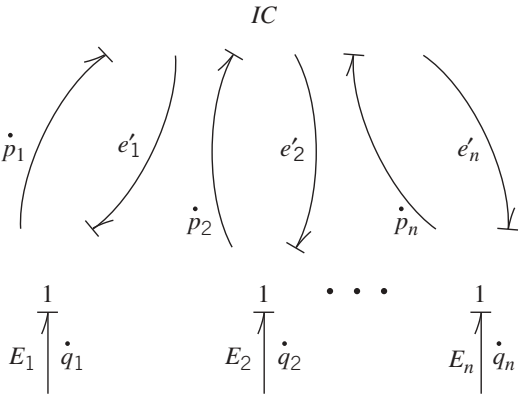


FIGURE 9.19. An *IC*-field representation of the Hamiltonian form of Lagrange's equations.

Equations (9.90) and (9.92) are elegantly summed up in the bond graph of Figure 9.19, which could be used to represent a complex mechanical subsystem connected to other elements readily described in bond graph terms. Of course, the e'_i force terms could even include forces derivable from potential energy, V, if e'_i in Eq. (9.90) were extended thus:

$$e'_i = \frac{\partial T}{\partial q_i} - \frac{\partial V}{\partial q_i}. \tag{9.93}$$

This would leave only nonpotential forces in the E_i. Often it is easy to put forces from elastic elements or gravity in the bond graph itself, and this may show system structure better than representing everything in the *IC* equations.

Although the equations for the *IC*-field may appear complex, one must remember that when many inertial elements are highly constrained, only a few degrees of freedom may be necessary to describe a subsystem that would otherwise have many cases of derivative causality. Furthermore, many cases of derivative causality are only local in nature so that a few *IC*-fields in combination with a number of elements in the bond graph with no causal problems may result in a useful model of a large and complex system.

An Example System. The system sketched in Figure 9.20 is elementary but yet shows a number of the difficulties associated with inertial elements in mechanics. It is not hard to create a bond graph for this system by considering the x, y, and z coordinates of the particle and the corresponding inertial space velocities $\dot{x}$, $\dot{y}$, and $\dot{z}$ as the flows on three simple I-elements. The torsion spring and a rotary inertia both have $\dot{\alpha}$ as their flow variable, and friction at the pendulum pivot has $\dot{\theta}$ as its flow variable.

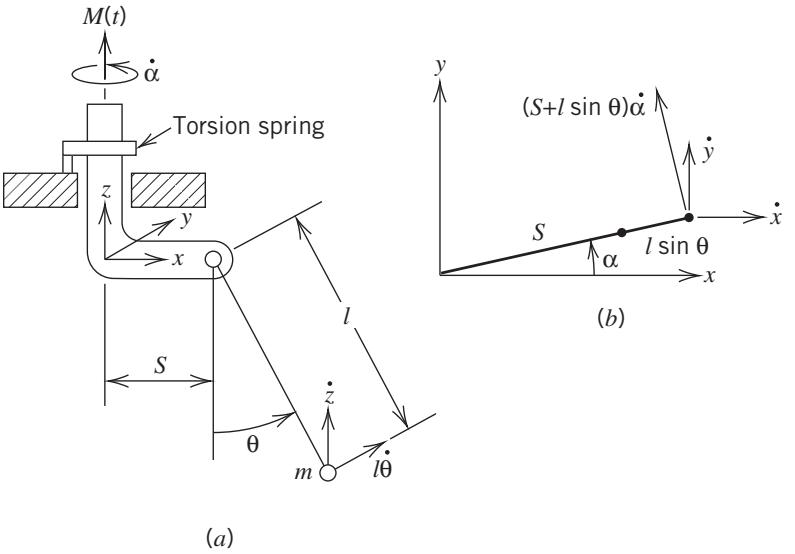


FIGURE 9.20. Example system: (*a*) three-dimensional view; (*b*) view from above.

An obvious way to enforce the geometric constraints in the problem is to consider α and θ as generalized coordinates which then determine x, y, and z:

$$\begin{aligned} x &= (S + \ell \sin\theta)\cos\alpha, \\ y &= (S + \ell \sin\theta)\sin\alpha, \\ z &= -\ell\cos\theta. \end{aligned} \tag{9.94}$$

Notice that time is not explicitly involved so that the **a** in Eqs. (9.91) and (9.92) will be missing.

The velocities $\dot{x}$, $\dot{y}$, and $\dot{z}$ are found by differentiating Eq. (9.94):

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -(S + \ell\sin\theta)\sin\alpha & \ell\cos\theta\cos\alpha \\ (S + \ell\sin\theta)\cos\alpha & \ell\cos\theta\sin\alpha \\ 0 & \ell\sin\theta \end{bmatrix} \begin{bmatrix} \dot{\alpha} \\ \dot{\theta} \end{bmatrix}. \tag{9.95}$$

Using these relationships, a bond graph can be assembled as shown in Figure 9.21. Either 2-port *MTF*s with 0- and 1-junctions can be used or a multiport *MTF* as shown with the 3×2 matrix in Eq. (9.95) as its transformation matrix. This bond graph is not easy to use, however, because if two I-elements are put into integral causality, the remaining I-elements will be forced into derivative causality. In the causality shown in Figure 9.21, $\dot{\alpha}$ and $\dot{x}$ are determined by I-elements in integral causality. By a manipulation of Eq. (9.95), $\dot{y}$, $\dot{\theta}$, and $\dot{z}$ could be related to $\dot{\alpha}$ and $\dot{x}$.

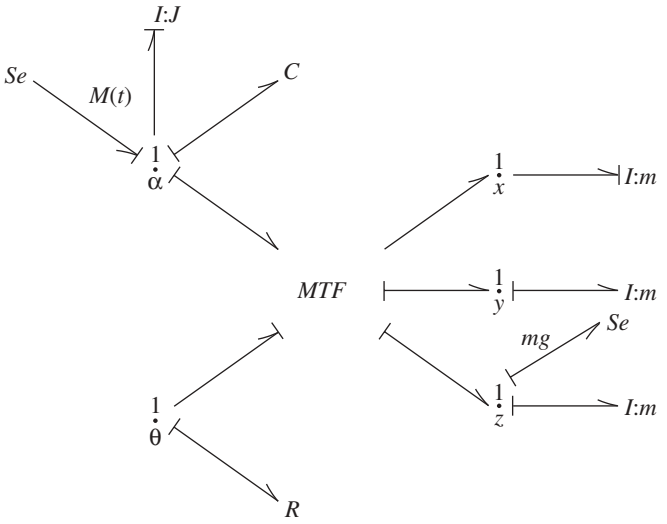


FIGURE 9.21. Bond graph for the system of Figure 9.20.

To develop an *IC*-field representation using α and θ as generalized coordinates, we first write the kinetic energy either with the help of Eq. (9.95) or, more simply, directly from the sketch of Figure 9.20:

$$T = \tfrac{1}{2}m\left\{[(S + \ell\sin\theta)\dot{\alpha}]^2 + (\ell\dot{\theta})^2\right\} + \tfrac{1}{2}J\dot{\alpha}^2, \tag{9.96}$$

where J is the moment of inertia about the z-axis.

The generalized momenta are

$$p_\alpha = \frac{\partial T}{\partial\dot{\alpha}} = m(S + \ell\sin\theta)^2\dot{\alpha} + J\dot{\alpha}, \tag{9.97}$$

$$p_\theta = \frac{\partial t}{\partial\dot{\theta}} = m\ell^2\dot{\theta}. \tag{9.98}$$

In this case, the inverted forms corresponding to Eq. (9.92) are simply

$$\dot{\alpha} = \frac{p_\alpha}{m(S + \ell\sin\theta)^2 + J}, \tag{9.99}$$

$$\dot{\theta} = \frac{p_\theta}{m\ell^2}. \tag{9.100}$$

The remaining equations of motion are

$$\dot{p}_\alpha = \frac{\partial T}{\partial\alpha} + E_\alpha = E_\alpha, \tag{9.101}$$

$$\dot{p}_\theta = \frac{\partial T}{\partial\theta} + E_\theta = m\dot{\alpha}^2(S + \ell\sin\theta)\ell\cos\theta + E_\theta. \tag{9.102}$$

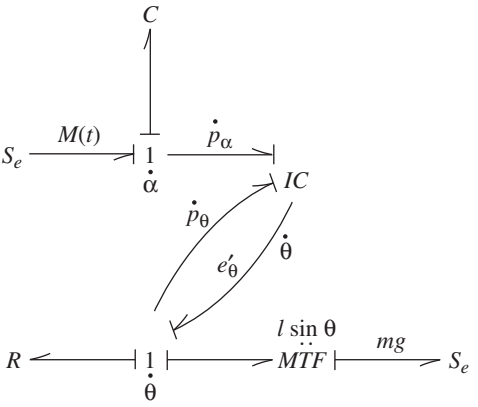


FIGURE 9.22. An *IC*-field representation for the example system of Figure 9.20.

Note that Eqs. (9.99)–(9.102) are essentially explicit first-order equations suitable for machine solution. One could use Eq. (9.99) to eliminate $\dot{\alpha}$ in Eq. (9.102) in favor of the state variables p_α and θ, but if the equations are integrated in the order given, this step is not really necessary in a computational sense since the numerical value of $\dot{\alpha}$ will be known before it is needed in Eq. (9.102).

Figure 9.22 shows a bond graph for the system in which an *IC*-field represents the mechanical subsystem and the S_e-, C-, and R-elements supply the generalized forces (torques) E_α and E_θ. Note that the gravity force acts through an *MTF*, which in Figure 9.21 is hidden in the multiport *MTF*. It is associated with the last line of Eq. (9.95). Also, since

$$e'_\alpha \equiv \frac{\partial T}{\partial \alpha} \equiv 0, \tag{9.103}$$

the bond conjugate to the $\dot{p}_\alpha$ bond is not needed, and, except for the torsion spring, α is not needed as a state variable.

Conclusions. Many problems with derivative causality can be solved by defining *I*- or *C*-fields for subsystems, as shown in Chapter 7, but mechanical inertia elements pose a particularly difficult set of problems when nonlinear geometric constraints are involved. The extension of the *I*-field idea to an *IC*-field, in the spirit of Lagrange's or Hamilton's equations, eliminates the derivative causality at the possible expense of an inertia matrix inversion which at worst may have to be made numerically and repeatedly during a computer simulation.

9.3 APPLICATION TO VEHICLE DYNAMICS

The bond graph fragment of Figure 9.6 is the fundamental building block for all rigid-body mechanics. Different transformations may be substituted, depending

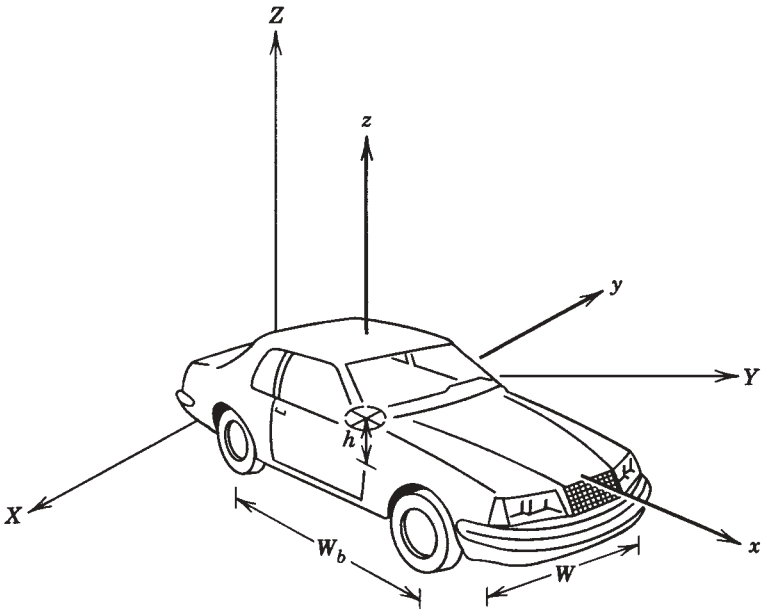


FIGURE 9.23. Ground vehicle with body-fixed frame.

on the rigid-body system being modeled, but the fundamental structure remains unchanged. In this section, the use of this building block for constructing sophisticated low-order models of ground vehicles is discussed.

Figure 9.23 shows a vehicle with a body-fixed frame attached at its center of mass and aligned with the principal directions. The vehicle body is considered rigid, with mass m_b and principal moments of inertia J_x, J_y, and J_z. It is a four-wheeled vehicle with pneumatic tires, front-wheel steering, front-wheel drive, and four-wheel independent suspension. It is a more or less conventional automobile. Through use of bond graphs, we can break the modeling effort into modeling of component pieces. We can then assemble the pieces, having been tested individually, into a system model.

Tire and Suspension. Pneumatic tires are extremely complicated in the way forces are generated at the contact patch. We will not attempt to develop this material thoroughly here and will keep the complexity within the context of demonstrating the procedure for building system models using 3-D rigid-body mechanics. The interested reader should see Reference [8] for a complete treatment of pneumatic tires.

Figure 9.24 shows a tire acted upon by forces at the ground contact patch and with an attached suspension unit. The tire is of mass m_t, which is representative of the unsprung mass associated with the wheel assembly, and the tire is compliant, indicated by the stiffness, k_t. It should be noted that the vertical compliance of

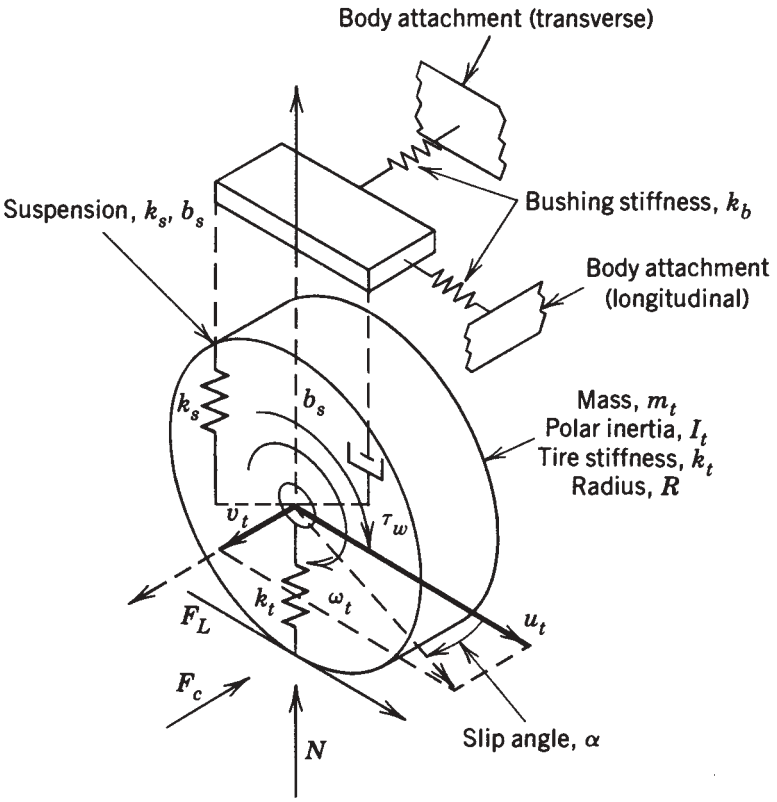


FIGURE 9.24. Schematic of tire and suspension.

a tire is nonlinear; however, for discussions here it is convenient to refer to a stiffness parameter. The realistic behavior can, of course, be included in the final computational model.

At the instant shown, the forward velocity of the tire is u_t, its transverse or sideways velocity is v_t, and its angular velocity is ω_t. It is acted upon by a longitudinal force F_L, a cornering force F_c, and a normal force N through the contact patch. The wheel torque τ_w is also shown in the figure, and it represents the torque from the drive train or torque due to braking.

The longitudinal force F_L is dependent upon the longitudinal *slip* of the tire. The slip s is expressed as a percentage and is calculated from

$$s = \frac{|R\omega_t - u_t|}{R\omega_t} \times 100 \tag{9.104}$$

during acceleration and

$$s = \frac{|u_t - R\omega_t|}{u_t} \times 100 \tag{9.105}$$

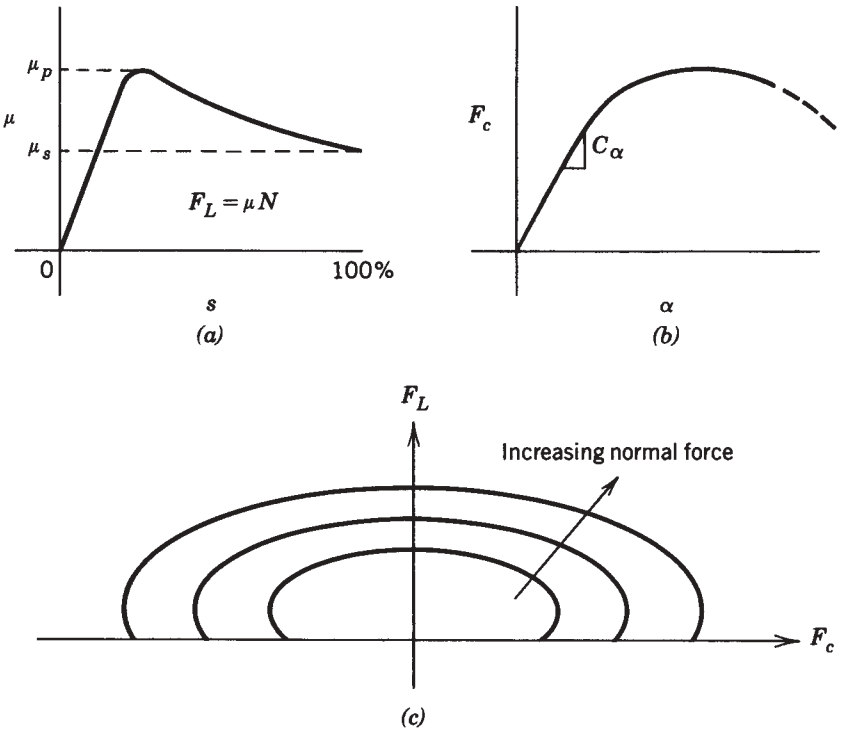


FIGURE 9.25. Qualitative force generation in pneumatic tires: (*a*) longitudinal force: (*b*) cornering force: (*c*) cross plot of longitudinal and cornering forces.

during braking. The slip determines the friction coefficient μ, through a functional dependence shown qualitatively in Figure 9.25*a*, and the force is given by

$$F_L = \mu N. \tag{9.106}$$

It is interesting that the tires generate a maximum force at some finite slip, and then the force decreases as the slip increases. When slip equals 100%, the wheel is spinning with no forward velocity (as on ice), or else the wheel is locked, but the vehicle is skidding with some forward velocity. The curve of Figure 9.25*a* is strongly dependent on road surface conditions and also the cornering force F_c.

When a tire is steered and the vehicle turns, the absolute velocity of the wheel center is not directed along the plane of the wheel. The tire deforms, and a side force F_c is developed. The slip angle α of the tire is defined to be the angle between the absolute velocity of the wheel and the plane of the wheel. This angle is shown in Figure 9.24 and can be calculated from

$$\alpha = \tan^{-1} \frac{v_t}{u_t}. \tag{9.107}$$

The cornering force is functionally related to α, as shown in Figure 9.25*b*. The linear part of the curve is characterized by the *cornering stiffness* C_α of the tire. The curve of Figure 9.25*b* is dependent on road surface conditions and on the longitudinal force. The trade-off between F_L and F_c is shown qualitatively in Figure 9.25*c*. For our discussion it is only important to realize that if α and s are both known, then F_L and F_c can be computed, regardless of the complexity of their interrelationship.

Figure 9.26 shows a bond graph of the wheel–tire system. The longitudinal dynamics are shown in part *a*, the cornering dynamics in part *b*, and the vertical dynamics in part *c*. The wheel is, in part, another rigid body that could be represented in a manner similar to the vehicle body. Figure 9.26 shows the wheel dynamics decoupled. For the wheel it is reasonable to neglect cross-coupling, since the angular velocity components about the vertical and horizontal longitudinal axes are very small. We are also neglecting gyroscopic effects due to the high angular velocity ω_t of the tire.

In Figure 9.26*a* the torque from the drive train or brakes is an input to the bond graph fragment, and a force on the body is an output. The modulated R-element represents the tire slip behavior and can be as complicated as necessary for the intended use of the model. In Figure 9.26*b* the cornering dynamics are represented. They too possess a modulated R-element for the cornering force generation. This fragment also outputs a force to the body. The vertical dynamics of Figure 9.26*c* allow for roadway unevenness under the tire contact patch, and this fragment also contains the suspension elements. A force on the body is an output from this fragment.

Notice in Figure 9.26 that integral causality exists for all models. Thus, state equations are straightforwardly derived. These equations will accept drive train torque τ as an input, and they will have forces as outputs. These forces then become inputs into the body model. Notice also that each wheel will have an identical but individual representation.

Drive Train Model. The final subsystem that requires discussion is the drive train, consisting of the engine, transmission, and differential (front only for this discussion). The drive train is not an essential element for simulating all vehicle systems. If we choose, we can simply specify the torque τ in the bond graph of Figure 9.26*a*. For completeness, we present the drive train model here.

Figure 9.27 shows, schematically, the drive train of a front-wheel drive vehicle. The engine power passes through a transmission and then is split by the differential and transmitted to the right and left front wheels. The differential applies equal torque, left and right, but permits the right and left wheels to have different angular velocities. A bond graph fragment for the drive train is also shown in Figure 9.27. In part *b*, no drive line compliance is included. The engine is modeled as a modulated torque source, and torques τ_L and τ_R are outputs from the model. In Figure 9.27*c*, drive line compliance is included. Torques are still outputs from the model. But this time the engine must be modeled as a modulated source of angular velocity. The output torques are, of course, the inputs to the wheel–tire dynamics of Figure 9.26.

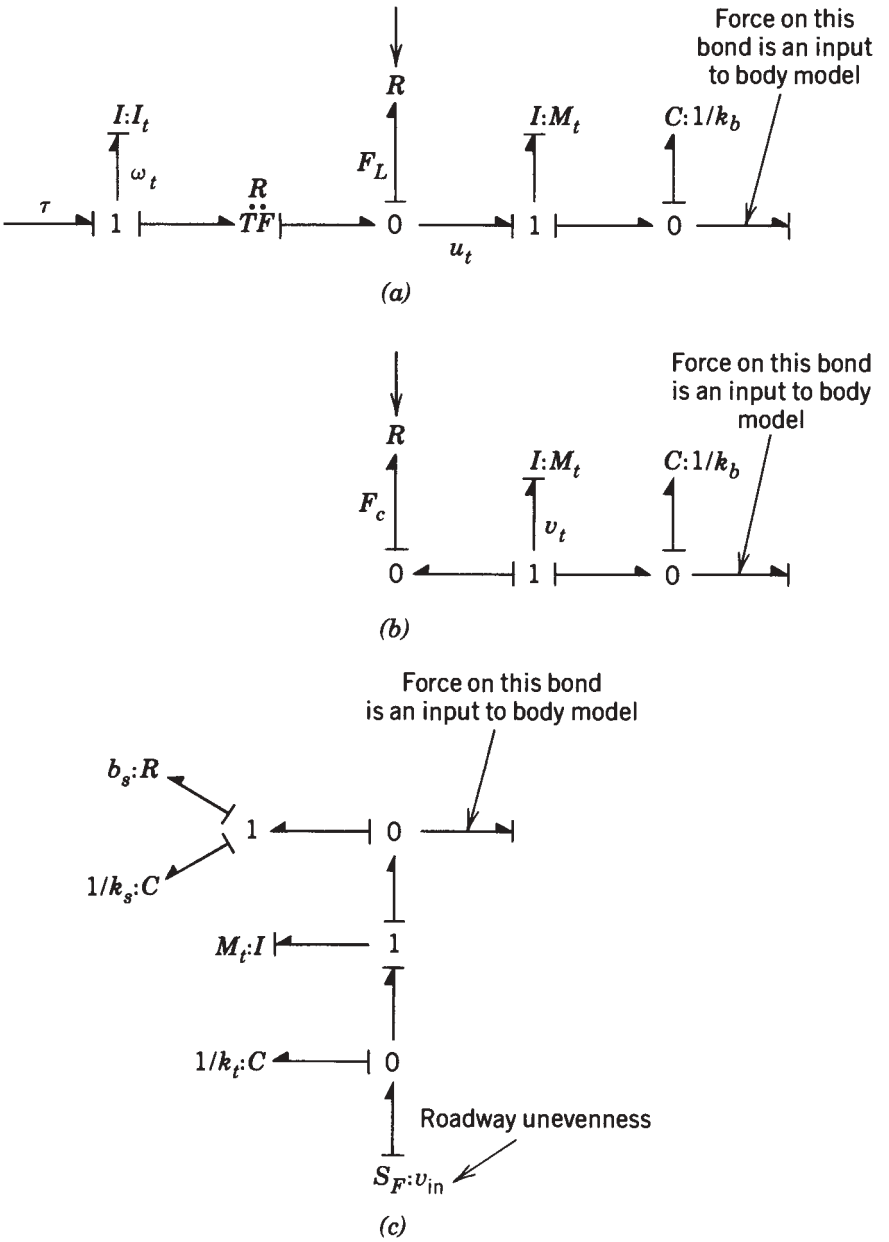


FIGURE 9.26. Bond graph of wheel–tire system: (*a*) longitudinal dynamics; (*b*) transverse dynamics; (*c*) vertical dynamics.

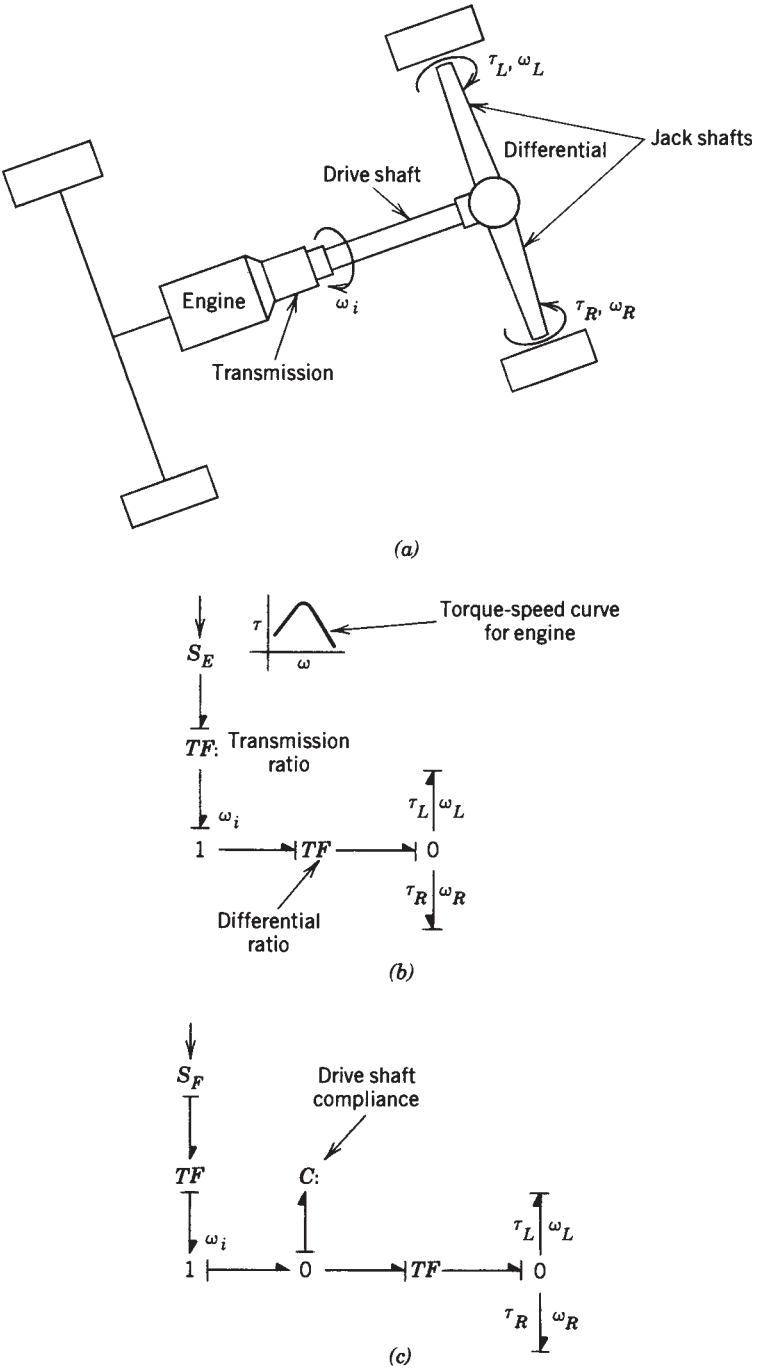


FIGURE 9.27. Drive train schematic and bond graph: (*a*) schematic; (*b*) bond graph without drive shaft compliance; (*c*) bond graph with drive shaft compliance.

9.4 SUMMARY

Three-dimensional rigid-body mechanics have been developed using bond graphs. A very elegant representation of Euler's equations resulted, but a prescribed causality must be used in order to allow the 3-D bond graph fragment to interact easily with a necessary coordinate transformation. Likewise, a complex coordinate transformation was presented which must be used with a prescribed causality for ease of coupling with other parts of an overall system model. The coordinate transformation shown here was the one particularly useful for vehicle dynamics.

The 3-D rigid-body development was demonstrated for vehicle dynamics applications. It was shown how wheel–tire models and drive train models could be developed separately from the 3-D mechanics and then made to interact in a computational model as long as prescribed causality was maintained. The complete model was not assembled here. However, it is hoped that sufficient detail was given to whet the reader's appetite so that uncontrollable hunger to read References [9–11] results. These references use this approach for constructing very complicated vehicle system models for use in design and control synthesis.

REFERENCES

[1] Sir William Thomson and P. G. Tait, *Principles of Mechanics and Dynamics*, New York: Dover, 1962.

[2] C. Lanczos, *The Variational Principles of Mechanics*, Toronto: Univ. of Toronto Press, 1957.

[3] R. C. Rosenberg, "Multiport models in mechanics," *Trans. ASME, J. Dyn. Syst. Meas. Control*, **94**, Ser. G, No. 3, 206–212 (Sept. 1972).

[4] S. Crandall, D. Karnopp, E. Kurtz, and D. Pridmore-Brown, *Dynamics of Mechanical and Electromechanical Systems*, New York: McGraw-Hill, 1968.

[5] D. Karnopp, "Power-conserving transformations: physical interpretations and applications using bond graphs," *J. Franklin Inst.*, **288**, No. 3, 175–201 (Sept. 1969).

[6] R. C. Rosenberg, "State-space formulation for bond graph models of multiport systems," *Trans. ASME, J. Dyn. Syst. Meas. Control*, **93**, Ser. G, No. 1, 35–40 (Mar. 1971).

[7] G. Martin, *Kinematics and Dynamics of Machines*, New York: McGraw-Hill, 1969, p. 7.

[8] H. B. Pacejka, "Introduction into the lateral dynamics of road vehicles," Third Seminar on Advanced Vehicle Dynamics, Amalfi, May 1986.

[9] D. C. Karnopp, "Bond graphs for vehicle dynamics," *Vehicle Syst. Dynam.*, **5** (1976).

[10] D. L. Margolis and J. Asgari, "Sophisticated yet insightful models of vehicle dynamics using bond graphs," ASME Symposium on Advanced Automotive Technologies, 89 ASME WAM, San Francisco, 1989.

[11] D. L. Margolis, "Bond graphs for vehicle stability analysis," *Int. J. Vehicle Design*, **5**, No. 4 (1984).

[12] B. Barasözen, P. Rentrop, and Y. Wagner, "Inverted n-bar model in description and in state space form," *Math. Model. Syst.*, **1**, 272–285 (1995).

[13] D. Karnopp, "Lagrange's equations for complex bond graph systems," *Trans. ASME J. Dynam. Syst. Meas. Control*, **99**, Ser. G, 300–306 (1977).

[14] D. Karnopp, "Understanding multibody dynamics using bond graphs," *J. Franklin Inst.*, **334B**, No. 4, 631–642 (1997).

[15] D. Karnopp, "An approach to derivative causality in bond graph models of mechanical systems," *J. Franklin. Inst.*, **329**, No. 1, 65–75 (1992).

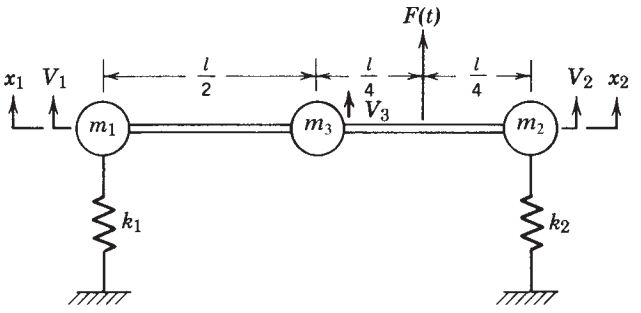
[16] D. Margolis and D. Karnopp, "Bond graphs for flexible multibody systems," *Trans. ASME*, *J. Dynam. Syst. Meas. Control*, **101**, 50–57 (1979).

[17] A. Zeid and C.-H. Chung, "Bond graph modeling of multibody systems: a library of three dimensional joints," *J. Franklin Inst.*, **329**, 605–636 (1992).

PROBLEMS

9-1. A particle can slide on a frictionless horizontal plane constrained by a linear spring pivoted at the origin of x, y and with spring constant k and free length l_0. Make a bond graph for this system by using r, θ for the vectors $\mathbf{q}_C$ and $\mathbf{q}_k$ and $\dot{x}$, $\dot{y}$ for the vector $\mathbf{v}_I$. Can the system accept all integral causality? Make another bond graph by using x, y for $\mathbf{q}_k$, $\dot{x}$, $\dot{y}$ for $\mathbf{v}_I$, and r for $\mathbf{q}_C$.

9-2. Three mass particles are attached to a spring-supported rigid massless bar that executes small vibratory motion. The vector $\mathbf{q}_C$ is x_1, x_2, the vector v_I is V_1, V_2, V_3, and several choices for $\mathbf{q}_k$ are possible. Show a bond graph for the system with the following choices for $\mathbf{q}_k$:

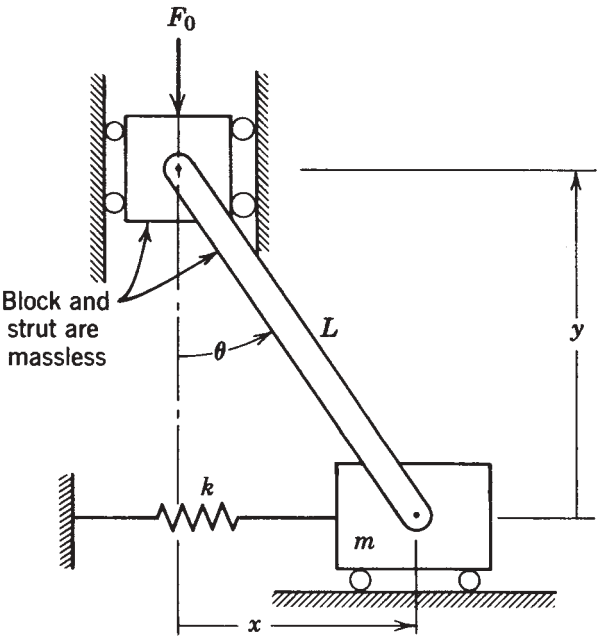


(a) $\mathbf{q}_k = x_1, x_2 = \mathbf{q}_C$;

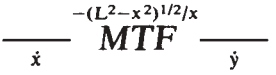
(b) $\mathbf{q}_k = x_3, \theta$, where x_3 is the displacement of m_3 and θ is the angle of inclination of the bar; and

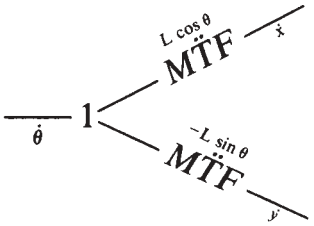
(c) $\mathbf{q}_k = x_1, x_3$.
Pick one bond graph, and write the equations of motion for the system using the graph.

9-3. Set up bond graph representations of the system shown below using the three alternative *MTF* forms shown. Write *complete* state equations in each case, using integration-causality methods or Lagrange equations. Comment on any advantages or disadvantages you see in the alternative formulations.

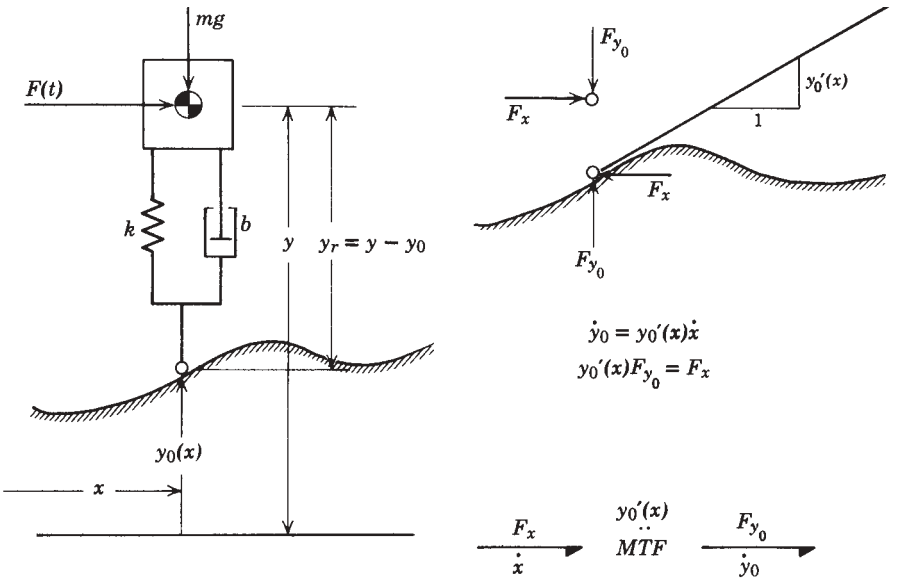


1. 

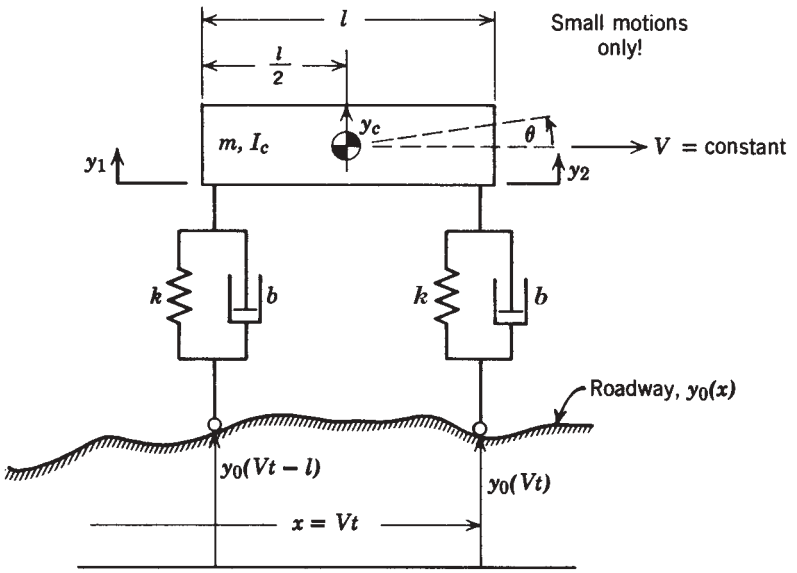
2. 

3. 

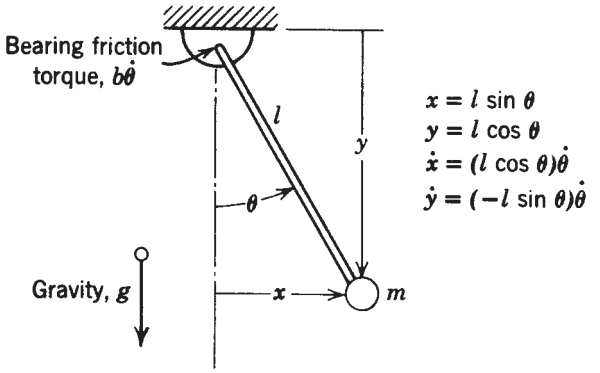
9-4. Consider the model of an off-road vehicle shown below. Note that the surface interaction may be represented by the equations and bond graph element shown. Develop a bond graph and state-space equations for this model.



9-5. Construct a bond graph for this two-degree-of-freedom vehicle model. Use two forms for the I-field representing the rigid body: one matrix representation using the $\dot{y}_1$, $\dot{y}_2$ port variables directly, the other using the auxiliary variables y_c and θ and a multiport TF.



9-6. (In which the inherent nastiness of geometric nonlinearity is exhibited.) Consider the simple pendulum of length l, mass m, swinging through large angles under the influence of gravity and bearing friction.



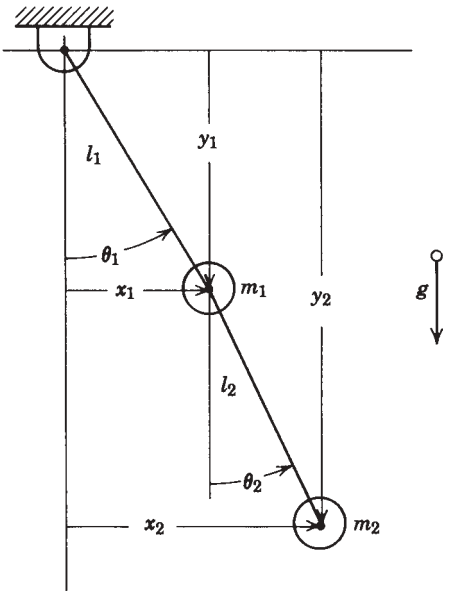
(a) Set up a modulated transformer representation of the relations between $\dot{x}$, $\dot{y}$, and $\dot{\theta}$ and the corresponding forces F_x, F_y and torque τ that are enforced by the rigid bar. Demonstrate that the multiport displacement-modulated transformer relations are power conservative.

(b) Show a bond graph for the system using your *MTF*. Set up sign conventions and a causal pattern in which integration causality is applied to the x-motion —I.

(c) Write the equations corresponding to your graph. Be sure that your state space is complete and in the standard form.

(d) Show that your state space is consistent with $ml^2\ddot{\theta} = -b\dot{\theta} - mgl \sin\theta$, which would be the result of many standard analyses.

9-7. Consider the double pendulum shown (see Problem 9-6).

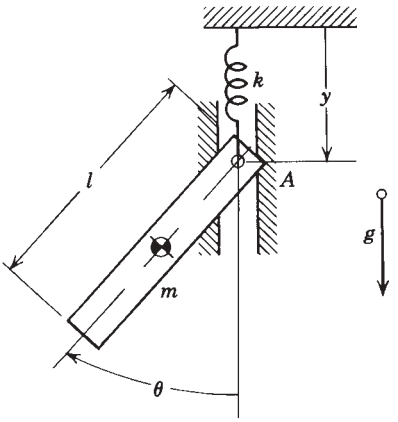


(a) Set up the multiport displacement-modulated transformer that relates $\dot{x}_1$, $\dot{y}_1$, $\dot{x}_2$, $\dot{y}_2$ to $\dot{\theta}_1$, θ_1, $\dot{\theta}_2$, θ_2, and exhibit both the velocity and force relations in matrix form.

(b) Using the *MTF*, show a system bond graph, and predict one possible set of state variables using the integration-causality method.

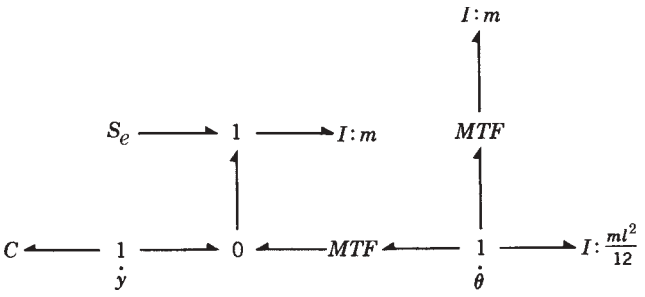
(c) Use the junction structure of (b) to compute $T(\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2)$—kinetic energy—and generalized torques so that Lagrange equations can be utilized as an alternative procedure to find state equations. What are the state variables if Lagrange equations are used?

9-8. The uniform thin bar shown below is pivoted at *A*. The pin slides freely in the vertical guideway. Represent the bar by three 1-port —*I*'s representing the horizontal and vertical motion of the center of mass and rotation about the center of mass.



(*a*)

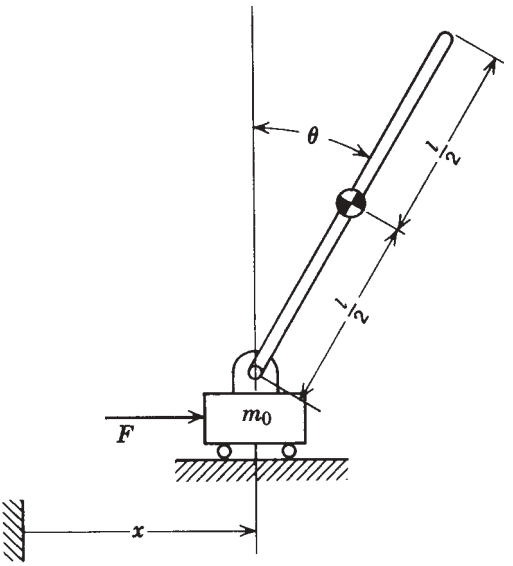
(a) Show that the junction structure of the multiport *MTF* involved may be rearranged to yield a bond graph for the system of the following form:



(*b*)

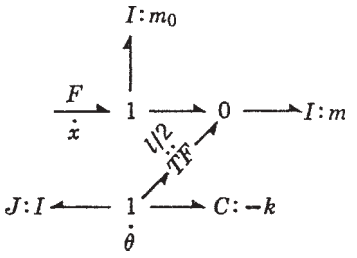
(b) Apply causality to the system, and write the equations of motion.

9-9. Let the pendulum have mass m and moment of inertia J. Assume $\theta \ll \pi$.



(*a*)

(a) Noting that the equilibrium point for the upside-down pendulum is $\theta = 0$, write expressions for kinetic energy and potential energy valid up to second order in θ and $\dot{\theta}$. Use the expressions to validate the linear bond graph shown below:



(*b*)

Find the negative spring constant for the —C that models the gravity effect.

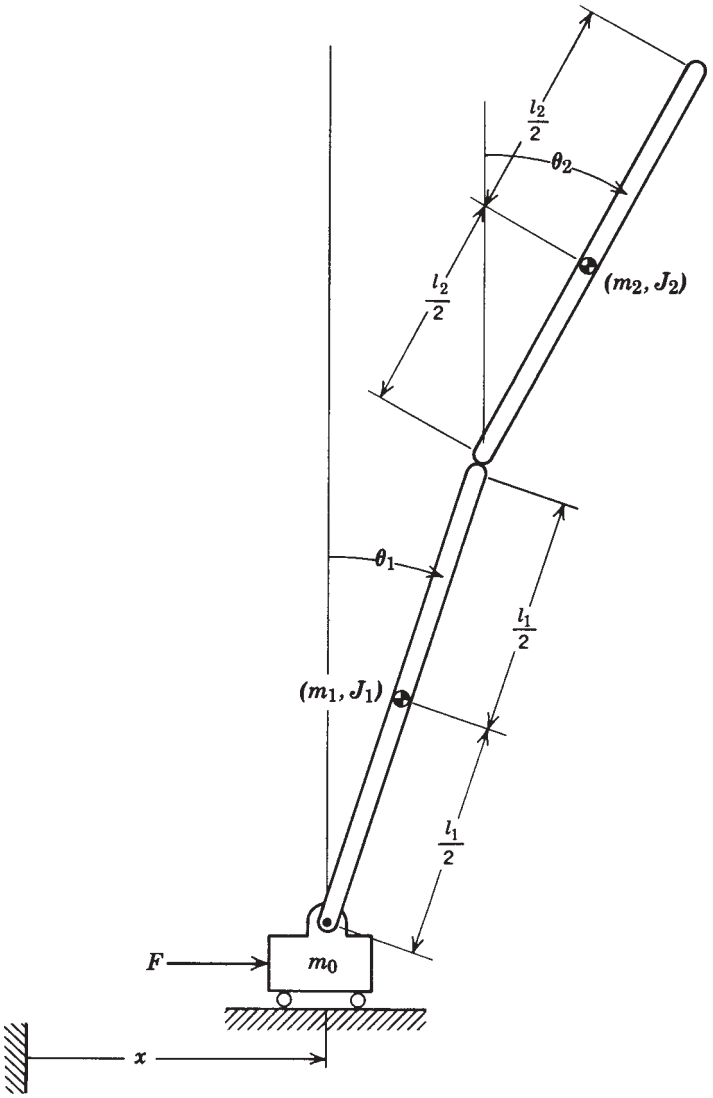
(b) Assuming F is an input, and that the —I's with parameters m_0 and J have integral causality, write state equations and indicate the matrix that must be inverted because of the differentiation causality.

(c) Show that the system can be represented by

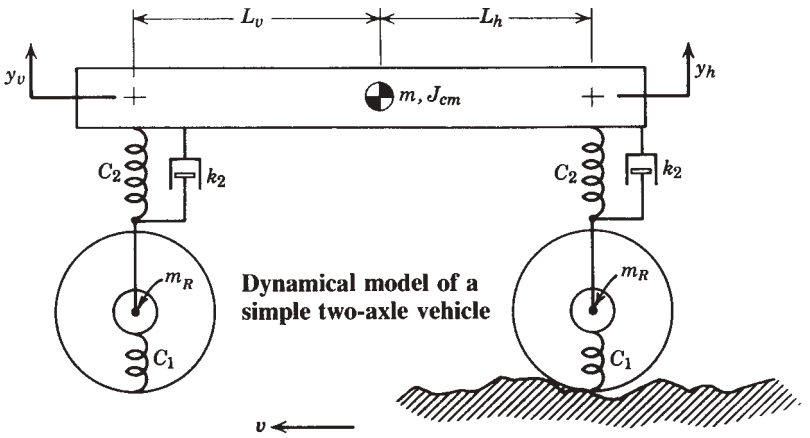
$$\underset{\dot{x}}{\overset{F=\dot{P}_x}{\rightarrow}} I \underset{\dot{\theta}}{\overset{\dot{P}_\theta}{\leftarrow}} C: + k,$$

and find the matrix representation of the I-field.

9-10. Following the pattern of analysis of Problem 9-9, find a bond graph for the double-inverted pendulum shown in the figure. Write state equations for the system.



9-11. The German-language automobile model shown below uses the subscripts, v, h, and R, for *vorn* (forward), *hinten* (rear), and *Rad* (wheel), respectively. Otherwise, it seems to be sufficiently international to be bond graphed.



Consider small pitch angles for the vehicle. Recall that the sum of the vertical forces is equal to the mass times the acceleration of the center of mass and that the sum of the torques about the center of mass is the centroidal moment of inertia times the angular acceleration.

(a) Construct a bond graph for the model in which the vehicle mass m and the centroidal moment of inertia appear as 1-port element parameters.

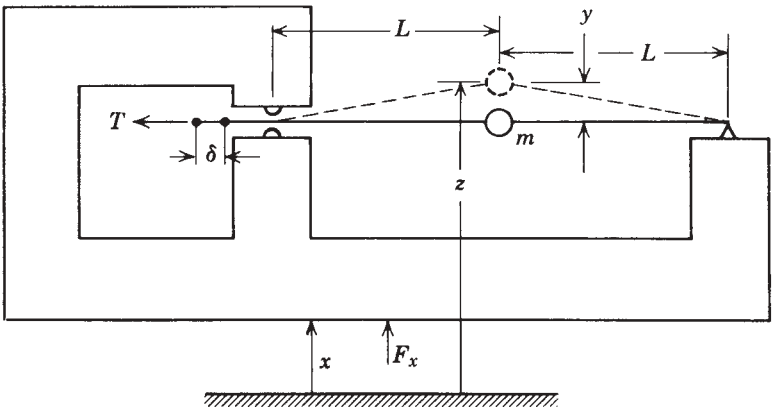
(b) Show that consistent causality may be assigned for your bond graph, and list the state variables. Discuss very briefly any difficulties you may find in assigning causality.

(c) Show that if the vehicle body is considered to be an I-field, the assignment of causality is simplified and the writing of state equations is easier.

(d) Either find the properties of the 2-port I-field describing the rigid body or outline a procedure to find them.

(e) Can you think of a way to represent the nonlinear behavior of the wheels when they leave the ground?

9-12. The mass m is supported on a thin inextensible wire stretched across the frame with tension T.



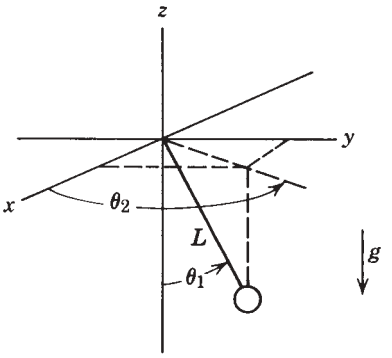
(a) Show that if $T = \text{const}$ and if $y/L \ll 1$, then the force on the mass is given approximately by $F = -(2y/L)T$, and a bond graph for the system (neglecting the inertia of the frame) is

$$k_{\text{eq}} = \frac{2T}{L} : C \leftarrow 1 \underset{\dot{y}}{\overset{F_y}{\rightharpoonup}} \overset{I}{0} \underset{\dot{x}}{\overset{F_z \mid \dot{z}}{\leftharpoondown}} F_x$$

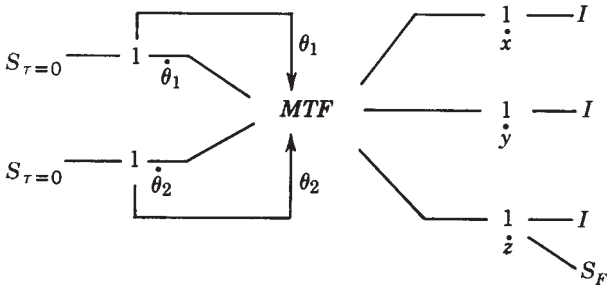
(b) When T is allowed to vary by connecting a force source to the wire, the equivalent spring seems to simply have a "variable constant." This concept is not very profound, however, since a —C element must be conservative, and with variable T the relation between F_y and y is not.

Show that if $y/L \ll 1$, the distance the end of the wire moves against the force source, δ, is $\delta = y^2/L$. Use this relation to show a bond graph that represents the effect of the wire by means of a force source and a displacement-modulated transformer. Note that this model shows a causal restriction that was not evident in the previous bond graph.

9-13. Consider the spherical pendulum shown in part a of the figure below. Verify the bond graph shown in part b by finding the constitutive laws for the *MTF* relating $\dot{\theta}_1$, $\dot{\theta}_2$ to $\dot{x}$, $\dot{y}$, $\dot{z}$. Write out the multiport *MTF* as a junction structure using 0- and 1-junctions and 2-port *MTF*s.



(a)



(b)

Now show that all three I's cannot accept integral causality, because the *MTF* cannot accept flow input causality on all three of the $\dot{x}$, $\dot{y}$, $\dot{z}$ bonds. Show that consistent complete causality can be achieved if only two of the I's have integral causality.

9-14. A brief outline of the relation between applied torque and angular momentum for general motion of a rigid body using components in a coordinate frame moving with the body and aligned with the principle axes is as follows: Following the notation of Reference [4], Section 4.4, a rigid body may be defined by a relation between angular velocity $\boldsymbol{\omega}$, angular momentum $\mathbf{H}$, and an inertia matrix $\mathbf{I}$, as follows:

$$\mathbf{I}\omega = \mathbf{H}. \tag{i}$$

In principal axis coordinates, this becomes

$$\begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix}. \tag{ii}$$

As long as torques are computed about a fixed point or the center of mass of the body, the following is true:

$$\frac{d\mathbf{H}}{dt} = \tau, \tag{iii}$$

where $\mathbf{H}$ is the angular momentum vector, and $\boldsymbol{\tau}$ is the torque vector. When the components of $\mathbf{H}$ and $\boldsymbol{\tau}$ are chosen in the moving coordinate frame, then Eq. (iii) becomes

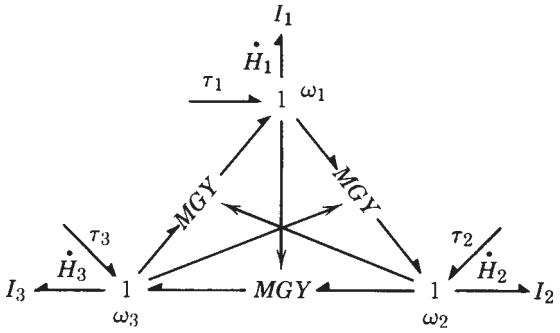
$$\frac{\partial \mathbf{H}}{\partial t_{\text{rel}}} + \omega \times \mathbf{H} = \tau \tag{iv}$$

or

$$\begin{bmatrix} \dot{H}_1 \\ \dot{H}_2 \\ \dot{H}_3 \end{bmatrix} + \begin{bmatrix} \omega_2 H_3 - \omega_3 H_2 \\ \omega_3 H_1 - \omega_1 H_3 \\ \omega_1 H_2 - \omega_2 H_1 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}, \tag{v}$$

in which the $\boldsymbol{\omega} \times \mathbf{H}$ term corrects the terms representing change relative to the moving frame to correctly portray the total change in $\mathbf{H}$ relative to the inertial frame.

Verify that the modulated gyrator–ring structure correctly represents Euler's equations (v):

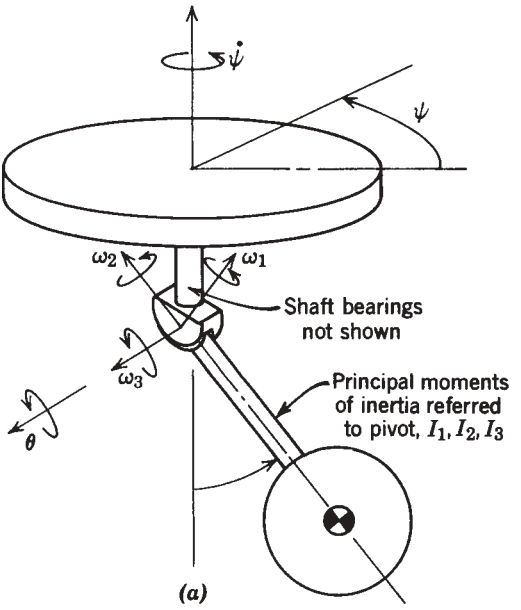


Select one of the *MGY*s, and write the equations that it represents. Verify power conservation in the form

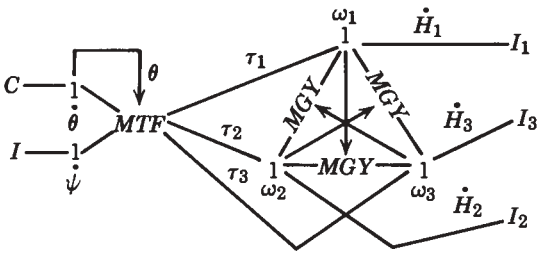
$$\dot{H}_1\omega_1 + \dot{H}_2\omega_2 + \dot{H}_3\omega_3 = \tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3.$$

9-15. The mechanical system shown in part *a* of the figure below contains one rigid body that rotates about a fixed axis and one that moves in more general motion about a single fixed point. Using the results of Problem 9-14, verify the bond graph representation shown in part *b*. By computing

$\omega_1, \omega_2, \omega_3$ in terms of θ, ψ, find the constitutive laws for the multiport *MTF*.



(a)

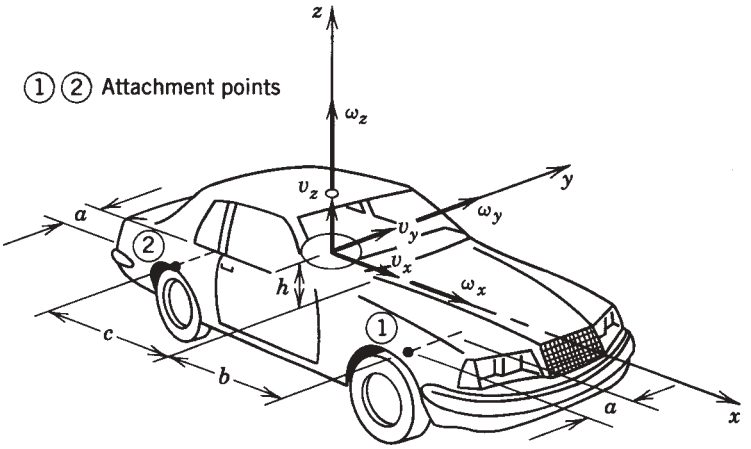


(b)

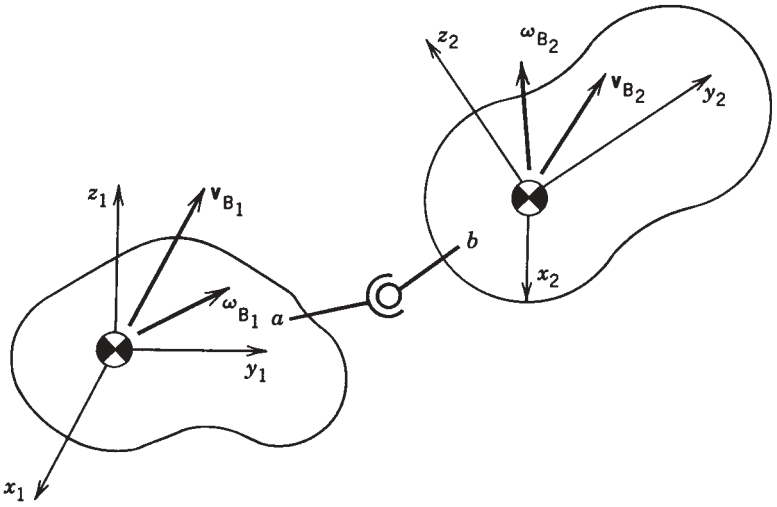
If θ and ψ were used as generalized coordinates, a nonlinear, fourth-order set of state equations could readily be found by applying Lagrange's equation to T—V. If all dynamic elements could accept integral causality, a fifth-order state space would result. This suggests that derivative causality may be necessary for some of this system. See if this is true by expanding the multiport *MTF* into a junction structure of 0- and 1-junctions and 2-port *MTF*s and applying causality.

9-16. The coordinate transformation of Figure 9.6 takes the body-fixed center-of-mass velocity and angular velocity components and outputs the inertial

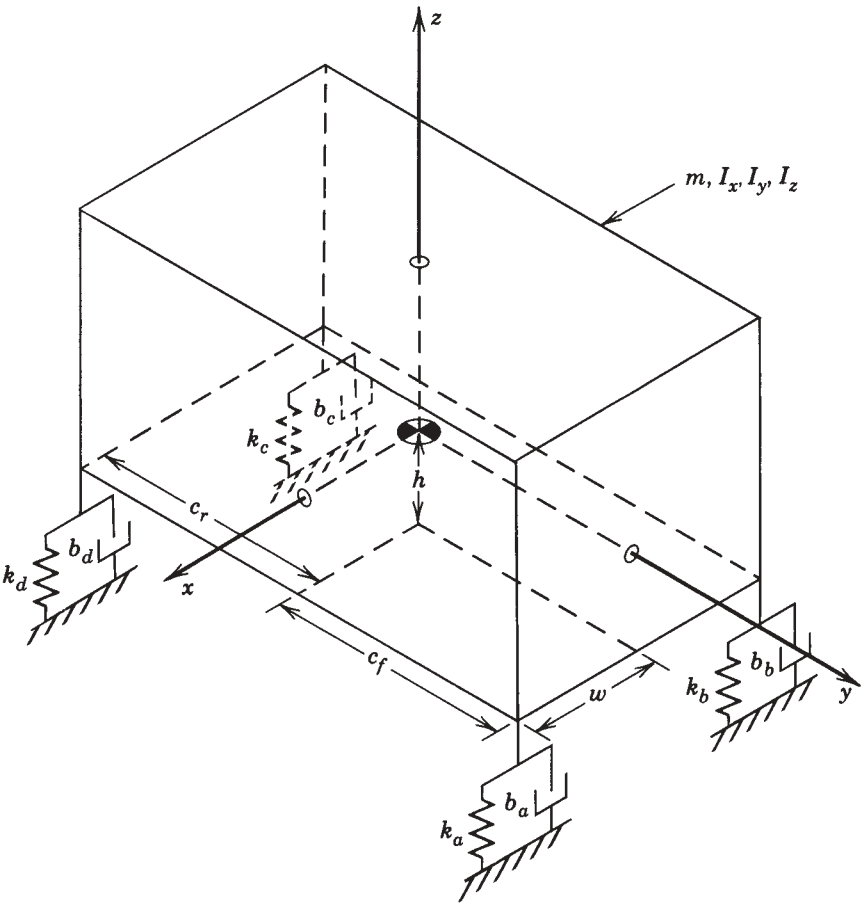
angular velocity of the body and the inertial velocity components of the center of mass. We frequently need the velocity components at attachment points on the rigid body. Using the figure shown, derive the coordinate transformation between center-of-mass body-fixed components and the body-fixed components at the attachment points shown.



9-17. The figure shows two rigid bodies connected by a frictionless spherical joint. Two body-fixed coordinate frames are shown. The spherical joint is at attachment point a on body 1 and attachment point b on body 2. Using a shorthand notation similar to that of Figure 9.6, construct a bond graph model for this interaction, assign causality, and discuss the problems associated with deriving a computational model.



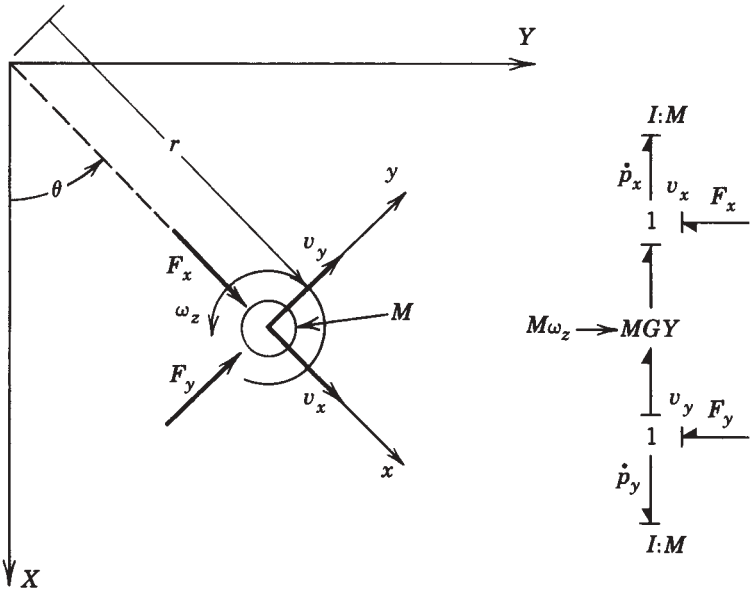
9-18. The figure shows a rigid body with a spring–damper suspension at each corner. A body-fixed coordinate frame is attached. A constraint system, not shown, ensures that ω_z (the angular velocity about the moving z-axis) remains zero. Using Figure 9.2, show the simplified bond graph fragments that will represent the rigid-body part of this system.



9-19. For small displacements it is reasonable to assume that the suspension units in Problem 9-18 generate only vertical forces and that the body-fixed z-axis remains virtually vertical. This assumption eliminates the need for the coordinate transformation from body-fixed components at the attachment points to inertial components. Construct a bond graph model of the system

under these assumptions. It is still convenient to use matrix transformations between center-of-mass motion and individual attachment points.

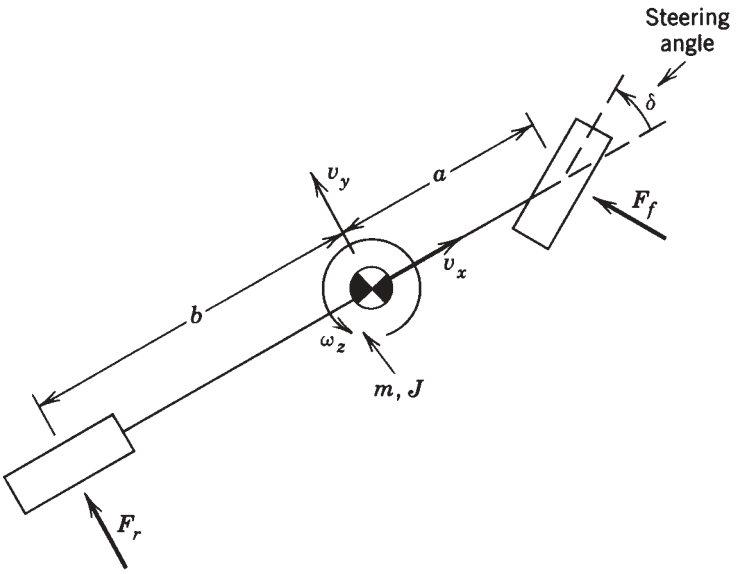
9-20. Planar motion of a mass particle is shown in the figure with respect to both a body-fixed frame and an inertial frame. The polar coordinates r, θ are used for the inertial description. A reduced bond graph fragment for planar motion is also shown in the figure. From classical dynamics, the absolute acceleration in r and θ directions are



$$a_r = \ddot{r} - r\dot{\theta}^2, \qquad a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta}.$$

Using the bond graph for planar motion, show that the absolute acceleration is the same as for polar coordinates.

9-21. A classical simplified vehicle model, known as the *bicycle* model, is shown in a top view in the figure. The vehicle travels in the xy-plane and has no width. There is also no suspension, and only cornering forces are considered. The front wheel can be steered through the angle δ. The only angular velocity considered is that about the body-fixed z (vertical) axis, ω_z. There is no wheel mass, nor any bushings to consider. Using Figures 9.2, 9.25, and 9.26, construct a complete bond graph model of this vehicle and assign causality.



9-22. One of the uses of the bicycle model of a ground vehicle is for stability considerations. For this purpose it is assumed that v_x is large compared to v_y, and v_x is virtually constant. Also, the steering angle δ is small, and the slip angles front and rear are small enough that the cornering forces are related to their respective slip angles by the linear relationships

$$F_f = C_f \alpha_f, \qquad F_r = C_r \alpha_r,$$

where

$$\alpha_f = \frac{\text{transverse velocity at front}}{v_x},$$
$$\alpha_r = \frac{\text{transverse velocity at rear}}{v_x}.$$

Modify the bond graph from Problem 9-21 to reflect these additional simplifications. Assign causality and identify state variables.

9-23. Derive the state equations for the system of Problem 9-22. Put them into standard matrix form for linear equations, remembering that v_x is constant for this model.

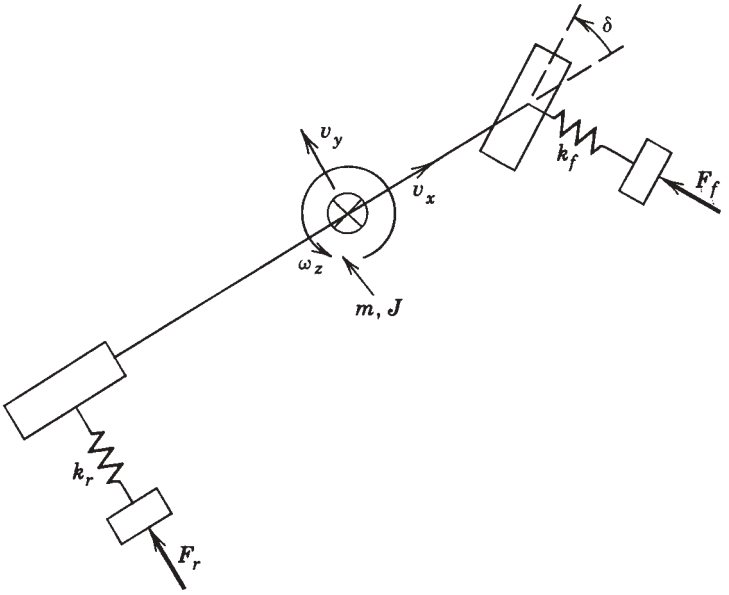
9-24. The equations from Problem 9-23 are

$$\frac{d}{dt}\begin{bmatrix} p_y \\ p_J \end{bmatrix} = \begin{bmatrix} \dfrac{-(C_r + C_f)}{mv_x} & \dfrac{bC_r - aC_f}{Jv_x} - \dfrac{mv_x}{J} \\ \dfrac{bC_r - aC_f}{mv_x} & \dfrac{-(a^2C_f + b^2C_r)}{Jv_x} \end{bmatrix} \begin{bmatrix} p_y \\ p_J \end{bmatrix} + \begin{bmatrix} C_f \\ aC_f \end{bmatrix} \delta.$$

Perform an eigenvalue analysis and determine the conditions for stability of the vehicle.

9-25. It is well known that pneumatic tires have sidewall compliance that can affect stability. The figure shows the bicycle model with springs, k_f and k_r, between the wheel and the contact patch. It is assumed that the springs are always in line with their respective axles as the vehicle moves.

Construct a bond graph of the new model.



10

DISTRIBUTED-PARAMETER SYSTEMS

Distributed-parameter systems are, analytically, those represented by partial rather than total differential equations. Physically, distributed systems are engineering components and devices that cannot be accurately approximated by the *lumped-parameter* models used thus far in the text. Remember that *all* engineering models are approximations and are made up by the modeler to answer specific questions concerning system understanding, design, and control. As modelers we are always trying to use the simplest model that will lead to understanding of the actual system. As Einstein once said, "Everything should be made as simple as possible, but not simpler."

So far in the text we have assumed that engineering systems are built from components exhibiting inertia-like, compliance-like, or resistance-like effects. A mass has been assumed rigid with no finite stiffness, and springs have been compliant with no mass, for example. Of course, all materials exhibit inertial and compliance effects, and it is our modeling assumptions that justify separating these effects into specific *lumps*.

There are real engineering systems in which this lumping process is far from obvious. An automobile for which we are interested in suspension design can be represented as a rigid body and characterized by its mass and moments of inertia. A tractor–trailer rig for which we are interested in suspension design cannot be represented as rigid bodies if we are to accurately model the actual system dynamics. This is because the frames of the tractor–trailer are flexible, and their bending motion is essential to understanding the system. Automobile frames bend also, but their bending frequencies are significantly higher than those of the truck and are typically out of the range of interest. Thus, we may be able to create a lumped-parameter model of the car, but not of the truck.

Multistory buildings, elevated roadways, flexible structures for outer space, and long hydraulic fluid lines are all examples of engineering systems that cannot be modeled without some consideration of distributed dynamic effects. This

Copyright © 2012 John Wiley & Sons, Inc.

chapter introduces the idea of distributed system modeling within the framework of bond graphs. Simple lumping techniques are discussed first, along with their limitations. Finite-mode modeling is then shown to produce more accurate low-order representations for distributed systems. How to combine lumped and distributed subsystems becomes obvious using this approach, and the use of causality still aids in formulating the system equations for subsequent computation.

10.1 SIMPLE LUMPING TECHNIQUES FOR DISTRIBUTED SYSTEMS

In the course of deriving a partial derivative equation model of a dynamic component or subsystem, it is typical to start with spatially distributed finite lumps and then to take the limit as the lumps become infinitesimal in size. One approach to creating lumped-parameter models for distributed systems is motivated by this means of deriving equations. We simply stop with the finite lumps and perform no limiting process. This approach is physically understandable; however, even though the continuous equations are approached as the number of finite lumps becomes infinite, it is difficult to predict the rate of convergence. Thus, a large number of lumps may be required to obtain accuracy response predictions even for low-frequency excitations. In addition, each new lump, while improving low-frequency prediction, introduces new, very inaccurate results at high frequencies. This approach to distributed system representation must be used cautiously, and with awareness on the part of the modeler. Reference [1] surveys the different approaches to distributed system modeling, of which this simple finite-lumping technique is one.

10.1.1 Longitudinal Motions of a Bar

Figure 10.1*a* shows a continuous representation of a uniform bar of length L, cross-sectional area A, and mass density ρ. It is fixed at the left end and acted upon by an external force $F(t)$ at the right. The continuous variable x locates any section of the bar, and the variable $\xi(x, t)$ represents the displacement of any cross section from an initially undisplaced state. Figure 10.1*b* shows a finite element of mass, $\rho A \Delta x$, with an imbalance in the normal stress σ across it. Newton's law yields

$$A\sigma(x + \Delta x) - A\sigma(x) = \rho A \, \Delta x \frac{\partial^2 \xi}{\partial t^2}. \tag{10.1}$$

Figure10.1*c* shows the same section subject to simple strain, yielding the relationship

$$\sigma(x) = \frac{E[\xi(x) - \xi(x - \Delta x)]}{\Delta x}, \tag{10.2}$$

where E is Young's modulus. Notice that Eqs. (10.1) and (10.2) are written as though the lumps occupied different points in space, when, in fact, as Δx becomes infinitesimal, the mass lump and stiffness lump relate to the same spatial location.

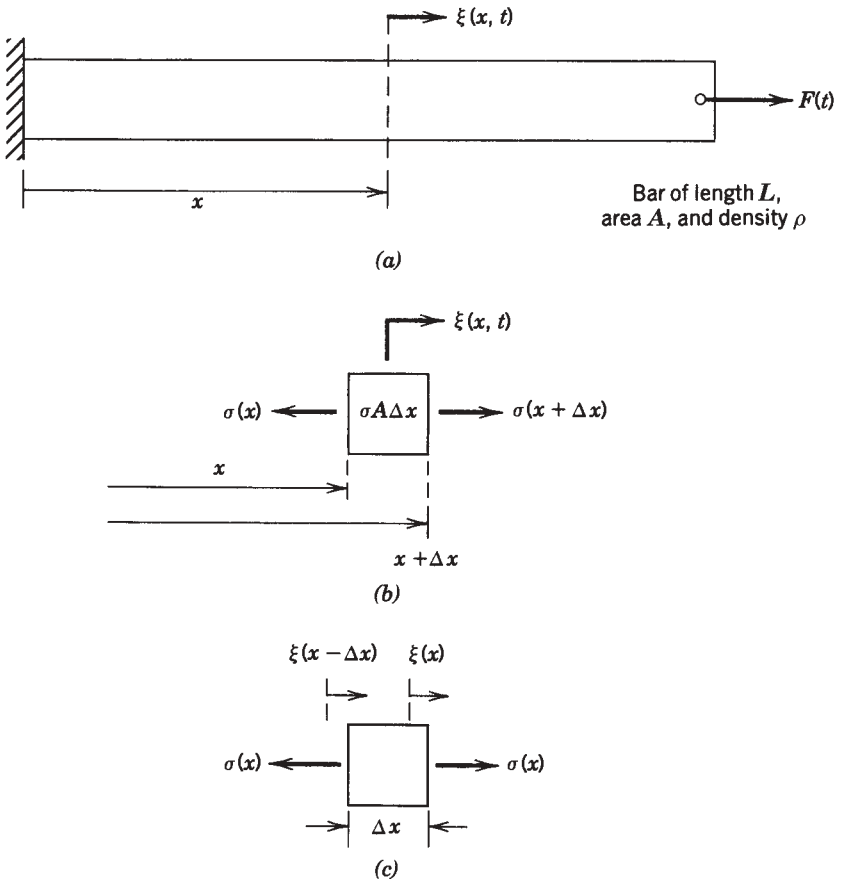


FIGURE 10.1. Vibrating bar.

We will soon let $\Delta x \to 0$, but first we interpret (10.1) and (10.2) as finite lumps, and construct a lumped approximation to the system. Let $A\sigma = F$, and write (10.1) as

$$F_{i+1} - F_i = \frac{d}{dt} p_i \tag{10.3}$$

and (10.2) as

$$F_i = \frac{EA}{\Delta x} q_i, \tag{10.4}$$

where

$$p_i = \rho A \, \Delta x \dot{\xi}_i \tag{10.5}$$

is the momentum of the ith lump and

$$q_i = \xi_i - \xi_{i-1} \tag{10.6}$$

is the relative displacement between the ith and the $(i-1)$st lump. This yields the equations

$$\frac{d}{dt} p_i = \frac{EA}{\Delta x}(q_{i+1} - q_i), \tag{10.7}$$

$$\frac{d}{dt} q_i = \frac{p_i - p_{i-1}}{\rho A\, \Delta x}. \tag{10.8}$$

These equations are those generated by the internal elements in the bond graph model of Figure 10.2. The first element and the nth element must be treated separately, since they are boundary elements.

From the bond graph of Figure 10.2, the first element yields

$$\frac{dq_1}{dt} = \frac{p_1}{\rho A\, \Delta x}, \tag{10.9}$$

and the nth element yields

$$\frac{dp_n}{dt} = F(t) - \frac{EA}{\Delta x} q_n. \tag{10.10}$$

To use this approach for an actual system, we would select the number of lumps we wish to include and let Δx equal the bar length divided by the number of lumps. This choice will then set the I and C parameters. We would like to choose the number of included lumps based on some accuracy criteria; however, as mentioned previously, it is not possible to determine, a priori, the convergence to the continuous model for general uses of this lumped model as part of an overall system. In fact, this approach to representing inherently distributed systems is

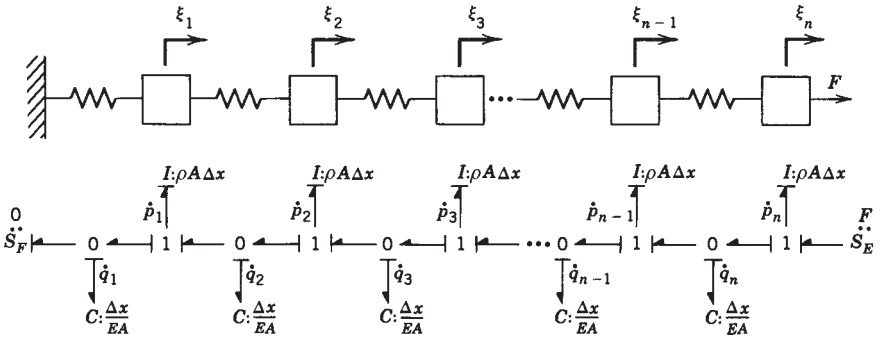


FIGURE 10.2. Bond graph finite-lump model of vibrating bar.

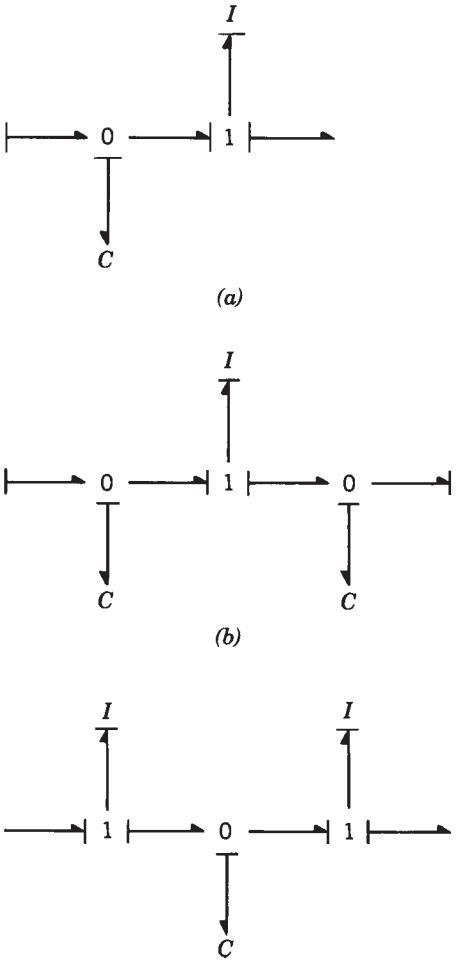


FIGURE 10.3. Typical configurations for finite-lump model of a bar.

usually used in configurations as shown in Figure 10.3, the choice being dictated by causal considerations of the entire system.

Returning to Eqs. (10.1) and (10.2), the continuous model can be derived by letting $\Delta x \to 0$ while letting small changes in variables become differentials. Thus, (10.1) becomes

$$A\frac{\partial \sigma}{\partial x} = \rho A \frac{\partial^2 \xi}{\partial t^2}, \tag{10.11}$$

for the internal domain of the bar, and (10.2) becomes

$$\sigma = E\frac{\partial \xi}{\partial x}. \tag{10.12}$$

Combining (10.11) and (10.12) yields the continuous representation for longitudinal vibration of a bar with no external forcing:

$$E\frac{\partial^2 \xi}{\partial x^2} = \rho\frac{\partial^2 \xi}{dt^2}, \tag{10.13}$$

which should be recognized as the simple wave equation.

When dealing with continuum elements, external forcing, such as $F(t)$ in Figure 10.1, is considered to be distributed over the spatial domain of the element so that the force acting on any incremental element, $F(x, t)$, would be

$$F(x, t) = f(x, t)\Delta x, \tag{10.14}$$

where $f(x, t)$ is the force per unit length and could be distributed over the spatial dimension x of the rod in this example.

For a point force, such as $F(t)$ acting at $x = L$ in Figure 10.1, the distributed force per unit length is mathematically expressed through use of the delta function as

$$f(x, t) = F(t)\delta(x - L). \tag{10.15}$$

(The delta function $\delta(x - L)$ is a function that is zero everywhere except in the immediate vicinity of $x = L$, but when integrated in $x$over any range that includes $x = L$ yields unity.)

The total external force acting on the rod is obtained through integration of $f(x)$ from Eq. (10.15) with respect to x over the total length, and turns out to be $F(t)$ as expected, since the area under the delta function is unity. The delta function correctly locates the force, since the delta function is zero except where its argument is zero. We shall make extensive use of the point force representation in some of the following development. For our rod example, Eq. (10.13) becomes

$$\frac{F(t)}{A}\delta(x - L) + E\frac{\partial^2 \xi}{\partial x^2} = \rho\frac{\partial^2 \xi}{\partial t^2}, \tag{10.16}$$

when the external forcing at the rod end is included.

We now have two types of representations for the bar. The first is a finite-lump representation, which will generate a high-order state space when many lumps are used. This lumped representation is viscerally satisfying because it is exactly the type of modeling we have been doing all along. We can imagine springs being compressed and stretched, causing the inertial elements in between to move. It is easy to "see" the physics of the system. Unfortunately, this finite-lump approach may not be particularly good as the number of included lumps increases. The state space grows very quickly while the accuracy does not.

The second representation for the bar is the linear partial differential equation (10.16). This recognizes that the bar model is actually of infinite order, and as long as the simple wave equation is assumed to be valid, quite a lot can be said analytically about the continuous behavior of the bar. However, if dissipation or

nonlinear effects are included (which is more often the case than not), then we are generally forced to solve the continuum equations numerically. This always means backing up into a finite representation of one form or another.

We shall see that the continuous representation can be manipulated using normal mode techniques to yield a much more accurate and computationally efficient finite-order model than does the lumping process described above, even especially as the continuum is approached by including more and more lumps.

Before proceeding to the use of the continuous model, transverse motion of a beam is discussed.

10.1.2 Transverse Beam Motion

We assume here that the beam is long and slender and that it is of uniform cross-sectional area A, mass density ρ, Young's modulus E, shear modulus G, area moment of inertia I, and length L. Figure 10.4 shows the beam displaced

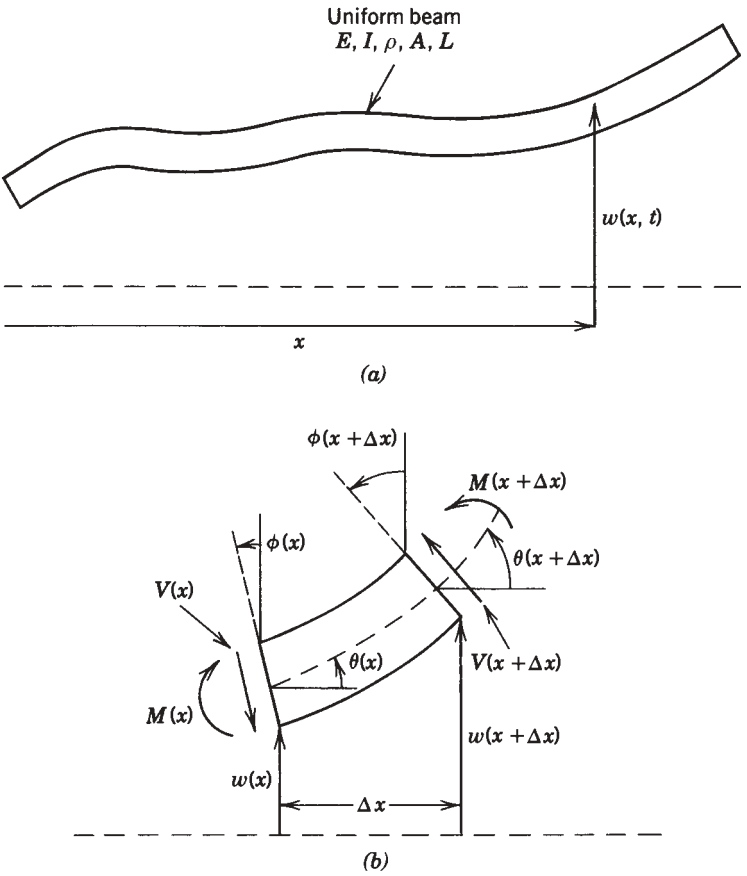


FIGURE 10.4. Uniform beam in transverse motion.

at some instant of time. The spatial variable x defines a position along the beam, and $w(x, t)$ is the transverse displacement of the position x at the time t.

Also shown in Figure 10.4 is a finite element of the beam with the various forces and moments acting on it. The angle θ is the rotation of the neutral axis with respect to a horizontal reference, and the angle ϕ is the rotation of a plane cross section with respect to a vertical reference. If $\phi = \theta$, then we say there is no shear deformation. In general $\phi \neq \theta$, and we define the shear angle γ as

$$\gamma = \theta - \phi. \tag{10.17}$$

We shall derive the equations of motion for a *Timoshenko* model of the beam. Reference [2] discusses this model thoroughly.

The Timoshenko model includes shear deformation and the rotary inertia of each cross section. We shall later neglect these effects to reduce our model to the perhaps more familiar *Bernoulli–Euler* model of the beam.

From Figures 10.4*b* and 10.5 we can use Newton's laws for the element as

$$V(x + \Delta x) - V(x) = \rho A \Delta x \frac{\partial^2 w}{\partial t^2}, \tag{10.18}$$

and

$$M(x + \Delta x) - M(x) + V(x + \Delta x)\Delta x = \rho I \Delta x \frac{\partial^2 \phi}{\partial t^2}, \tag{10.19}$$

where it should be noted from Figure 10.5 that the inertia properties and compliant properties of the beam are assumed to be spatially separate.

From Figure 10.5*b*, we assume that the bending moment can be calculated as though pure bending took place. Thus,

$$M(x) = \frac{EI}{\Delta x}[\phi(x + \Delta x) - \phi(x)]. \tag{10.20}$$

From Figure 10.5*c*, we assume that the shear $V(x)$ can be related to the shear angle by

$$V(x) = kGA(\theta - \phi), \tag{10.21}$$

where

$$\theta(x) = \frac{w(x + \Delta x) - w(x)}{\Delta x}, \tag{10.22}$$

the slope of the neutral axis. In Eq. (10.21), k is a parameter that accounts for the actual nonuniform shear distribution across a cross section.

Equations (10.18)–(10.22) suggest the lumped representation shown in Figure 10.6 and as a bond graph in Figure 10.7. Each element has mass $\rho A \Delta x$ and rotary inertia $\rho I \Delta x$. The spring elements k_b provide the bending moment

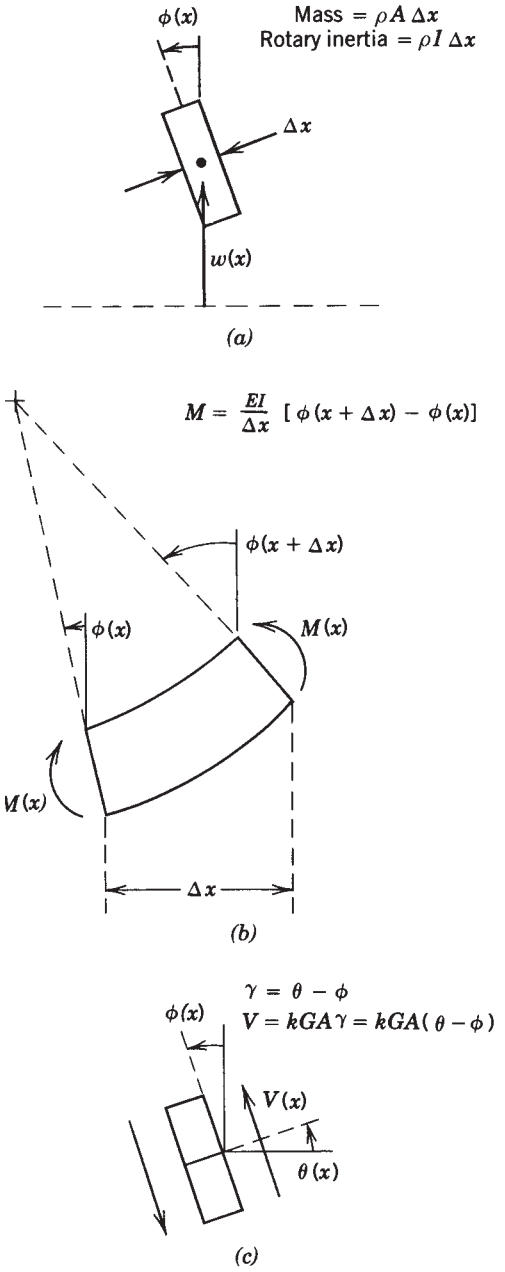


FIGURE 10.5. Beam elements for use in Timoshenko model.

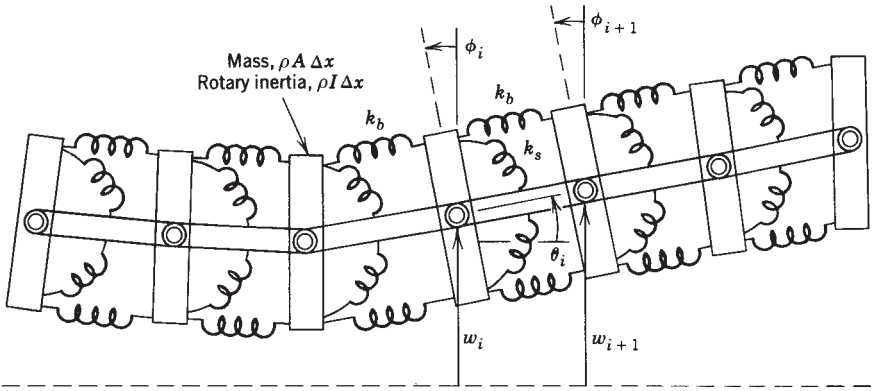


FIGURE 10.6. Lumped representation of Timoshenko beam.

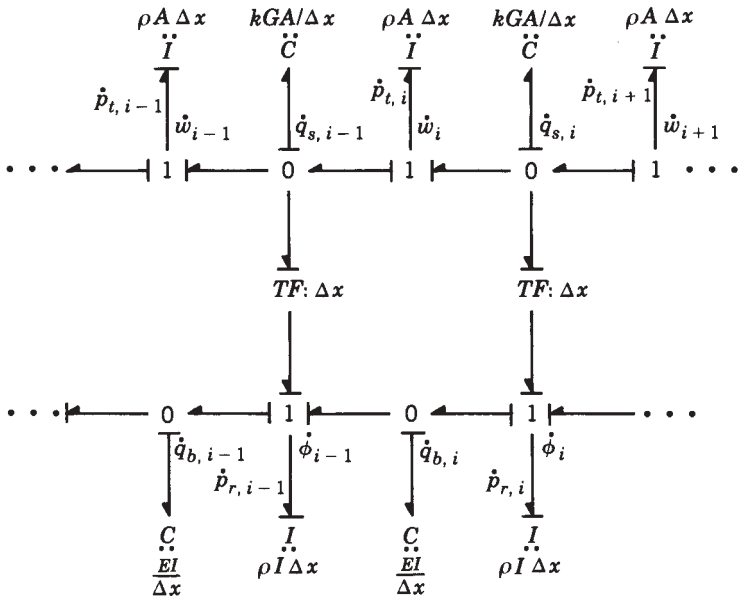


FIGURE 10.7. Bond graph finite-lump model of Timoshenko beam.

and react to the difference in ϕ between successive elements, as indicated in Eq. (10.20). The spring elements k_s are the shear springs, and they react to the difference between θ_i and ϕ_i, at any location i, as indicated in Eq. (10.21), where

$$\theta_i \cong \frac{w_{i+1} - w_i}{\Delta x}, \tag{10.23}$$

as indicated by Eq. (10.22).

The bond graph of Figure 10.7 shows the lumped model of the Timoshenko beam. The parameters would be chosen in a similar manner as for the bar discussed previously. We would hope to base the number of included lumps on some use criteria for the model, then to simply select Δx to be the beam length divided by the number of included lumps.

As used in Figure 10.7, the bond graph possesses all integral causality. Thus, state equations can be straightforwardly derived using the procedures described in an earlier chapter. Care must be exercised when using this representation as part of an overall model, to terminate the lumping process in a manner that preserves integral causality. This will greatly facilitate computation.

We are now in a position to derive the continuous representation of the beam. Before doing this, it is interesting to reduce our lumped model to a Bernoulli–Euler beam model. This model neglects rotary inertia and shear deformation. In Figure 10.7, we simply remove the lower inertial elements and the upper shear compliances. This results in the lumped model of Figure 10.8. Notice that integral causality still exists, although the shear forces associated with the upper 0-junctions are no longer set by the shear stiffness elements, but now are algebraically set by the bending stiffness.

The Timoshenko model can be reduced to a continuum representation by letting $\Delta x \to 0$ and letting small differences divided by Δx become derivatives. Thus, from Eq. (10.18),

$$\lim_{\Delta x \to 0} \frac{V(x + \Delta x) - V(x)}{\Delta x} = \frac{\partial V}{\partial x} \tag{10.24}$$

and

$$\frac{\partial V}{\partial x} = \rho A \frac{\partial^2 w}{\partial t^2}. \tag{10.25}$$

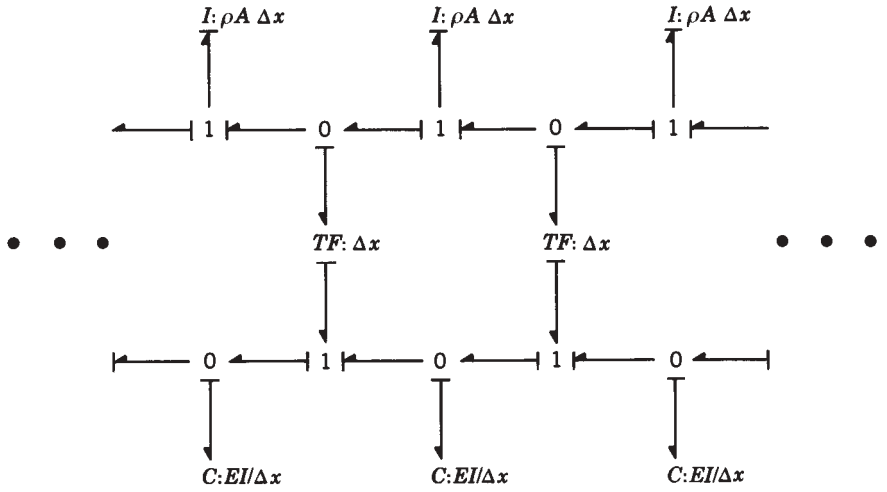


FIGURE 10.8. Bond graph finite-lump model for Bernoulli–Euler beam.

Similarly, from Eq. (10.19),

$$\frac{\partial M}{\partial x} + V = \rho I \frac{\partial^2 \phi}{\partial t^2}, \tag{10.26}$$

and from (10.20),

$$M = EI \frac{\partial \phi}{\partial x}. \tag{10.27}$$

From Eqs. (10.21) and (10.22)

$$\theta = \frac{\partial w}{\partial x}, \tag{10.28}$$

and

$$V = kGA \left(\frac{\partial w}{\partial x} - \phi \right). \tag{10.29}$$

We can now combine Eqs. (10.25) and (10.29) to obtain

$$kGA \frac{\partial}{\partial x} \left(\frac{\partial w}{\partial x} - \phi \right) = \rho A \frac{\partial^2 w}{\partial t^2}, \tag{10.30}$$

and we can combine Eqs. (10.26), (10.27), and (10.29) to obtain

$$EI \frac{\partial^2 \phi}{\partial x^2} + kGA \left(\frac{\partial w}{\partial x} - \phi \right) = \rho I \frac{\partial^2 \phi}{\partial t^2}. \tag{10.31}$$

Equations (10.30) and (10.31) are the two partial differential equations in the two unknowns, $w(x, t)$ and $\phi(x, t)$, for the continuous uniform Timoshenko beam. These can be reduced to the Bernoulli–Euler model by neglecting rotary inertia in (10.31) and letting the shear stiffness become infinite, so that

$$\theta = \phi = \frac{\partial w}{\partial x}. \tag{10.32}$$

Then, substituting into Eq. (10.30) yields

$$-\frac{\partial}{\partial x} EI \frac{\partial^2 \phi}{\partial x^2} = -EI \frac{\partial^4 w}{\partial x^4} = \rho A \frac{\partial^2 w}{\partial t^2}, \tag{10.33}$$

or

$$EI \frac{\partial^4 w}{\partial x^4} + \rho A \frac{\partial^2 w}{\partial t^2} = 0. \tag{10.34}$$

If point forces $F_1(t)$ and $F_2(t)$ were acting on the Bernoulli–Euler beam at locations x_1 and x_2, respectively, along the beam, Eq. (10.34) would become

$$F_1 \delta(x - x_1) + F_2 \delta(x - x_2) - EI \frac{\partial^4 w}{\partial x^4} = \rho A \frac{\partial^2 w}{\partial t^2}. \tag{10.35}$$

While the Timoshenko model is a more realistic representation in that the high-frequency behavior is more accurately predicted, the solution is too cumbersome to pursue here. The interested reader should see Reference [2]. We will henceforth deal exclusively with the Bernoulli–Euler model.

Before continuing into the main emphasis of this chapter, it is perhaps useful to summarize the salient points that have been demonstrated here by example.

1. A lumped-parameter model can usually be found for a distributed system by following the derivation for the continuous equations up to the point where a typical element of the continuum is allowed to become infinitesimal and then simply retaining a finite size for the element. Such a lumped model may not be very accurate without retaining a large number of small elements.
2. The lumped-parameter representations have the advantage that they can be straightforwardly combined into an overall system model. Also, nonlinear effects pose no particular problems. The disadvantage to the lumped-parameter approach described so far is that the model will generate a large state space and can cause severe computational problems due to large disparities in the time scales of the distributed portion of the model and the remainder of the system.
3. True continuum models can yield insight into system behavior when analytical solutions are available, which typically restricts us to the linear case. However, it is usually difficult to incorporate a continuum model into an overall system model with interactions with complex lumped systems (or other continuous systems) external to the continuum. When this is attempted, it becomes a problem of solving the partial differential equations for the continuum part, subject to complicated boundary conditions for the attached parts. This is really not a practical approach for obtaining an insightful representation of an overall system. It is much better to find an accurate, low-order, lumped-parameter representation for the continuous part of the system which can interact easily with external dynamics. We do this next.

10.2 LUMPED MODELS OF CONTINUA THROUGH SEPARATION OF VARIABLES

Many physical elements from which engineering systems are constructed possess a continuous distribution of inertia and compliance properties. The rod and beam presented previously are examples of such elements, as are stretched cables, membranes, plates, and shells. An aircraft fuselage is an example of a complex system composed of distributed elements. When it is justified, and when we as modelers choose to do so, these distributed elements can be represented by linear partial differential equations in space and time.

The actual physical dynamics of these distributed elements and systems are composed of propagating waves that reflect from boundaries and add together to produce what is actually observed at any point and time. There is mathematical development that we could pursue to demonstrate that continuous systems with distributed inertia and compliance properties possess wavelike behavior. However, this will not be carried out here, and the interested reader should see Reference [3].

Instead, we will pursue separation of variables, an alternative approach to representing the motion time history of linear, distributed wavelike systems. As will be seen, this approach yields an elegant lumped-parameter bond graph representation of distributed elements which can easily be coupled with bond graph models of attached dynamic systems to yield an extremely accurate overall system model possessing all the formulational and computational virtues associated with bond graphs. It is possible to pursue separation of variables in a very general way and demonstrate some very useful characteristics of the approach that apply to any appropriate continuum. Again, Reference [3] is recommended for that development. Here, we will demonstrate the approach through example.

10.2.1 The Bar Revisited

The equation for forced longitudinal vibration of a bar is Eq. (10.16), repeated here as

$$\rho \frac{\partial^2 \xi}{\partial t^2} - E \frac{\partial^2 \xi}{\partial x^2} = \frac{F(t)}{A} \delta(x - L). \tag{10.36}$$

Separation of variables begins by temporarily setting the force $F(t)$ to zero and assuming that the displacement at any point and time, $\xi(x, t)$, can be separated into a product of a function of x only, $Y(x)$, and a function of time only, $f(t)$:

$$\xi(x, t) = Y(x) f(t). \tag{10.37}$$

Note that this general function of time $f(t)$ is *not* the force density of Eqs. (10.14) and (10.15).

Substitution into the homogeneous form of (10.36) yields

$$\rho Y \frac{d^2 f}{dt^2} - E f \frac{d^2 Y}{dx^2} = 0. \tag{10.38}$$

Dividing by $\rho Y f$ yields

$$\frac{1}{f} \frac{d^2 f}{dt^2} = \frac{E}{\rho} \frac{1}{Y} \frac{d^2 Y}{dx^2}. \tag{10.39}$$

The collection of terms on the left of (10.39) is, by assumption, dependent on time only, while the collection of terms on the right is dependent on x only. For

this to be true for all values of the independent variables x and t, each collection of terms must be equal to the same constant. By convention we let

$$\frac{1}{f}\frac{d^2 f}{dt^2} = -\omega^2, \tag{10.40}$$

with the result

$$\frac{d^2 f}{dt^2} + \omega^2 f = 0, \tag{10.41}$$

$$\frac{d^2 Y}{dx^2} + \frac{\rho}{E}\omega^2 Y = 0. \tag{10.42}$$

For a separated solution to exist, the two ordinary differential equations (10.41) and (10.42) must be satisfied.

First consider Eq. (10.42), for the spatial function $Y(x)$. For this simple example, the spatial differential equation is easily solved:

$$Y(x) = A\cos kx + B\sin kx, \tag{10.43}$$

where

$$k^2 = \frac{\rho}{E}\omega^2. \tag{10.44}$$

To complete the solution for $Y(x)$, we must now specify the boundary conditions.

For the bar of Figure 10.1, the displacement at the left end, $\xi(0, t)$, is zero. At the right end, we have a choice of ways to express the boundary condition. We can include the force $F(t)$ in the boundary condition by balancing it with the normal stress at $x = L$. Thus,

$$\sigma(L, t)A = F(t), \tag{10.45}$$

or

$$EA\frac{\partial \xi}{\partial x}(L, t) = F(t). \tag{10.46}$$

We can also think of the force as being applied a very small distance upstream of the right end, so that the right end itself is stress-free. Thus,

$$\sigma(L, t) = E\frac{\partial \xi}{\partial x}(L, t) = 0. \tag{10.47}$$

This replacement of the boundary forcing with a limiting form of forcing in the bar interior is analytically convenient, and it causes no problems, since we are only interested in solutions to within a finite scale of fineness. The boundary conditions for our example become

$$\xi(0, t) = Y(0)f(t) = 0,$$

or

$$Y(0) = 0, \tag{10.48}$$

and

$$\frac{\partial \xi}{\partial x}(L, t) = \frac{dY}{dx}(L) \cdot f(t) = 0,$$

or

$$\frac{dY}{dx}(L) = 0. \tag{10.49}$$

Application of (10.48) and (10.49) to (10.43) yields

$$A = 0, \tag{10.50}$$

and

$$Bk \cos kL = 0. \tag{10.51}$$

Equation (10.51) is called the frequency equation, and it is analogous to the characteristic equation for the eigenvalues of a lumped system.

The solution continues by recognizing that if B or k is zero, then $Y(x) = 0$ and $\xi(x, t) = 0$, a possible but uninteresting solution. Instead, we let

$$\cos kL = 0, \tag{10.52}$$

with the result

$$k_n L = (2n - 1)\frac{\pi}{2}, n = 1, 2, 3, \ldots, \tag{10.53}$$

and, from Eq. (10.44),

$$\omega_n^2 = \frac{E}{\rho} k_n^2 = \frac{E}{\rho} \frac{(k_n L)^2}{L^2},$$

or

$$\omega_n = \sqrt{\frac{E}{\rho}} \frac{(2n - 1)}{L} \frac{\pi}{2}, n = 1, 2, 3, \ldots \tag{10.54}$$

Accompanying each ω_n is a special shape function, $Y_n(x)$, from Eq. (10.43):

$$Y_n(x) = B_n \sin\left(k_n L \frac{x}{L}\right) = B_n \sin\left((2n - 1)\frac{\pi}{2}\frac{x}{L}\right), n = 1, 2, 3, \ldots \tag{10.55}$$

The constants B_n are arbitrary, and it is convenient here to set them equal to unity.

The shape functions $Y_n(x)$ are frequently called *eigenfunctions*, *mode shapes*, or *normal modes*. Corresponding to each mode shape is a *natural frequency* ω_n. Figure 10.9 shows a few mode shapes for a bar fixed at the left end and

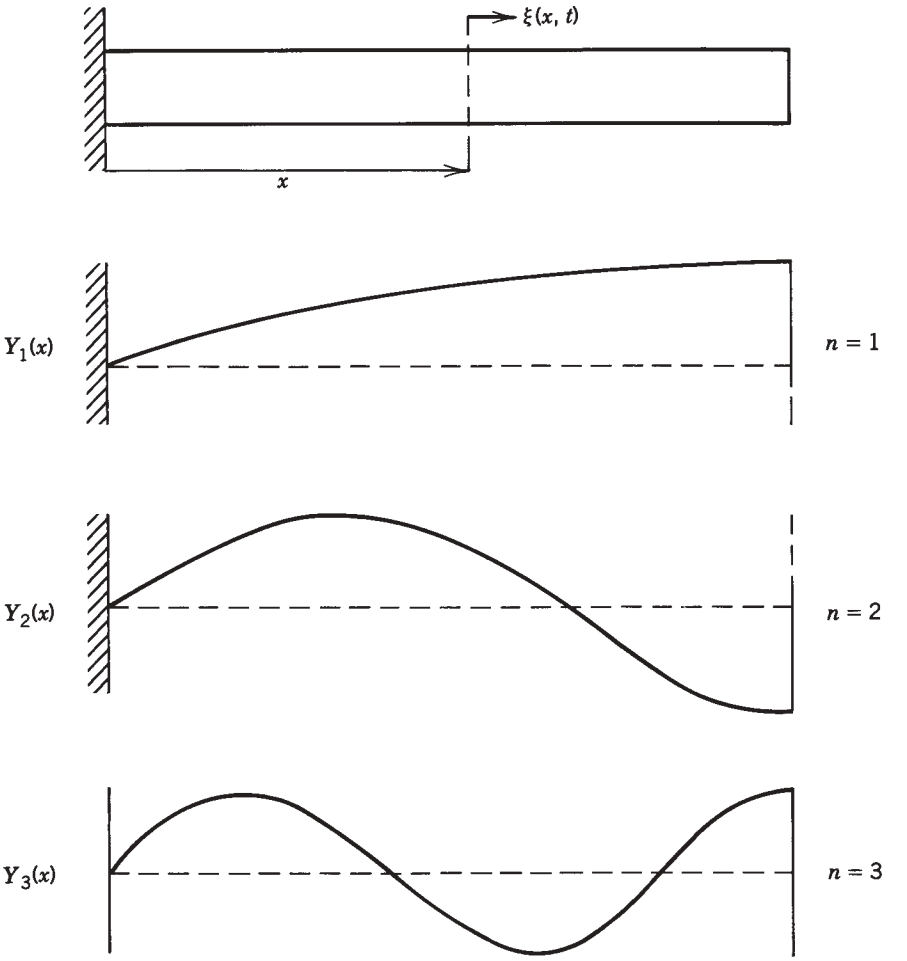


FIGURE 10.9. Mode shapes for the bar.

free at the right. One interpretation of a mode shape and its frequency is that, if the rod is initially displaced in the shape of a mode and released, it will oscillate harmonically in that modal configuration at the corresponding modal frequency. Another fact about these modes and frequencies is that any motion time history that the bar executes, from any initial conditions, is a linear combination of the mode shapes oscillating at their respective natural frequencies. And another most useful fact about these mode shapes is that they are orthogonal, in that

$$\int_0^L Y_n(x) Y_m(x)\, dx = 0, n \neq m, \tag{10.56}$$

or

$$\int_0^L \sin\left((2n-1)\frac{\pi}{2}\frac{x}{L}\right)\sin\left((2m-1)\frac{\pi}{2}\frac{x}{L}\right)dx = \begin{cases} 0, & n \neq m, \\ L/2, & n = m, \end{cases} \tag{10.57}$$

for the bar example. We shall make use of this fact soon.

We are interested in the forced response of the bar, but so far we have dealt only with the unforced response. We now make use of the fact [3] that the response of the bar, forced or unforced, can always be expressed as a linear combination of the mode shapes. We write

$$\xi(x,t) = \sum_{n=1}^{\infty} Y_n(x)\eta_n(t) \tag{10.58}$$

and use this assumption in the forced equation (10.36). Thus,

$$\sum_n \rho A Y_n \ddot{\eta}_n - \sum_n AE\frac{d^2Y_n}{dx^2}\eta_n = F(t)\delta(x-L). \tag{10.59}$$

Now, multiply each term by the mth mode shape $Y_m(x)$, and integrate term by term over the bar length:

$$\sum_n \left(\int_0^L \rho A Y_n Y_m dx\right)\ddot{\eta}_n + \sum_n \left(\int_0^L \rho A Y_n Y_m dx\right)\omega_n^2\eta_n = \int_0^L F(t)\delta(x-L)Y_m dx, \tag{10.60}$$

where Eq. (10.42) has been used in the second term by substituting

$$\frac{d^2Y_n}{dx^2} = -\frac{\rho}{E}\omega_n^2 Y_n. \tag{10.61}$$

As the summation is carried out for $n = 1, 2, \ldots$, each integral on the left-hand side of (10.60) is zero except for $n = m$, leaving only one equation,

$$\left[\int_0^L \rho A Y_m^2\, dx\right]\ddot{\eta}_m + \left[\int_0^L \rho A Y_m^2\, dx\right]\omega_m^2\eta_m = \int_0^L F(t)\delta(x-L)Y_m\, dx. \tag{10.62}$$

The integral on the left side is called the modal mass m_m:

$$m_m = \int_0^L \rho A Y_m^2\, dx, m = 1, 2, \ldots. \tag{10.63}$$

For the bar, it is simply

$$m_m = \frac{\rho AL}{2}, \tag{10.64}$$

independent of the mode number.

The integral on the right side of (10.62) is called the modal forcing function and is particularly simple to evaluate for point forcing. The delta function assures that no contribution to the integral occurs until $x = L$, at which point $Y_m(x)$ is virtually constant as we sweep across the delta function and equals $Y_m(L)$. Thus the integral becomes

$$\int_0^L F(t)\delta(x - L)Y_m(x)\,dx = F(t)Y_m(L)\int_0^L \delta(x - L)\,dx = F(t)Y_m(L). \tag{10.65}$$

Thus, Eq. (10.62) becomes

$$m_m\ddot{\eta}_m + k_m\eta_m = F(t)Y_m(L), \tag{10.66}$$

where m_m is the modal mass

$$m_m = \int_0^L \rho AY_m^2(x)\,dx = \frac{\rho AL}{2}, \tag{10.67}$$

k_m is the modal stiffness

$$k_m = m_m\omega_m^2, \tag{10.68}$$

and $Y_m(x)$ and ω_m are known from the unforced analysis for the mode shapes.

The beauty of this approach to distributed systems is that, due to mode orthogonality, all the equations in Eq. (10.66) are decoupled from one another, and each $\eta_m(t)$ can be solved for separately and then combined with the mode shapes $Y_m(x)$ in Eq. (10.58) to produce the actual response,

$$\xi(x, t) = \sum_{m=1}^{\infty} Y_m(x)\eta_m(t). \tag{10.69}$$

Obviously, only a finite number of modes can be retained in the solution.

This development is becoming long, so we move quickly to a very interesting result. Define the modal momentum p_m as

$$p_m = m_m\dot{\eta}_m \tag{10.70}$$

and the modal displacement q_m as

$$q_m = \eta_m. \tag{10.71}$$

Then Eq. (10.66) can be written as

$$\frac{d}{dt}p_m = -k_m q_m + F(t)Y_m(L), \tag{10.72}$$

and

$$\frac{dq_m}{dt} = \frac{p_m}{m_m}. \tag{10.73}$$

We have reduced the mth modal equation to two first-order state equations in terms of bond graph energy variables, p_m and q_m. Figure 10.10 shows a bond graph that would duplicate these modal equations for $m = 1, 2, 3, \ldots$ Notice that integral causality exists for all m included modes and that the external force F properly excites each mode according to Eqs. (10.72) and (10.73). The transformers are simply the mode shapes from Eq. (10.55) evaluated at the location of the force. The bond graph also enforces that the flow on the external force bond, which must be $\partial\xi(L,t)/\partial t$, is the summation of the modal flows multiplied by appropriate transformer moduli:

$$\frac{\partial\xi}{\partial t}(L,t) = Y_1(L)\frac{p_1}{m_1} + Y_2(L)\frac{p_2}{m_2} + \cdots + Y_m(L)\frac{p_m}{m_m}, \tag{10.74}$$

which is exactly Eq. (10.69) evaluated for the velocity at $x = L$. In fact, from Eq. (10.69),

$$F(t)\frac{\partial\xi}{\partial t}(L,t) = \sum_{m=1}^{\infty} F(t)Y_m(L)\dot{\eta}_m(t), \tag{10.75}$$

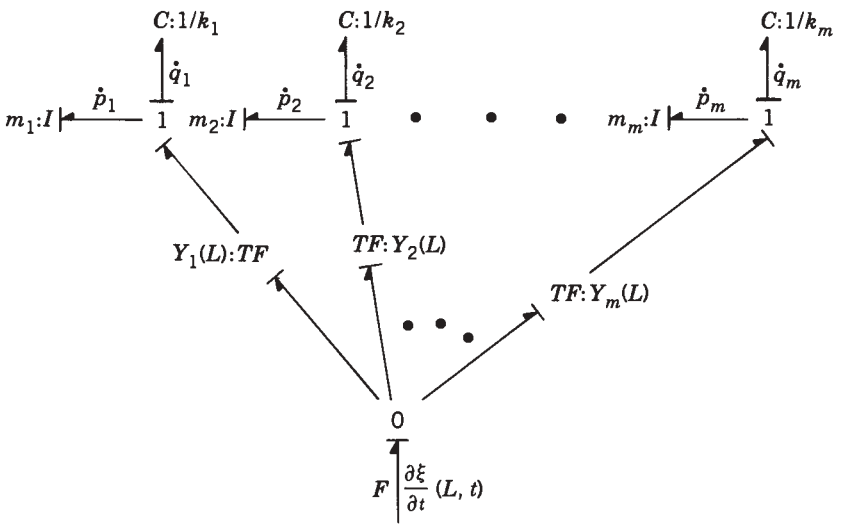


FIGURE 10.10. Bond graph for modal vibration of a bar forced at $x = L$.

but $F(t)Y_m(L)$ is the mth modal force, and $\dot{\eta}_m(t)$ is the mth modal flow. Thus the external force times the velocity at the input port equals the summation of the modal forces times corresponding modal flows. This is a statement of a power-conserving transformation, and the transformer fan of Figure 10.10 represents this transformation from physical variables to modal variables and back.

A good question at this point is: Why construct a bond graph when we have already tediously derived the equations (10.66) subject to (10.58)? The answer becomes obvious when we realize that the external force $F(t)$ is probably not a specified input, but rather resulted from interaction with some external system. For instance, if the end of the rod had an attached transducer which we decided to represent by a simple mass–spring–damper system, the bond graph of the interaction would be as shown in Figure 10.11. The modal dynamics are identical to Figure 10.10, and the external dynamic system simply attaches to the external port, where the end forcing $F(t)$ is now an internal force, as shown in Figure 10.11. Since integral causality still exists, equations for the entire interacting system can be straightforwardly derived. Of course, only a finite number of modes can be retained for the distributed part of the system.

If the rod end were terminated with a compliant member and then attached to another distributed rod, the model would be as shown in Figure 10.12, where n modes have been retained for each rod. Again, integral causality exists, and equation derivation is straightforward.

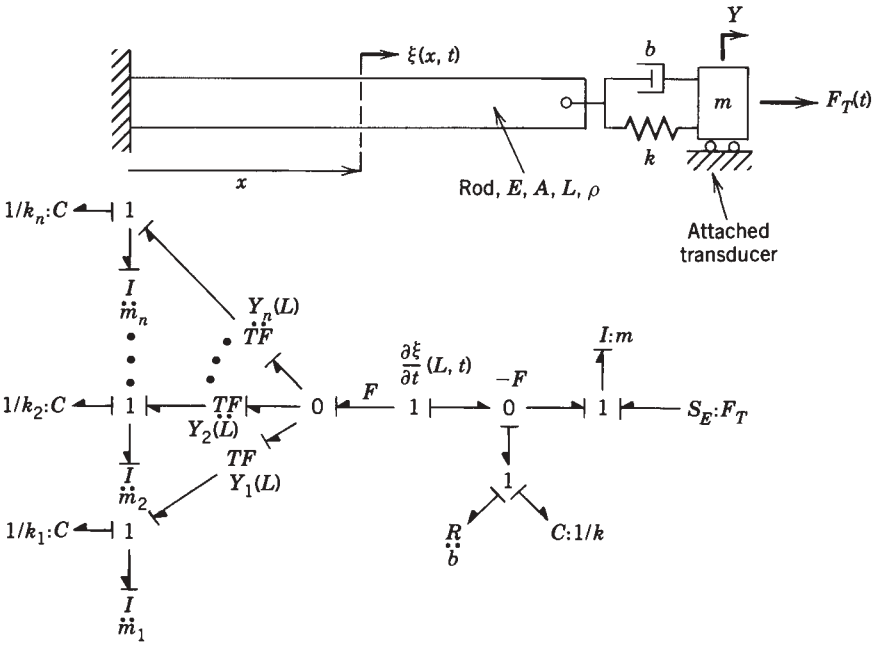


FIGURE 10.11. Finite-mode model of a rod with a lumped system attached at $x = L$.

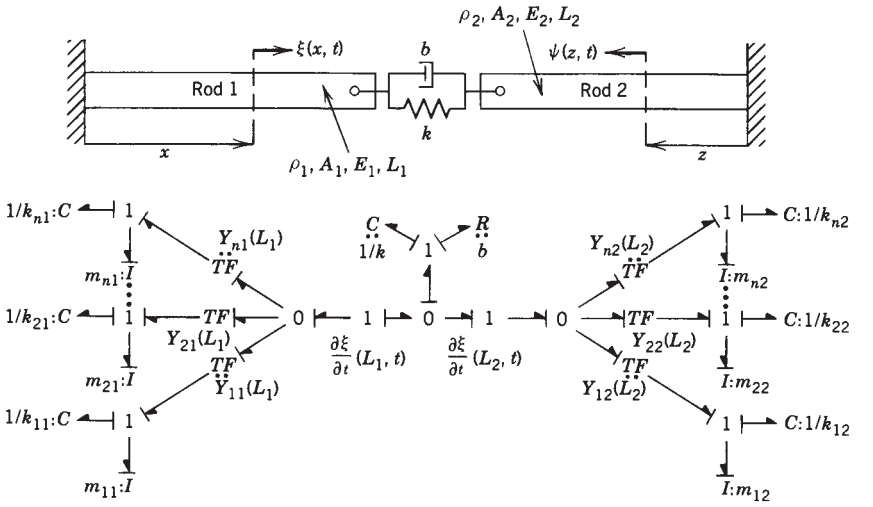


FIGURE 10.12. Two distributed rods interacting through a compliant and dissipative element.

The reader who is gaining some familiarity with bond graph procedures should be starting to appreciate the utility and elegance of constructing models using this finite-mode approach. We will allow this good feeling to persist a little longer before discussing some of the pitfalls of this approach.

Before leaving the rod example, it should be noted that any location on the rod can be used as an output. We simply consider any desired output point as a location for a force input equal to zero. The bond graph for two such output points, x_1 and x_2, is shown in Figure 10.13. As can be seen, drawing every bond in the bond graph can become very messy. In practice we do not include output points on the bond graph. Instead we use Eq. (10.69) and recognize that

$$\frac{\partial \xi}{\partial t}(x_i, t) = Y_1(x_i)\frac{p_1}{m_1} + Y_2(x_i)\frac{p_2}{m_2} + \cdots + Y_n(x_i)\frac{p_n}{m_n}. \tag{10.76}$$

Some excellent questions that may have arisen by now are

1. Is this approach only good for longitudinal vibration of rods fixed at one end and free at the other?
2. How many retained modes are enough?
3. What happens when causal flow inputs to the modes exist?

We will answer these questions in the following sections.

10.2.2 Bernoulli–Euler Beam Revisited

One criticism of the finite-mode approach for representing distributed system dynamics is that the problem must be reformulated any time the boundary conditions change. Also, the boundary conditions in physical systems are ill-defined.

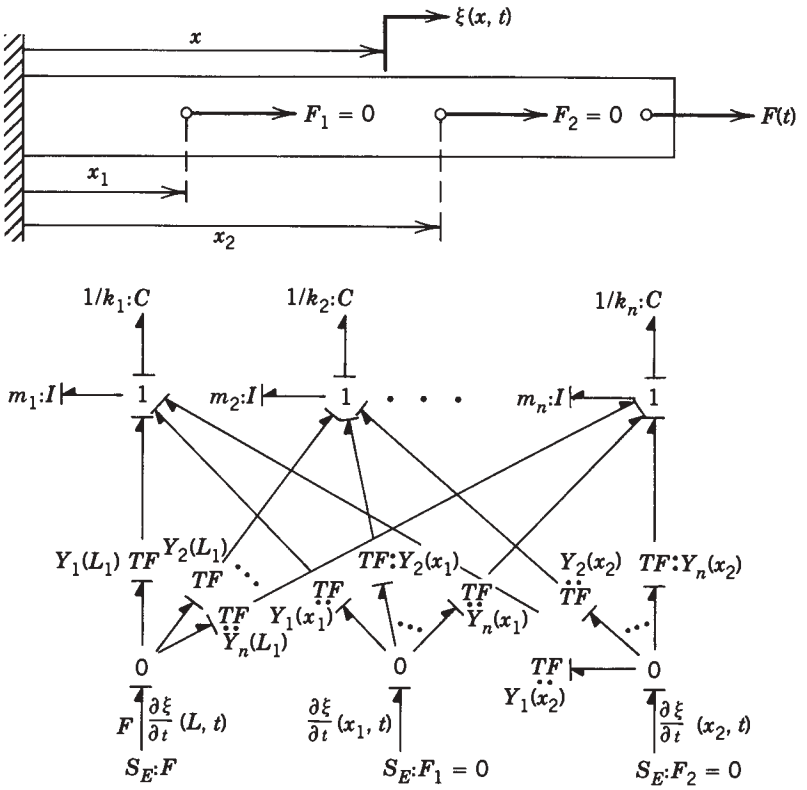


FIGURE 10.13. Bond graph showing how to obtain output at any point along the rod.

For instance, if we decide to represent a truck frame as a beam, there are no definite boundary conditions. There are only additional physical elements, such as suspension components, engine mounts, and cab mounts, attached at various points. The frame is not pinned or built in, so what modes should we use?

It turns out that the most general modes to use for finite-mode modeling of distributed systems are those modes associated with force-free boundaries. A heuristic argument is that modes associated with fixed boundaries, such as the modes for our bar example evaluated at the fixed end, can never produce a response at the fixed end other than zero, regardless of how hard we might push on that location. The modes are all zero at the fixed end, and from Eq. (10.58), the motion will always be zero there. Thus, if we decide that some previously assumed fixed boundary was not really fixed, but actually has a mass attached, then we must reformulate the problem and use different mode shapes.

On the other hand, if we had started with force-free modes and later decided that a boundary was attached to ground through a stiff spring, the force-free modes are perfectly capable of adding up, according to Eq. (10.58), to a very small motion. The motion can even be zero if we decide to fix a previously

force-free boundary. Thus, when using the finite-mode approach, it is best to use force-free modes unless the boundary conditions are well known and not subject to change.

The use of force-free modes adds a variation to the bond graph structure of Figure 10.13. This will be presented through the example of the Bernoulli–Euler beam described in Section 10.1. Unfortunately, the force-free modes are more difficult to derive than are modes for some other boundary conditions. However, once obtained (analytically, approximately, or even experimentally), they become part of our modeling arsenal, to be stored in our computer files, for use over and over again.

Figure 10.14 shows the Bernoulli–Euler beam acted upon by two external forces, $F_1(t)$, and $F_2(t)$. The equation of motion is repeated here from Eq. (10.35):

$$EI\frac{\partial^4 w}{\partial x^4} + \rho A\frac{d^2 w}{\partial t^2} = F_1\delta(x - x_1) + F_2\delta(x - x_2). \tag{10.77}$$

The force-free boundary conditions are that no shear force and no moment exist at $x = 0$ and $x = L$. From Eqs. (10.26), (10.27), and (10.28), these conditions can be written in terms of $w(x, t)$ as

$$\frac{\partial^2 w}{\partial x^2}(0, t) = \frac{\partial^2 w}{\partial x^2}(L, t) = 0 \tag{10.78}$$

for the zero-moment constraint and

$$\frac{\partial^3 w}{\partial x^3}(0, t) = \frac{\partial^3 w}{\partial x^3}(L, t) = 0 \tag{10.79}$$

for the zero-shear constraint.

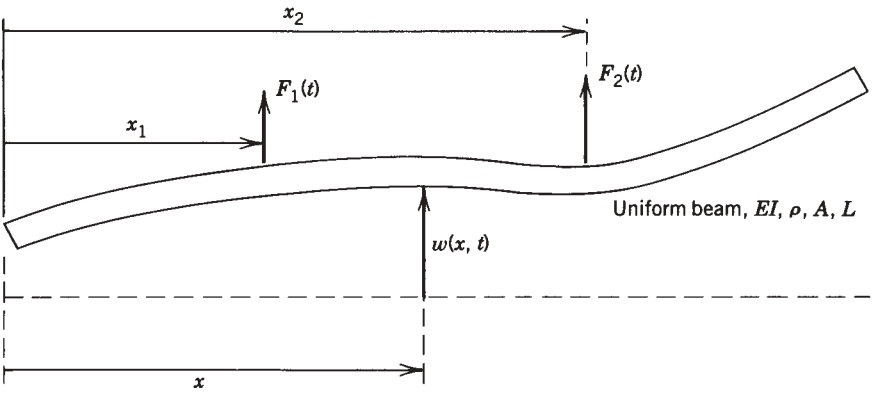


FIGURE 10.14. Uniform Bernoulli–Euler beam with point forcing.

Proceeding as we did in the bar example, we try a separated solution for the homogenous form of (10.77) and assume

$$w(x,t) = Y(x)f(t). \tag{10.80}$$

Substituting into (10.77) yields

$$EI\frac{d^4Y}{dx^4}f + \rho AY\frac{d^2f}{dt^2} = 0. \tag{10.81}$$

Dividing each term by ρAYf yields

$$\frac{EI}{\rho A}\frac{1}{Y}\frac{d^4Y}{dx^4} + \frac{1}{f}\frac{d^2f}{dt^2} = 0. \tag{10.82}$$

This equation can be satisfied at all x and all t only if both terms equal the same constant. By convention we let the second term equal $-\omega^2$, yielding

$$\frac{d^2f}{dt^2} + \omega^2 f = 0 \tag{10.83}$$

and

$$\frac{d^4Y}{dx^4} - \frac{\rho A}{EI}\omega^2 Y = 0, \tag{10.84}$$

or

$$\frac{d^4Y}{dx^4} - k^4 Y = 0, \tag{10.85}$$

where

$$k^4 = \frac{\rho A}{EI}\omega^2. \tag{10.86}$$

Equation (10.85) is a total differential equation, which, when solved subject to the boundary conditions (10.78) and (10.79), will yield the mode shapes and associated mode frequencies. Using Eq. (10.80) in the boundary conditions yields

$$\frac{d^2Y}{dx^2}(0) = \frac{d^2Y}{dx^2}(L) = \frac{d^3Y}{dx^3}(0) = \frac{d^3Y}{dx^3}(L) = 0. \tag{10.87}$$

The spatial equation (10.85) has the general solution [3]

$$Y(x) = A\cosh kx + B\sinh kx + C\cos kx + D\sin kx. \tag{10.88}$$

We will not cover all the algebra here. However, using Eq. (10.87) in Eq. (10.88) yields the frequency equation

$$\cosh k_n L \cos k_n L = 1, \tag{10.89}$$

and the mode shape functions

$$Y_n(x) = (\cos k_n L - \cosh k_n L)(\sin k_n x + \sinh k_n x) - (\sin k_n L - \sinh k_n L)(\cos k_n x + \cosh k_n x). \tag{10.90}$$

We now solve (10.89) for the special values of $k_n L$, and use these in (10.86) to obtain the mode frequencies as

$$\omega_n^2 = \frac{EI}{\rho A}\frac{(k_n L)^4}{L^4}. \tag{10.91}$$

At this point we must recognize that the force-free boundary conditions permit $\omega_n = k_n = 0$ to be a mode frequency. Using this information yields

$$\frac{d^4 Y}{dx^4} = 0, \tag{10.92}$$

or

$$Y = c_1 x^3 + c_2 x^2 + c_3 x + c_4. \tag{10.93}$$

There are two possible solutions to (10.93) that satisfy the boundary conditions: They are $Y = \text{const}$ and $Y = ax + b$. These are called rigid-body modes, and it is convenient to think of them as rigid-body vertical translation of the entire beam and rigid-body rotation of the beam about the centrally located center of mass. Thus,

$$Y_{00} = 1, \tag{10.94a}$$

and

$$Y_0 = x - \frac{L}{2}. \tag{10.94b}$$

The bending modes are given by (10.90) and are rather complicated functions. We will find that the rigid-body modes are orthogonal in the same sense as previously presented,

$$\int_0^L \rho A Y_n(x) Y_m(x) dx = 0, n \neq m, \tag{10.95}$$

for $n, m = 00, 0, 1, 2, 3, \ldots$

The calculation of the forced response proceeds in a similar fashion to the bar example. We assume that the forced solution has the form

$$w(x, t) = \sum_{n=0}^{\infty} Y_n(x)\eta_n(t) \tag{10.96}$$

and use this in the forced equation (10.77). We then multiply each term by $Y_m(x)$ and integrate with respect to x from $x = 0$ to $x = L$. Then we use the orthogonality property of the modes and obtain

$$\left(\int_0^L \rho A Y_n^2 dx\right) \ddot{\eta}_n + \left(\int_0^L \rho A Y_n^2 dx\right) \omega_n^2 \eta_n = F_1 Y_n(x_1) + F_2 Y_n(x_2), \quad (10.97)$$

which has exactly the same form as Eq. (10.66) for the bar example. The first zero-frequency mode yields

$$\left[\int_0^L \rho A(1)^2 dx\right] \ddot{\eta}_{00} = F_1 + F_2, \quad (10.98)$$

or

$$m\ddot{\eta}_{00} = F_1 + F_2, \quad (10.99)$$

which simply states that the external forces accelerate the center of mass of the beam. The other zero-frequency mode yields

$$\left[\int_0^L \rho A \left(x - \frac{L}{2}\right)^2 dx\right] \ddot{\eta}_0 = F_1 \left(x_1 - \frac{L}{2}\right) + F_2 \left(x_2 - \frac{L}{2}\right), \quad (10.100)$$

or

$$J_g \ddot{\eta}_0 = F_1 \left(x_1 - \frac{L}{2}\right) + F_2 \left(x_2 - \frac{L}{2}\right), \quad (10.101)$$

which simply states that the moment of the external forces about the center of mass produces angular acceleration $\ddot{\eta}_0$, where J_g is the centroidal moment of inertia of the beam.

The bond graph for a beam with force-free boundary conditions is shown in Figure 10.15. The rigid-body modes, which will be present whenever force-free boundaries are assumed, appear simply as inertia elements with no associated modal stiffness. The first inertia parameter is beam mass m, and the second is the beam moment of inertia J_g. The —*TF*— elements connected to the rigid-body modes correctly apply the forces and moments to these elements. The rest of the structure is identical to the bar example. The modal masses are still, symbolically,

$$m_n = \int_0^L \rho A Y_n^2 dx, n = 1, 2, \ldots, \quad (10.102)$$

and the modal stiffnesses are still

$$k_n = m_n \omega_n^2.$$

Only this time the modes Y_n are given by Eq. (10.90), and the frequencies ω_n come from solving Eq. (10.89). Nobody said life would be easy, but the

integration required in (10.102) need only be carried out once in your life in the nondimensional form

$$\frac{m_n}{m} = \int_0^1 Y_n^2\left(\frac{x}{L}\right) d\frac{x}{L} \tag{10.103}$$

and then stored away for use whenever needed. The integration would probably be carried out numerically anyway.

The true virtue and elegance of this modeling procedure comes when we recognize that the external forces, F_1 and F_2 in this example, probably result from some attached dynamic system and can be appended to Figure 10.15 in straightforward bond graph fashion. For instance, the beam in Figure 10.16 has suspension components at the two ends and a one-degree-of-freedom system attached at $x = x_F$. Perhaps this is an optical bench with a laser turning mirror at $x = x_F$, and we are trying to design the suspension to isolate the mirror from ground motion. The finite-mode bond graph is also shown in Figure 10.16. We simply treat all attached components as external forcing and show these with 0-junctions for each force location. Each 0-junction is attached to each 1-junction with the modal components, through —*TF*— elements with moduli equal to appropriate mode functions evaluated at the location of the force. The attached systems are then appropriately appended to the external 0-junction bonds to form a complete, low-order, very accurate system model for all frequencies of excitation up to the highest natural frequency included o the model.

While we have not been as general in our development as is possible, it should be realized now that the structure of Figure 10.15 is entirely general and represents *all* distributed structures that exhibit normal-mode behavior. The mode shapes change and the modal masses and stiffnesses change from system to system. But the bond graph structure is invariant. The modes can be obtained analytically, as was done here for the bar and the beam, or they can be obtained approximately, as described in Reference [4] for vibrating plates. They can be

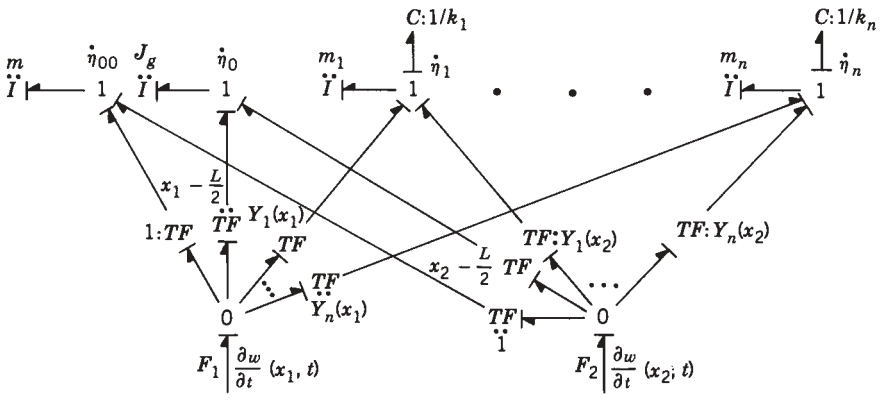


FIGURE 10.15. Bond graph for Bernoulli–Euler beam with force-free boundaries.

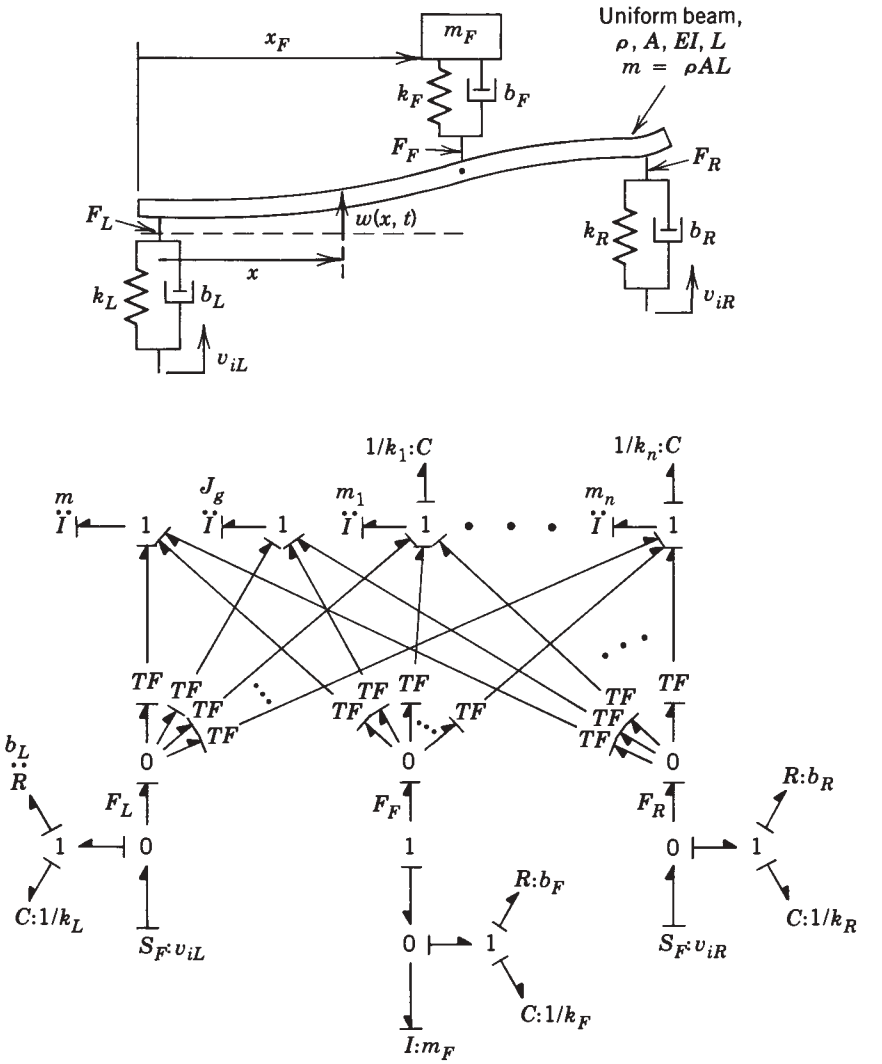


FIGURE 10.16. An example of a distributed system interacting with attached lumped systems.

obtained by finite-element methods and then used in the bond graph structure to construct low-order, understandable models. And the modes can even be obtained experimentally using structural dynamic testing devices and programs. No matter how the modes are obtained, their finite-mode representation is as shown in Figure 10.15, and they are ready to interact with any external dynamic elements, even nonlinear ones.

Also, while not derived here, external moments can also be applied to the structure of Figure 10.15. If a moment is applied instead of a force, the

associated —*TF*— elements, coming from the moment 0-junction, have the mode shape slopes, (dY_n/dx), evaluated at the location of the moment as their moduli instead of the mode shapes themselves.

10.3 GENERAL CONSIDERATIONS OF FINITE-MODE BOND GRAPHS

The reader who has begun to appreciate the virtues of bond graph modeling with regard to ease in putting all types of systems together, ease in identifying physical state variables, ease in deriving state equations, and ease in obtaining computer solutions through use of some very automated software has certainly begun to appreciate the ease with which distributed systems can be made to interact with dynamic subsystems. Some questions still remain in the practical application of the finite-mode concept, and some of the most prominent are answered in this section.

10.3.1 How Many Modes Should Be Retained?

When a distributed component is part of an overall physical dynamic system model, we often have some idea of the frequency band of interest. This comes from our engineering judgment, and, unfortunately, there is no substitute for exercising this judgment when we construct any system model, regardless of the modeling approach. If we were constructing a model of a long-haul truck interacting with the roadway unevenness, and decided to model the frame as a beam with the wheels attached through their suspension elements, we would make use of our knowledge that input frequencies from the road rarely exceed 20 Hz at normal driving speeds, wheel hop in trucks is around 15 Hz, and suspension frequencies are usually less than 2 Hz. It would then be obvious that very high-frequency beam modes would contribute very little to the response in the 20-Hz frequency range of interest.

We can quantify this thought somewhat by considering how a lightly damped mode responds to a harmonic input. Figure 10.17 shows a general finite-mode representation with one input location. The equation governing the contribution of the ith mode is

$$m_i\ddot{\eta}_i + k_i\eta_i = FY_i. \tag{10.104}$$

If there is some damping associated with the mode, then (10.104) can be written as

$$\ddot{\eta}_i + 2\xi_i\omega_i\dot{\eta}_i + \omega_i^2\eta_i = \frac{F}{m_i}Y_i, \tag{10.105}$$

where ξ_i is the mode damping ratio.

For F harmonic, the mode frequency response for $\dot{\eta}_i$ has the magnitude

$$\left|\frac{\dot{\eta}_i}{\frac{F}{m_i}Y_i}\right| = \frac{\frac{\omega}{\omega_i}\frac{1}{\omega_i}}{\left[\left(1-\frac{\omega^2}{\omega_i^2}\right)+\left(2\xi_i\frac{\omega}{\omega_i}\right)^2\right]^{1/2}}. \tag{10.106}$$

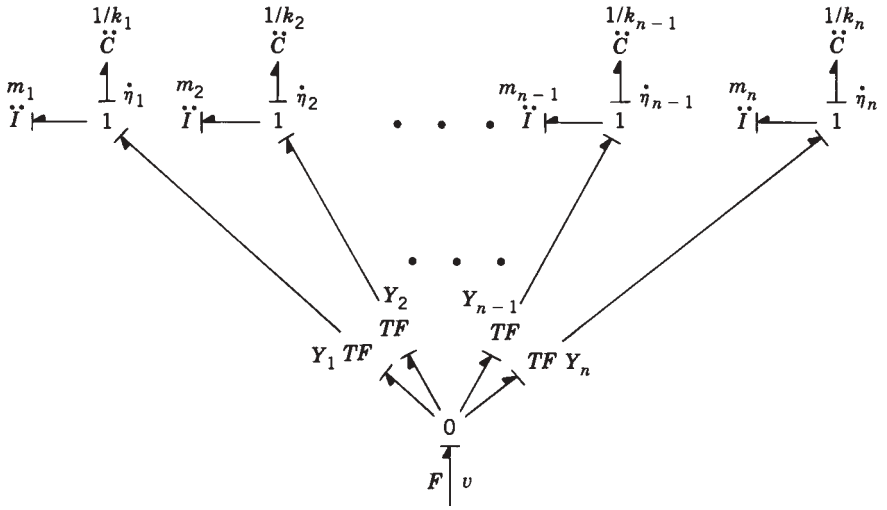


FIGURE 10.17. General finite-mode bond graph.

This response is sketched in Figure 10.18. At low excitation frequency, (10.106) becomes

$$\dot{\eta}_i|_{\omega/\omega_i \ll 1} \sim Y_i F \frac{1}{m_i \omega_i^2} \omega = Y_i \frac{F}{k_i} \omega, \tag{10.107}$$

and the response is governed by the modal stiffness, k. We call this low-frequency range the *stiffness-controlled region* in Figure 10.18. At frequencies near ω_i, the response is

$$\dot{\eta}_i|_{\omega/\omega_i \cong 1} \cong \frac{F}{m_i} Y_i \frac{1}{2\xi_i \omega_i} F Y_i \frac{1}{b_i}, \tag{10.108}$$

where b_i is the mode damping constant, $b_i = 2\xi_i \omega_i m_i$. We call the frequency range near $\omega/\omega_i = 1$ the *resistance-controlled region*. Finally, at frequencies high compared to ω_i, we get

$$\dot{\eta}_i|_{\omega/\omega_i \gg 1} \sim \frac{F}{m_i} Y_i \frac{1}{\omega}, \tag{10.109}$$

and we see that the response is governed by the modal mass only. This high-frequency range is called the *mass-controlled region*.

If we have reasonable knowledge of input frequency content, then we know which modes will be excited significantly below their mode frequencies and will therefore be in the stiffness-controlled region. For those infinity of modes in the stiffness-controlled region, their inertia and resistance are of no consequence and can be ignored. At this point we can either prune these high-frequency modes

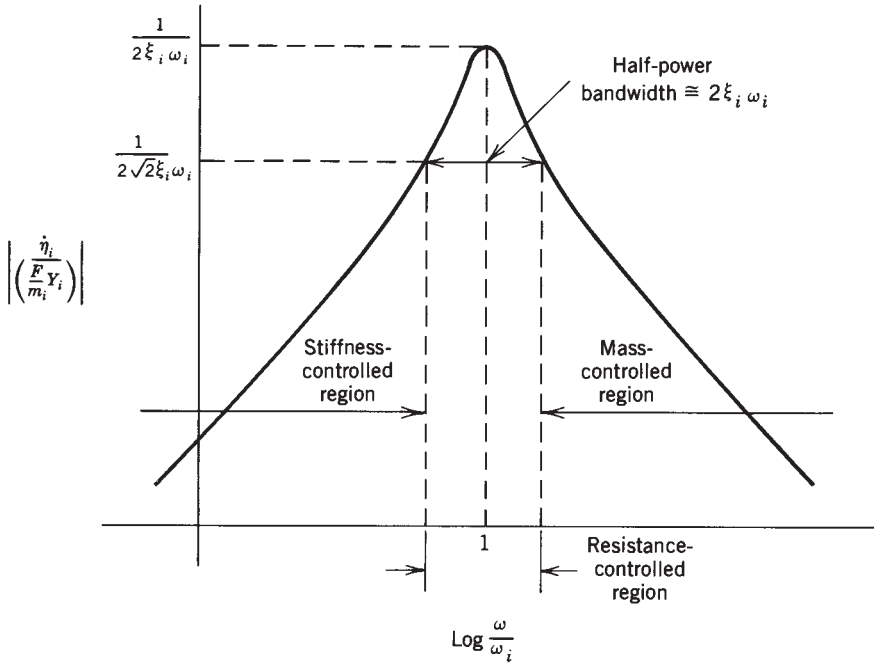


FIGURE 10.18. Frequency response of a single mode.

altogether or choose to retain their modal stiffness to improve the static stiffness representation of the distributed system.

In fact, for distributed systems with only one input location, the residual compliance due to the infinity of stiffness-controlled modes can be calculated straightforwardly. Consider Figure 10.19, where n dynamic modes have been retained, and n has been selected to extend well beyond the frequency range of interest. The remaining modes are represented by their modal compliance only. Notice that all the stiffness-controlled modes are in derivative causality.

The contribution of the modal flows ($\dot{q}_{n+1}, q_{n+2}, \ldots$) of the stiffness-controlled modes to the output velocity v we will call the *residual velocity*, v_R. From the bond graph,

$$v_R = Y_{n+1}\dot{\eta}_{n+1} + Y_{n+2}\dot{\eta}_{n+2} + \cdots, \tag{10.110}$$

but

$$\dot{\eta}_{n+1} = \frac{1}{k_{n+1}}\dot{F}_{n+1}, \dot{\eta}_{n+2} = \frac{1}{k_{n+2}}\dot{F}_{n+2}, \text{etc.} \tag{10.111}$$

and

$$F_{n+1} = Y_{n+1}F, F_{n+2} = Y_{n+2}F, \text{etc.} \tag{10.112}$$

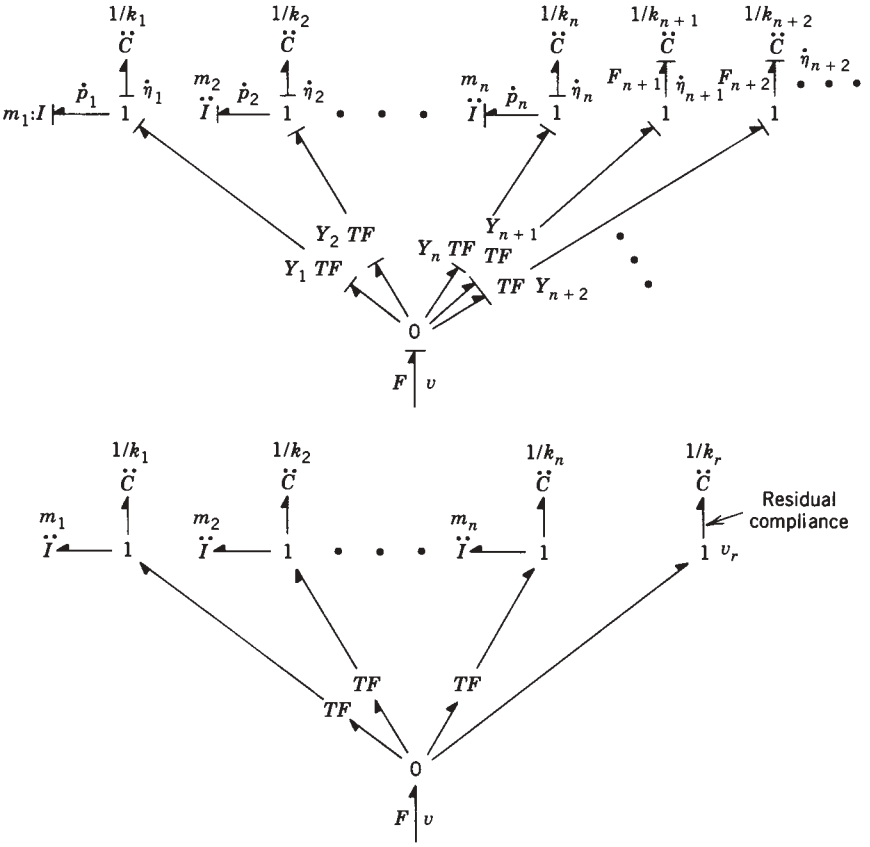


FIGURE 10.19. Finite-mode bond graph including residual compliance.

Thus,

$$v_R = \frac{Y_{n+1}^2}{k_{n+1}}\dot{F} + \frac{Y_{n+2}^2}{k_{n+2}}\dot{F} + \cdots, \tag{10.113}$$

or

$$v_R = \left(\sum_{j=1}^{\infty} \frac{Y_{n+j}^2}{k_{n+j}}\right)\frac{d}{dt}F. \tag{10.114}$$

This same contribution to the output velocity would occur on introducing the single residual stiffness element k_r, where

$$\frac{1}{kr} = \sum_{j=1}^{\infty} \frac{Y_{n+j}^2}{k_{n+j}}, \tag{10.115}$$

as shown also in Figure 10.19. The summation of Eq. (10.105) converges very quickly and is easily performed numerically.

It should be pointed out that, when more than one input is acting on the distributed system, the residual compliance is not calculated from the simple formula (10.115). Instead, a C-field representation is needed for the stiffness-controlled modes. The field constitutive properties can be determined, but frequently it is easier to simply eliminate the pruned modes altogether.

As a rule of thumb, it is recommended that modes be retained up to a frequency at least a factor of 2, but no more than a factor of 5, higher than the highest frequency of interest.

10.3.2 How to Include Damping

The normal-mode representation for distributed system dynamics has been formulated without damping. And, strictly speaking, separation of variables and the general bond graph structure of Figures 10.15 and 10.17 are correct only in the absence of damping altogether or the presence of damping in a very special form. In general, distributed elements that we want to include in overall system models are typically lightly damped. Lightly damped modes are the ones that cause undesirable system response or difficulties for control systems.

We cannot identity exactly from where on the structure the damping arises. We only know for certain that, once excited, the energy will be dissipated. The mechanisms are complex; some dissipation is due to minute plastic deformations of the material, and some to the radiation of energy from material surfaces. One thing that is certain is that there are no little dampers attached to the structure that one can point to and therefore include in a model in order to represent the damping of the distributed elements.

It is customary to not attempt to perform detailed modeling of the damping mechanisms, but instead to include damping functionally by incorporating it into the individual modes. This is accomplished by simply appending R-elements to the mode oscillators of the modal bond graph. The general structure of Figure 10.17 will then become as shown in Figure 10.20. Treating each mode oscillator as an individual single-degree-of-freedom system, the resistance parameter R_i is set by

$$R_i = 2\xi_i \omega_i m_i, \tag{10.116}$$

where ξ_i is the modal damping ratio. Typically structures are lightly damped, with ξ_i in the range of 0.01–0.1, where the high value would perhaps be representative of some composite structures. If experimental data were available, we could improve upon this seemingly crude inclusion of damping. However, usually the inclusion of modal damping as described above produces very accurate and useful models.

10.3.3 Causality Consideration for Modal Bond Graphs

From Chapter 2, the reader should be familiar with the concept of causality and with the useful information it supplies with regard to equation formulation and the computability of one's model. We know that integral causality throughout a

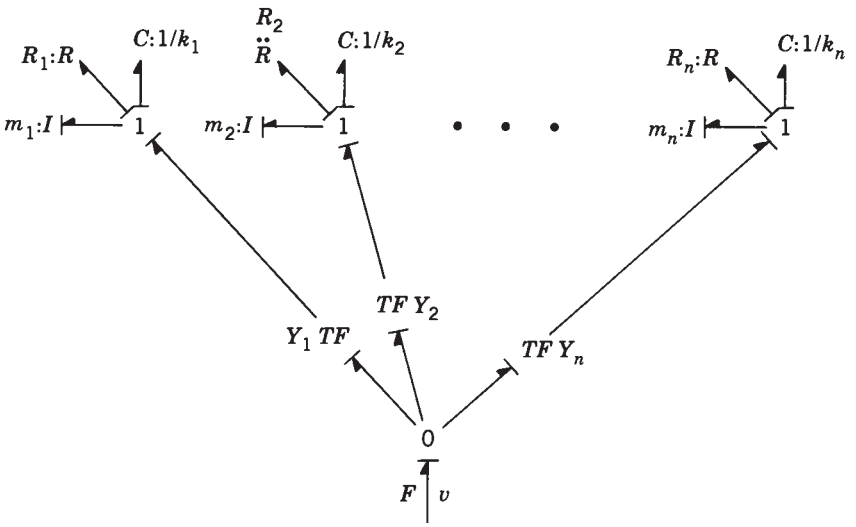


FIGURE 10.20. General finite-mode bond graph with modal damping.

model indicates that formulation will be straightforward. If any derivative causality exists or some arbitrary assignment of causality is required, then we know that algebraic problems exist in the model, and formulation of equations can be difficult, if not impossible, when nonlinear elements are involved.

Another implication of causality relates to the general unification of some part of a system model with the models of other parts of the system. For instance, the bond graph modal representation of Figure 10.20 is not a complete model, because we intend to connect the external bond, carrying F and v, with some other dynamical elements. In Figure 10.20, the force F is a causal input to the modes, and the velocity v is a causal output. The bond graph possesses all integral causality. We thus could straightforwardly formulate the state equations for this bond graph fragment, and, as long as we agreed to use this fragment only in its present causality, these equations would never change. We would simply need to know the input F and the equations would deliver the output v. Of course, any attached dynamical model will have some causal restrictions, since it must have F as an output and accept v as an input. Such is the construction of computable models.

As we have seen in every finite-mode model presented so far, integral causality exists for all I- and C-elements, and the effort variable has been the causal input on all external bonds. Consider Figure 10.21, where a general n-mode representation is shown along with two external input locations. The causal effort F_1 is an input at one input location, but the causal flow v_2 is an input at the other location. The reader should assign causality to show that, with a causal flow input, one I-element in the modes must be in derivative causality. For this example, the nth modal inertia has been selected to be put into derivative causality. In

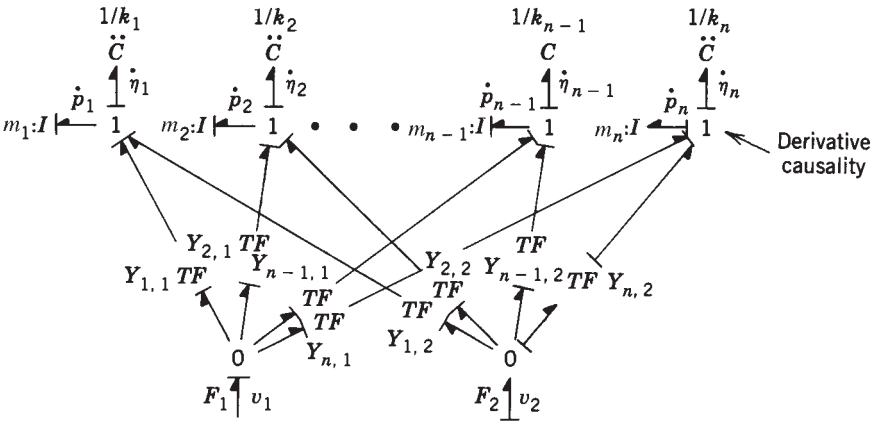


FIGURE 10.21. Finite-mode bond graph with causal flow input at one input location.

general, it can be shown that for each additional causal flow input to the modes, one additional modal inertia will be forced into derivative causality.

One result of the derivative causality is that a state variable is lost, since p_n is no longer an indepent state variable, and the accuracy of the n-mode description is somewhat reduced. This really presents no problem, since we can always include additional modes and, in fact, have probably already included more than enough modes if recommended procedures have been followed in formulating the original model.

The real disadvantage of allowing derivative causality is in the formulation of system equations for analysis or simulation. If we attempt to use procedures described in Chapter 5 to deal directly with the derivative causality, we write

$$p_n = m_n \dot{\eta}_n \tag{10.117}$$

and follow the causality for $\dot{\eta}_n$, obtaining

$$p_n = m_n \frac{1}{Y_{n,2}} \left(v_2 - Y_{n-1,2} \frac{P_{n-1}}{m_{n-1}} - \cdots - Y_{2,2} \frac{p_2}{m_2} - Y_{1,2} \frac{p_1}{m_1} \right). \tag{10.118}$$

We see that p_n depends on all $n-1$ other modal momenta in the mode description, and it also depends on the external input, v_2. If we attempt to write the state equations for I-elements in integral causality, we find that $\dot{p}_1, \dot{p}_2, \ldots, \dot{p}_{n-1}$ will all require $\dot{p}_n$ in their equations, and $\dot{p}_n$, from (10.118), is

$$\dot{p}_n = m_n \frac{1}{Y_{n,2}} \left(\dot{v}_2 - Y_{n-1,2} \frac{\dot{p}_{n-1}}{m_{n-1}} - \cdots - Y_{2,2} \frac{\dot{p}_2}{m_2} - Y_{1,2} \frac{\dot{p}_1}{m_1} \right). \tag{10.119}$$

Thus, all the modal momenta are coupled, and obtaining explicit state equations becomes very difficult. To make matters worse, the derivative of the input velocity, $\dot{v}_2$, is also required. The velocity v_2 probably is causally determined by some

inertial elements external to the modes; thus, additional momentum variables will couple with the modal momenta, making the algebraic problem worse. In fact, this coupling of dynamic elements external to the mode structure practically precludes the derivation of equations for the modes as separate entities to be used over and over with different external dynamic systems.

A matter of computational importance is that when derivative causality is permitted, we invariably end up with a mode shape factor (transformer modulus) in the denominator of the algebraic expression (10.119). In Eq. (10.119), $Y_{n,2}$ (mode shape n evaluated at position 2) is in the denominator of the right-hand side. If it should happen that $Y_{n,2}$ is zero or very small (position 2 is at or very near a node of mode n), then computational problems will result. Reference [5] discusses this problem further. The bottom line of this discussion is that it would be very nice to avoid derivative causality when using finite-mode models for distributed systems.

It turns out that the derivative causality can be avoided while still permitting causal flow inputs to the mode. We simply include an additional modal compliance in the model *without* the associated modal inertia. And we include as many additional modal compliances as there are external inputs with flow-in causality. This addition of modal C-elements without their associated I-elements will pose no accuracy problems, since the original choice of n modes was based on ensuring that the higher frequency modes were already stiffness controlled. Thus, additional modes are certainly stiffness controlled, and their modal inertia is of no consequence to the system response.

Figure 10.22 shows the general n-mode structure of Figure 10.21 repeated with the addition of the $(n+1)$st modal compliance. Notice that v_2 is still a causal

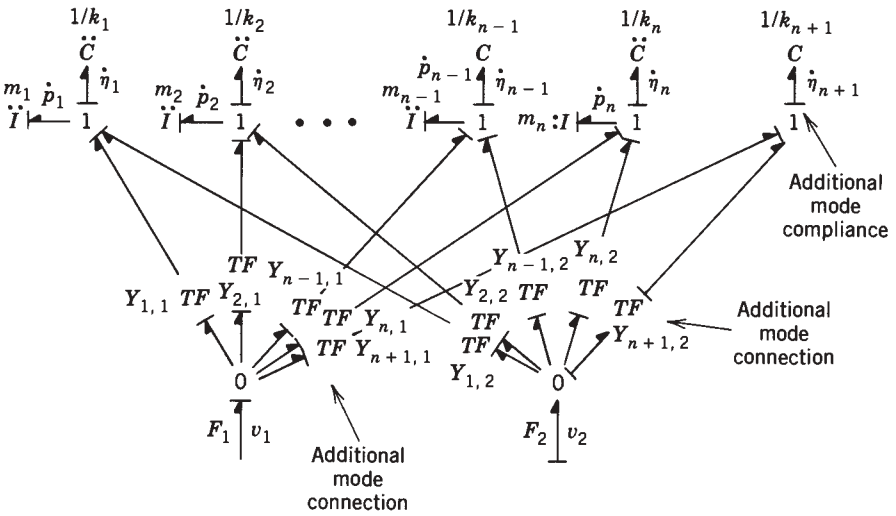


FIGURE 10.22. Finite-mode bond graph with causal flow input plus an additional modal compliance.

flow input to the modes, but no derivative causality exists. Thus, no algebraic problems exist, and equation formulation is straightforward.

Before leaving this section it must be admitted that the act of appending additional modal compliances to avoid causality problems does present an algebraic problem when more than one causal flow input is present. Figure 10.23 exposes this problem. In part *a*, a three-mode model is shown with two causal flow inputs. Readers should convince themselves that the only permissible causality forces modal inertias 2 and 3 into derivative causality. Figure 10.23*b* shows the

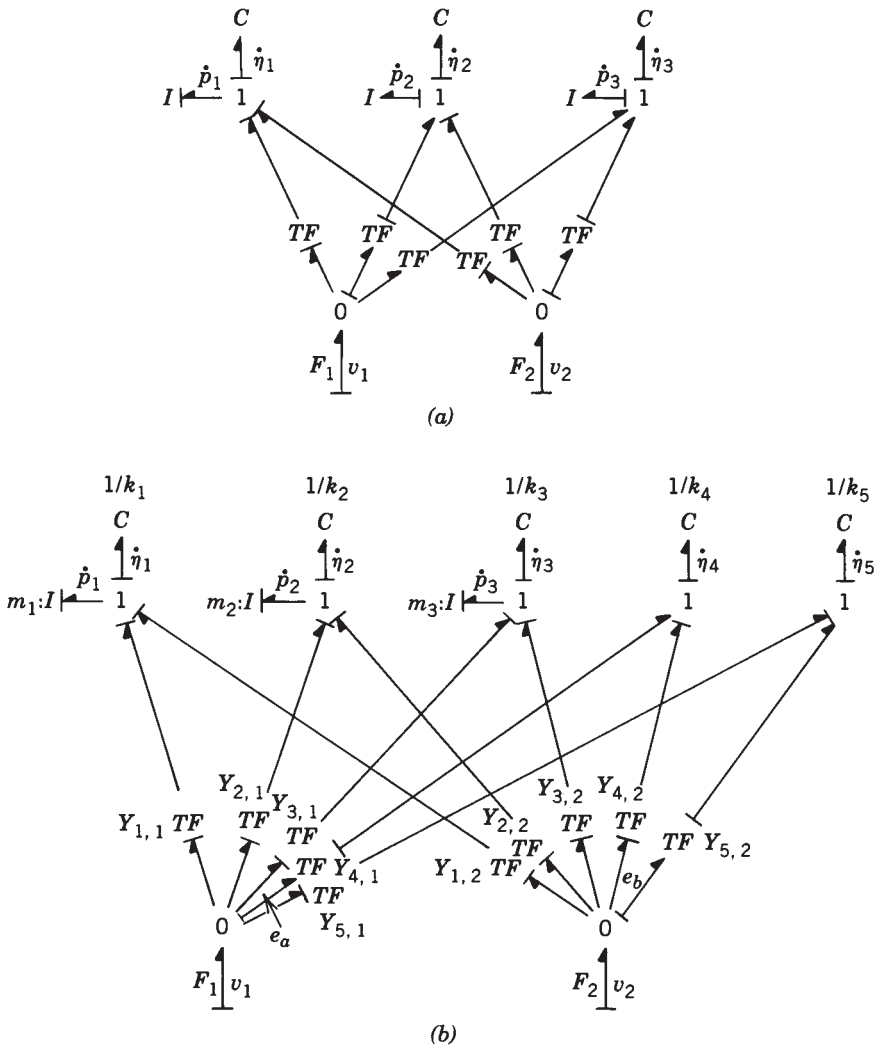


FIGURE 10.23. Finite-mode bond graph with two causal flow inputs.

same model after using the trick of appending two additional modal compliances without their associated inertias. Readers should assign causality to this model and show that, after assigning integral causality to all I- and C-elements, there are still some unassigned bonds. This occurrence always indicates the presence of an algebraic loop. The effort e_a has been chosen arbitrarily to set the effort on the 0-junction indicated, and, as a result, the causal picture is complete. The effort e_b has also been labeled in Figure 10.23*b*.

Unlike the algebraic problem when derivative causality is permitted, the present algebraic loop can be solved, once and for all, internally to the bond graph structure. It is straightforward to derive

$$e_a = \frac{1}{Y_{4,1}}(k_4\eta_4 - Y_{4,2}e_b) \tag{10.120}$$

and

$$e_b = \frac{1}{Y_{5,2}}(k_5\eta_5 - Y_{5,1}e_a). \tag{10.121}$$

This can obviously be solved explicitly for e_a and e_b and then used in the state-space equation formulation. The algebraic problem created by appending extra modal compliances in order to avoid derivative causality is very manageable, while dealing with the derivative causality is generally not.

10.4 ASSEMBLING OVERALL SYSTEM MODELS

In the previous sections a procedure was developed for representing all types of distributed continua as finite-mode bond graph models. Only beams and bars were introduced through example; nevertheless, multidimensional elements can be represented with the same bond graph structure as was developed for beams and bars. Only the mode shapes, modal masses, and modal frequencies differ from continuum to continuum. This statement is made somewhat tongue in cheek, since obtaining modal information for multidimensional structures can be virtually impossible.

If an overall structure is composed of many (more or less conventional) elements, then an overall structural model can be constructed by appropriately connecting finite-mode models of the elements. Figure 10.24 shows an A-frame for a ground vehicle composed of three uniform beams. In Figure 10.24*b* the frame is shown with internal forces and moments exposed. The forces F_1, F_2, and F_3 are external forces which ultimately will represent suspension forces coming through the force locations. We are assuming that beams 1 and 2 do not twist due to the motion of beam 3.

In Figure 10.25 a first cut at an assembled model is shown. The words "modes of beam 1" and so forth are used to represent the 1-junctions with attached modal

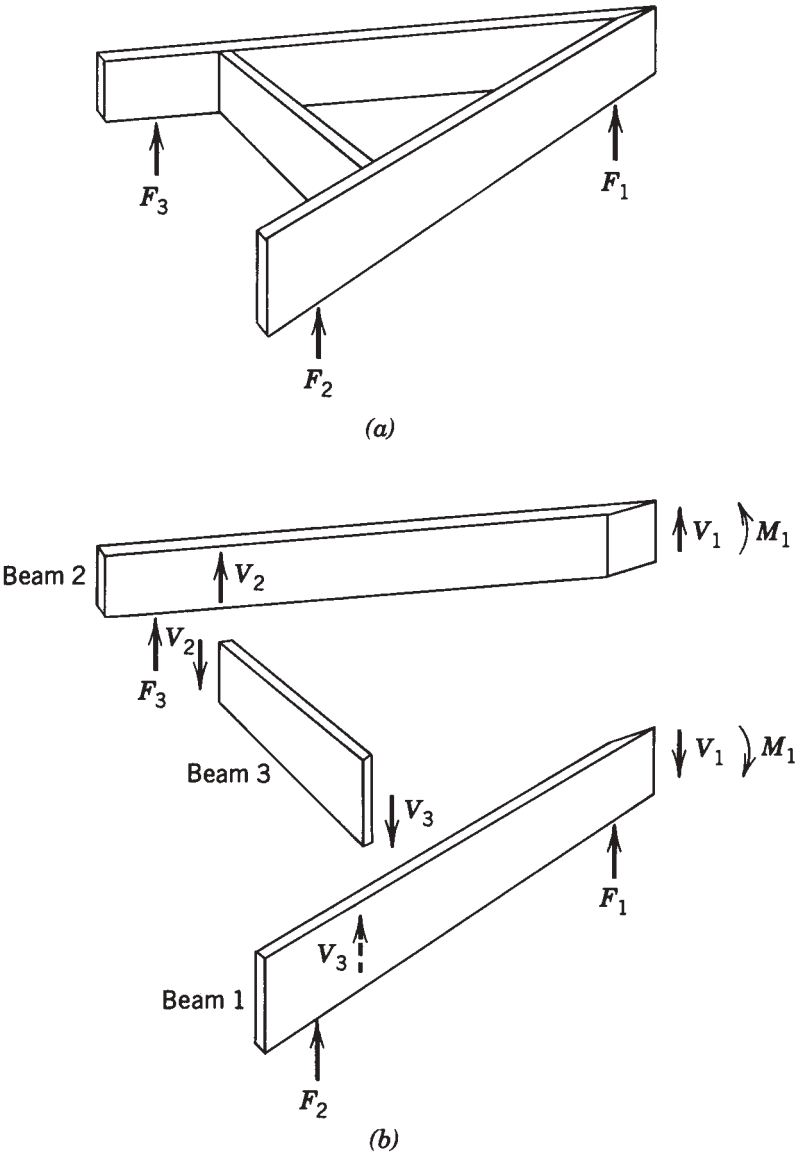


FIGURE 10.24. An A-frame composed of three uniform beams.

oscillators and the two rigid-body inertias. The forces and moments are shown with 0-junctions, and the fan of bonds emanating from the 0-junctions must connect to each 1-junction in the appropriate mode set through transformers whose moduli are force-free beam mode shapes evaluated at the location of the

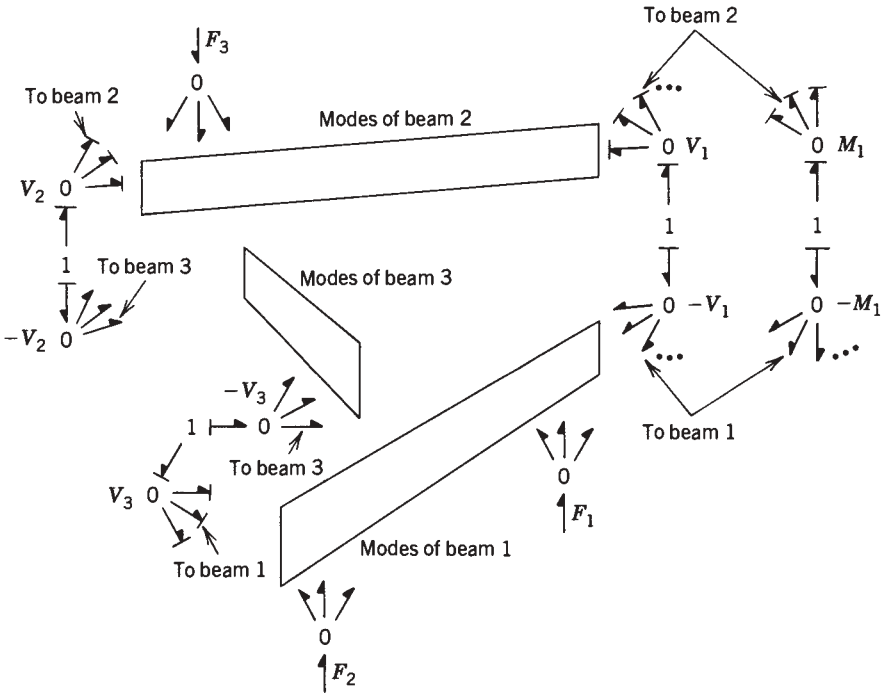


FIGURE 10.25. Schematic diagram of the finite-mode bond graph for the A-frame.

forces. Thus, V_1 acts on beams 2 and 1, M_1 acts on beams 2 and 1, V_2 acts on beams 2 and 3, and V_3 acts on beams 3 and 1. Causality considerations tell us that derivative causality will exist because beam contact interfaces will share common velocities and angular velocities. Thus, our second and final model will have additional modal compliances to rid us of the derivative causality. Reference [6] shows this example fully developed.

As a final example (Figure 10.26), consider a longitudinal bar attached at $z = z_1$ to a uniform beam. A spring k is attached to the beam at one end, and a velocity excitation $v_i(t)$ exists at the left end of the bar. Frequency considerations have led us to a two-mode (plus the rigid-body mode) representation for the bar and a one-mode (plus two rigid-body modes) representation for the beam. Force-free modes are used for both distributed elements in order to have the most general representation possible. The final bond graph is also shown in Figure 10.26. Notice that the bar has one causal flow input; thus one additional modal compliance is included for the bar. The beam also has one causal flow input; thus, the beam modes also require an additional modal compliance. The overall model has all integral causality and is ready for straightforward formulation.

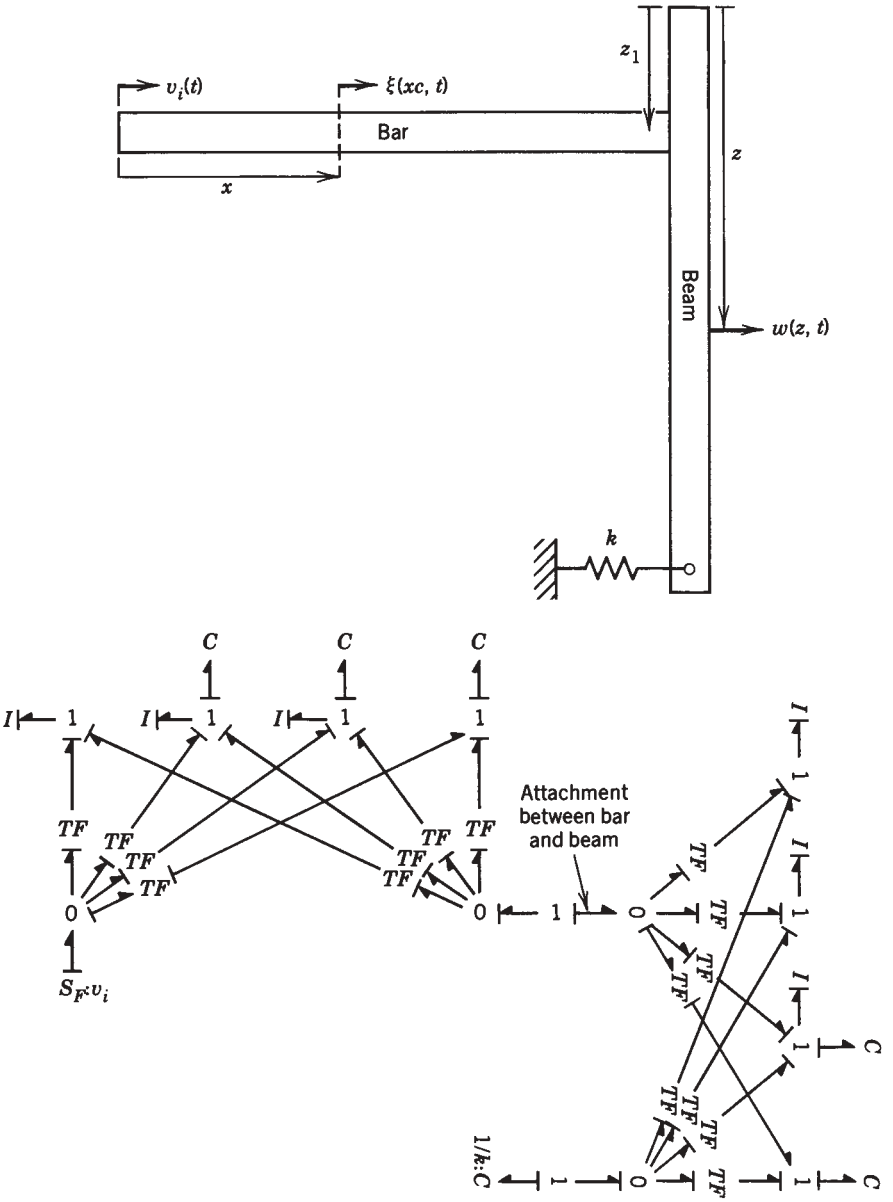


FIGURE 10.26. A bar interacting with a beam.

10.5 SUMMARY

This chapter has dealt with the modeling of inherently distributed systems so that the distributed elements could be included in an overall system model in a reasonable manner. The first modeling approach described dealt with the use of finite lumps, contiguously aligned, and retained in sufficient numbers that low-frequency accuracy would result. This approach is all right for very crude approximations, using one or two lumps, but not practical when many lumps are required.

The main emphasis of this chapter is the finite-mode approach for obtaining very accurate, low-order, overall system models incorporating all types of distributed and lumped dynamic elements. It is hoped that the reader appreciates the elegance and utility of this approach. If not, it is suggested that References [6–8] be read. These show applications of finite-mode modeling to very realistic systems, such as the assembled frame of a ground vehicle, the dynamics of the structure holding the world's largest laser, and the dynamics of the human lung.

REFERENCES

[1] D. L. Margolis, "A survey of bond graph modeling for interacting lumped and distributed systems," *J. Franklin Inst.*, **19**, No. 1/2 (Jan. 1985).

[2] H. Crandall et al., *Dynamics of Mechanical and Electromechanical Systems*, New York: McGraw-Hill, 1968.

[3] L. Meirovitch, *Analytical Methods in Vibrations*, New York: Macmillan, 1967.

[4] A. W. Leissa, "Vibration of plates," NASA Report SP-160, 1969.

[5] D. L. Margolis, "Bond graphs for distributed system models admitting mixed causal inputs," *ASME J. Dyn. Syst. Meas. Control*, **102**, No. 2, 94–100 (June 1980).

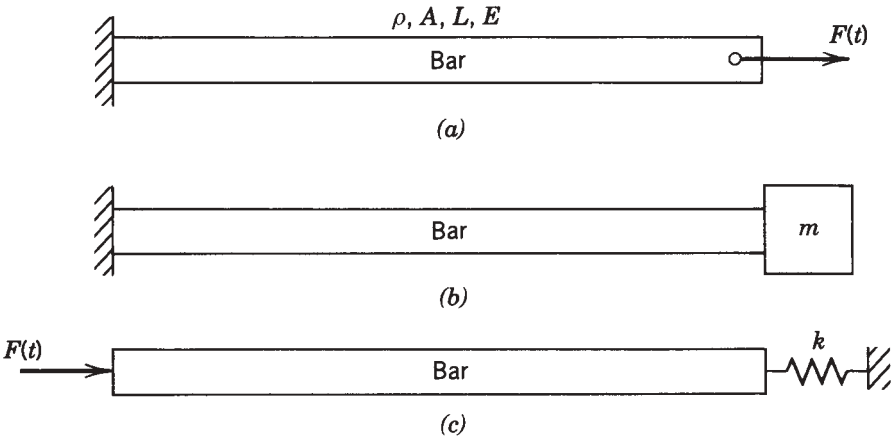
[6] D. L. Margolis, "Bond graphs, normal modes, and vehicular structures," *Vehicle Syst. Dyn.*, **7**, No. 1, 49–63 (1978).

[7] D. L. Margolis, "Dynamical models for multidimensional structures using bond graphs," *ASME J. Dyn. Syst. Meas. Control*, **102**, No. 3, 180–187 (Sept. 1980).

[8] D. L. Margolis and M. Tabrizi, "Acoustic modeling of lung dynamics using bond graphs," *J. Biomech. Eng.*, **105**, No. 1, 84–91 (Feb. 1983).

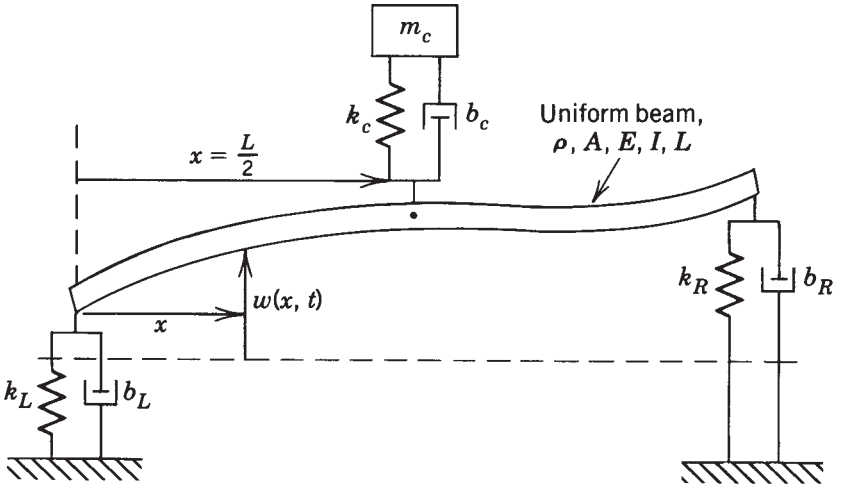
PROBLEMS

10-1. Using the simple lumping technique of Section 10.1, construct as low-order a model as possible for the bars shown in the figure such that integral causality exists for all elements.



10-2. For the systems of Problem 10-1, assign causality and derive state equations. Write down, using symbols from Problem 10-1, reasonable parameters for each system.

10-3. A uniform Bernoulli–Euler beam is suspended at its two ends and has a dynamic system attached at its center. Construct a low-order, finite-lump model using lumps similar to those in Figure 10.8.



10-4. For the system of Problem 10-3, derive a state representation.

10-5. The transverse motion of a tightly stretched string is governed by the simple wave equation, identically to the bar in longitudinal motion,

$$T\frac{\partial^2 \omega}{\partial x^2} + F\delta(x - x_1) = \rho\frac{\partial^2 w}{\partial t^2}.$$

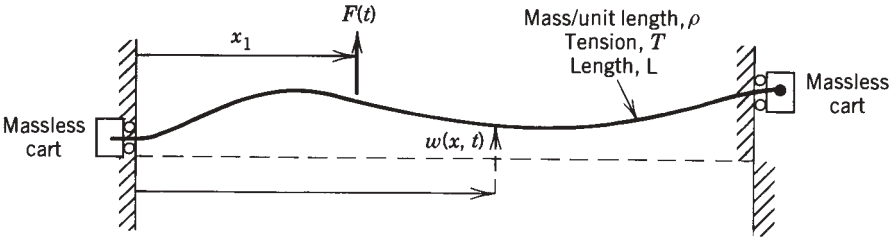
In the figure the boundaries are force free because the attachment points are massless. The boundary conditions are

$$\frac{\partial \omega}{\partial x}(0, t) = \frac{\partial w}{\partial x}(L, t) = 0.$$

Assume a separated solution

$$\omega(x, t) = Y(x) f(t),$$

and derive the frequency equation and mode shapes, including the rigid-body mode.

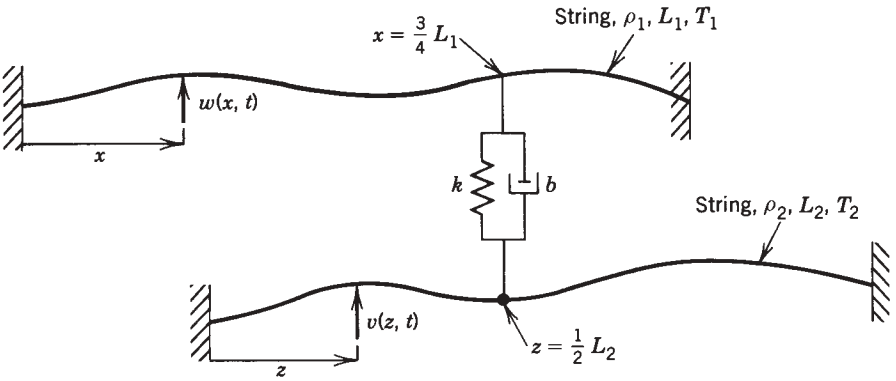


10-6. For Problem 10-5, assume the forced response is given by

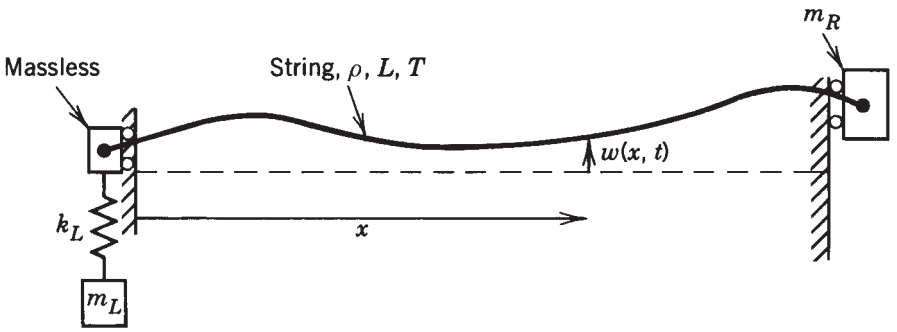
$$w(x, t) = \sum_{n=0}^{\infty} Y_n(x) \eta_n(t),$$

and derive expressions for the modal masses and stiffnesses. Convince yourself that a bond graph structure similar to Figure 10.17 (only including the rigid-body mode) is appropriate for this system.

10-7. Two tightly stretched strings are connected as shown in the figure. Construct a finite-mode model for this system including two modes for each string. Since the string ends are fixed, use the mode shapes for fixed ends, and use the symbol $Y_n(x)$ to represent these modes.



10-8. A tightly stretched string has dynamic elements on its boundaries as shown in the figure. Construct a bond graph for this system including two dynamic modes for the string (plus the rigid-body mode, of course). Assign causality, and convince yourself that the derivative causality exists. Use symbols $Y_n(x)$ for the mode shapes and modal parameters.



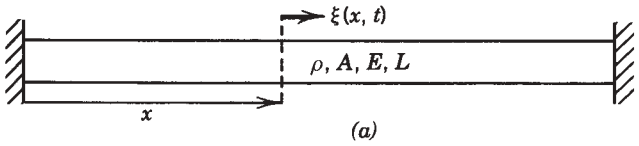
10-9. For the system of Problem 10-8, add an additional modal compliance, and show that no derivative causality exists. Derive the state equations.

10-10. The figure shows a bar, fixed at its left end, with three different right-end boundary conditions. In part a of the figure, the right end is fixed, in (b) it is attached to a spring, and in (c) it is attached to a velocity source. For reference, the mode shapes and frequencies for the fixed–fixed bar are

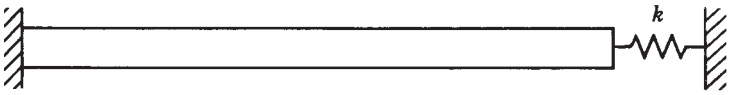
$$\omega_n^2 = \frac{E}{\rho}\left(\frac{n\pi}{L}\right)^2, \quad Y_n(x) = \sin n\pi\frac{x}{L}, \qquad n = 1, 2, \ldots,$$

and the mode shapes and frequencies for a bar fixed at $x = 0$ and free at $x = L$ are

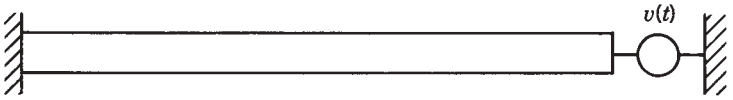
$$\omega_n^2 = \frac{E}{\rho}\left((2n-1)\frac{\pi}{2L}\right)^2, \quad Y_n(x) = \sin(2n-1)\frac{\pi}{2}\frac{x}{L}, \quad n = 1, 2, \ldots.$$



(a)



(b)



(c)

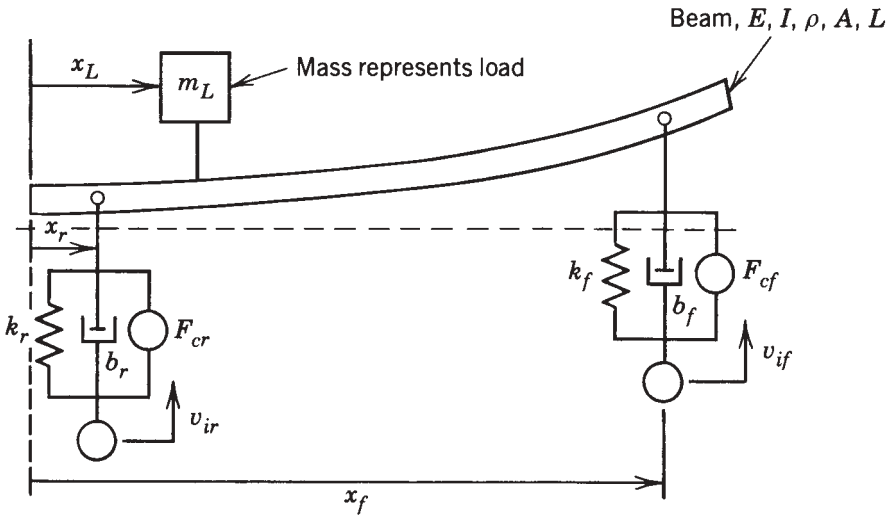
(a) Construct a two-mode bond graph model for the bar in part b using the fixed–free mode shapes and properly attaching the spring k. Assign causality, derive equations, and derive the characteristic equation. Let $k \to \infty$, effectively fixing the right end, and determine the natural frequencies of the constrained system. Compare with the frequencies for a fixed–fixed bar.

(b) Construct a two-mode model for the bar in part c, again using the fixed–free modes. Assign causality, and notice the derivative causality caused by the velocity input $v(t)$. Let $v(t) = 0$, effectively fixing the right end; deal with the derivative causality algebraically; and derive the characteristic equation for this system. Compare the calculated frequency with the frequency for the fixed–fixed bar.

(c) Finally, rid yourself of derivative causality by appending an additional modal compliance. Again derive the characteristic equation, and compare the calculated frequencies with those for a fixed–fixed bar.

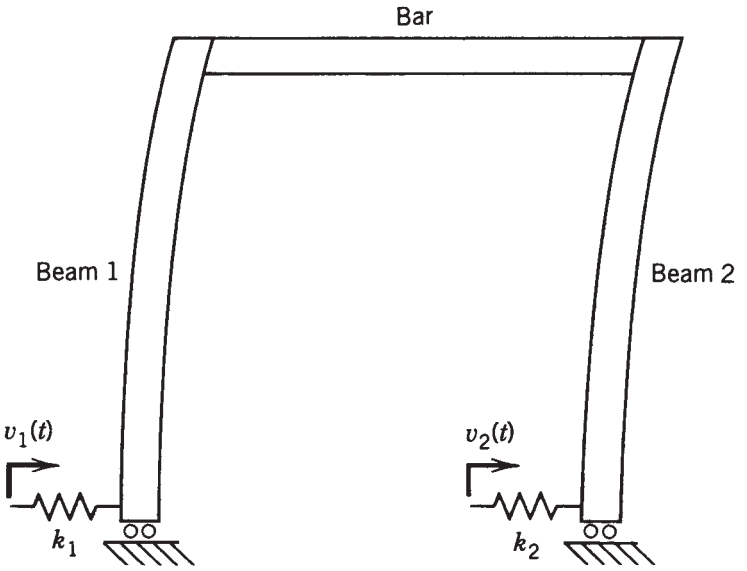
The reader should be impressed at how well the force-free modes converge to fixed-boundary modes even when only two modes are used.

10-11. A truck frame is modeled as a uniform beam in the figure. We would like to design a control system, using actuators at the suspension locations, that can provide damping for the bending motion of the frame. Construct a one-mode model (two rigid-body modes plus one dynamic mode) for the frame, and properly attach the suspension elements. Derive a set of state equations from which we can launch our control studies. If derivative causality exists, modify your model to fix it, and then derive equations.



10-12. The figure shows a simple structure composed of two vertical beams and a horizontal bar. Construct a finite-mode model of this system, assuming

the beams execute only transverse motion and the bar only longitudinal motion. Use symbols for modes and modal parameters, but state what modes you are using. Predict any formulation problems, and tell how you would modify the model to avoid these problems.



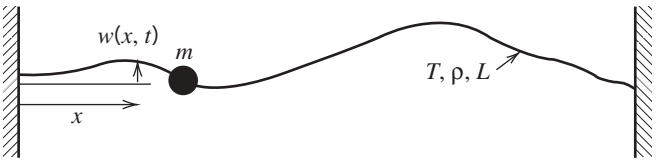
10-13. A tightly stretched string is fixed at both ends, as in Problem 10-7. Perform a modal analysis of a tightly stretched string and convince yourself that the mode shapes and frequencies are

$$\omega_n^2 = \sqrt{\frac{T}{\rho}}\frac{n\pi}{L}, \quad n = 1, 2, \ldots,$$

$$Y_n(x) = \sin n\pi \frac{x}{L},$$

where T is the string tension and ρ is the mass per unit length. Derive expressions for the modal masses, m_n, and modal stiffnesses, k_n.

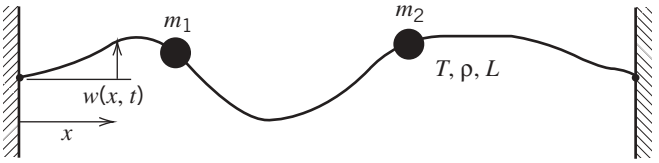
10-14. The tightly stretched string from Problem 10-13 now has a mass, m, attached at $x = \frac{1}{4}L$:
Treat the string lengths $\frac{1}{4}L$ to the left and $\frac{3}{4}L$ to the right of the mass as nonmodal tension springs and derive an expression for the natural frequency of the system.



10-15. Model the system in Problem 10-14 using the finite-mode bond graph. Include the first two modes for the string and, of course, attach the mass. Assign causality, and you will discover that derivative causality exists. Use formulation methods from Chapter 5 and derive the state equations. Perform a linear analysis and derive an expression for the system natural frequencies. Compare this result to that from Problem 10-14.

10-16. Add an additional modal compliance to the model in Problem 10-15 and show that no derivative causality exists. Derive state equations.

10-17. The tightly stretched string in Problem 10-13 now has two masses, m_1 and m_2, attached at positions $x = \frac{1}{4}L$ and $x = \frac{3}{4}L$. Formulate a two-mode bond graph model. Assign causality starting with the modal masses and show that both attached masses are in derivative causality. Append one additional modal compliance and show that one attached mass is in derivative causality. Append an additional modal compliance and show that all integral causality exists but there are unassigned bonds indicating an algebraic loop. Perform the algebra using procedures from Chapter 5 and derive state equations.



11

MAGNETIC CIRCUITS AND DEVICES

Many useful electrical and electromechanical devices contain magnetic circuits. Multiport models for some of these devices have already been studied in previous chapters, but here the magnetic flux paths will be modeled in detail. A detailed model is required, of course, if one is to design a motor, solenoid, transformer, or the like, although a system analyst may be content with an overall multiport model that adequately predicts only the external port behavior.

11.1 MAGNETIC EFFORT AND FLOW VARIABLES

The toroidal coil shown in Figure 11.1 can serve as an ideal configuration for defining magnetic variables and establishing bond graph representations for magnetic circuits. If the toroid is made of soft iron, we expect the device to behave at the single electrical port as an inductance with a linear relation between current i and flux linkage λ, at least for moderate values of i. This model shows no physical effects inside the coil, however.

Inside the iron core, magnetic *flux* is induced whenever current flows through the coil. Let the total flux be called φ; it is measured in the SI system in *webers*, Wb. (In this unit system one weber is equivalent to one volt-second. There are several units for magnetic variables in use, but for simplicity, only the units appropriate to the SI system will be presented here.) The magnetic field may be described by the *magnetic flux density vector* **B** with the units of *teslas*, T, or *webers per square meter,* Wb/m^2. The length of **B** corresponds to the amount of flux passing through an area element perpendicular to the flux lines, and the direction of **B** is along the flux lines. The **B**-vector points in the direction the north pole of a compass needle would point if it were free to line itself up with the magnetic field. The **B**-vector is also often called the *magnetic induction*.

Copyright © 2012 John Wiley & Sons, Inc.

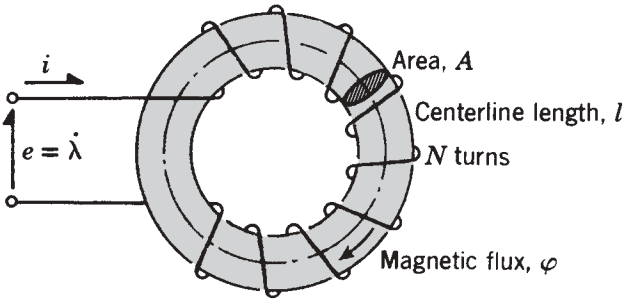


FIGURE 11.1. Toroidal coil.

In Figure 11.1, if we can assume that each of the N turns of conductor around the toroidal core links all the flux φ in the toroid, then the flux linkage λ is related to φ by

$$\lambda = N\varphi. \tag{11.1}$$

In practice, when many layers of wire are wound around on a core, some turns of wire do not link all of the flux lines. In this case, sometimes described by saying that some flux "leaks" out of the coil, Eq. (11.1) can be still valid except that N is a nondimensional number representing an *effective number of turns*, rather than an actual number of turns. In what follows, we shall refer to N as a number of turns without necessarily including the adjective *effective*.

By differentiating Eq. (11.1) with respect to time, the relation between the port voltage e and the magnetic variable $\dot{\varphi}$ is found:

$$e = \dot{\lambda} = N\dot{\varphi}, \tag{11.2}$$

which is just Faraday's law applied to the coil.

The driving force that tends to set up φ in the core is the *magnetomotive force*, M, which is proportional both to the number of turns of conductor and to the current flowing in the coil:

$$M = Ni. \tag{11.3}$$

The magnetomotive force, which is analogous to the electromotive force, has the units of current, amperes (although it is conventionally given in ampere-turns, despite the fact that the number of turns is really dimensionless). The magnetomotive force is often abbreviated as MMF. In the ideal, dissipationless case, M and φ for the coil are related by a nonlinear or linear relation. This relation is sketched in Figure 11.2.

In Figure 11.2*a*, a typical relation for a ferromagnetic material core is sketched. A *soft* ferromagnetic material is one that is easily magnetized and demagnetized by means of a current-carrying coil, that is, a material for which M and φ are related by a single valued curve so that $\varphi = 0$ when $M = 0$. *Hard* ferromagnetic materials, used for permanent magnets, show a hysteresis loop when M is cycled,

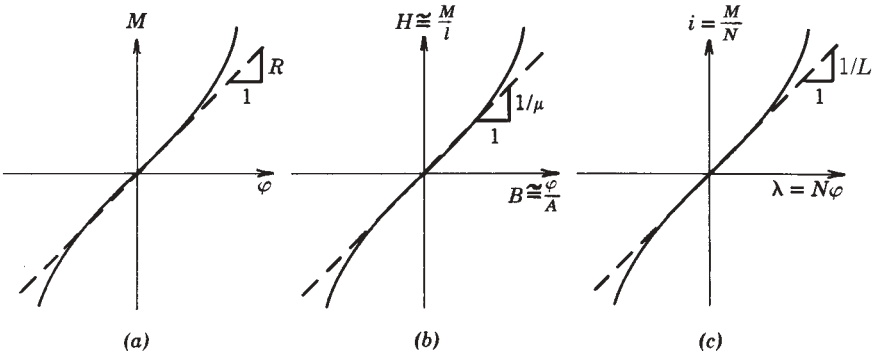


FIGURE 11.2. Core properties: (*a*) magnetomotive force versus flux; (*b*) field strength versus flux density; (*c*) current versus flux linkage.

and φ can remain high even when M is zero. For now, we shall concentrate on soft ferromagnetic materials. Such materials exhibit a saturation effect as shown. The increase in φ as M is increased slows down, and, practically, there is a limiting value of flux that the core can attain when M is very large. (Most steels saturate at B-values of less than 2 T.)

In order to characterize a core material, it is convenient to exhibit a curve independent of any particular core configuration. For this purpose the so-called $B-H$ curve is frequently used. In Figure 11.2*b* the magnitude of the **B**-vector, B, is shown as φ divided by the core cross-sectional area A. This way of computing B would be exact if the flux were uniformly distributed throughout the core. The *magnetizing force*, or *field strength* H, is the magnetomotive force per unit length, with units of ampere-turns per meter. For the toroidal coil, H is approximately M divided by the centerline length of the coil, l. When only a single $B-H$ curve is given, it is implied that the material is *isotropic*, that is, that the relation between B and H is the same no matter how the field is oriented in the material. For certain crystal structures, this is not the case, and several different $B-H$ curves for different orientations of **B** can be found.

In Figure 11.2*c*, the port characteristics of the inductor are sketched by modifying the $M-\varphi$ or $B-H$ curve using N. We see that the saturation effect corresponds to a nonlinear region of the —I element representing the inductance. Since most $B-H$ curves exhibit saturation, electrical chokes are really nonlinear, though in normal operation of these devices a linear inductance model is often sufficiently accurate.

It is of interest to compute the inductance for the coil. Assuming that the curves of Figure 11.2 can be approximately represented by a straight line near the origin, L may be found as follows:

$$L = \frac{\lambda}{i} = \frac{N\varphi}{M/N} = \frac{N^2AB}{lH}. \tag{11.4}$$

The initial slope of the $B-H$ curve is given the symbol μ and called the *permeability*; it has units of tesla-meters per ampere

$$\mu = \frac{B}{H}, \tag{11.5}$$

so that $L = N^2 A\mu/l$. (Free space itself has a permeability, μ_0, and sometimes μ is expressed as a relative permeability, that is, the value of B/H is given by $\mu\mu_0$ rather than just μ.)

For the entire coil, one may express the slope of the $M - \varphi$ curve in the linear region in two ways. The *permeance*, **P**, is given by the following formula:

$$\mathbf{P} = \frac{\varphi}{M} = \frac{\mu A}{l}. \tag{11.6}$$

The *reluctance*, **R**, is given by

$$\mathbf{R} = \frac{1}{\mathbf{P}} = \frac{M}{\varphi} = \frac{l}{\mu A}. \tag{11.7}$$

As a mnemonic device, one may say that a long, small-area flux path in a material of small permeability is reluctant to permit the establishment of much flux. The parameters L, **R**, and μ are shown as slopes in the sketches in Figure 11.2. Table 11.1 summarizes the magnetic variables and parameters discussed, together with their SI units and some remarks concerning typical values.

In order to proceed further with the analysis of magnetic circuits and devices, it is useful to begin classifying variables such as M and φ and parameters such as **P** and **R**. Historically, the reluctance **R** was sometimes thought of as analogous to electrical resistance, so that flux was thought of as analogous to current and the magnetomotive force as analogous to the electromotive force. Although this analogy can be used, it is not satisfying as a basis for a bond graph, since an electrical resistor dissipates power while a coil exhibiting reluctance stores energy. In fact, the reluctance and the permeance are linear parameters of a —C or —I element. From a bond graph point of view it seems reasonable that M, the magnetomotive force, should be an effort quantity, but the flow should not be flux, as in the more traditional analogy, but rather the time rate of change of flux, $\dot{\varphi}$.

The deficiencies of the reluctance–resistance analogy are apparent when one attempts to study a dynamic system containing electrical and mechanical elements as well as a magnetic current (see Reference [1]). It was only after the gyrator had been accepted as a useful network element, however, that the analogy most natural for bond graphs could be proposed [2].

Suppose we return to the basic equations, (11.2) and (11.3), and identify $\dot{\varphi}$ as a flow variable and M as an effort variable. Then N is clearly a gyrator parameter. In addition, since φ is a displacement variable, then Eq. (11.6) shows that **P** is a capacitance parameter and **R**, the reluctance, is the inverse capacitance or stiffness parameter. All the relations are elegantly summed up in the bond graph of Figure 11.3. Of course, the combination $\rightharpoonup GY \rightharpoonup C$ does behave like $\rightharpoonup I$ at

TABLE 11.1. Summary of Magnetic Variables and Parameters

Symbol	Name	SI Unit	Equivalent Units, Remarks
φ	Total magnetic flux	Weber (Wb)	Volt-second (V-s)
$\dot{\varphi}$	Flux rate	Wb/s	V
B	Magnetic flux density, magnetic induction	Tesla (T)	Weber per square meter, (Wb/m^2); magnitude of a vector quantity
M	Magnetomotive force, MMF	Ampere (-turn) (A)	Number of turns of a conductor is dimensionless, but is included if $M = Ni$ is used, where N = turns of wire carrying current i
H	Magnetic field strength	Ampere (-turn) per meter (A/m)	Magnitude of a vector quantity
μ_0	Permeability of free space, $4\pi \times 10^{-7} \cong 1/800{,}000$	Henry per meter (H/m)	Tesla-meter per ampere (T-m/A); relates B to H in air: $B = \mu_0 H$ or $H \cong 800{,}000B$
μ	Permeability of material	H/m	T-m/A; relates B to H in isotropic material: $B = \mu H$
μ_r	Relative permeability of material	Dimensionless	$\mu = \mu_r \mu_0$; for ferromagnetic materials, μ_r varies from about 400 to 400,000 for such materials, $2B < H < 2000B$
P	Permeance of circuit element	Henry (H)	Weber per ampere (Wb/A), used in formula $\varphi = \mathbf{P}M$
R	Reluctance of circuit element	H^{-1}	A/Wb; $\mathbf{R} = 1/\mathbf{P}$ used in formula $M = \mathbf{R}\varphi$
E	Energy	Joule (J)	$\mathbf{E} = \int M d\varphi$(A-Wb = A-V-s)

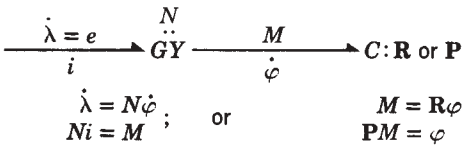


FIGURE 11.3. Bond graph for coil of Figure 11.1.

the external port, as it must. The gyrator is necessary if we wish to consider both electromotive force and magnetomotive force as effort variables. (This seems to be the way classical physicists thought of these variables.)

The classification of variables that will be used subsequently is listed in Table 11.2 for convenience. Note that the capacitance parameter only is to be

TABLE 11.2. Mechanical, Electrical, and Magnetic Bond Graph Variable and Parameters

General	Mechanical	Electrical	Magnetic
Effort	Force, F	Electromotive force, e	Magnetomotive force, M
Flow	Velocity, V	Current, i	Flux rate, $\dot{\varphi}$
Displacement	Distance, X	Charge, q	Flux, φ
Capacitance parameter	Compliance, C	Capacitance, C	Permeance, $\mathbf{P}$
Stiffness parameter	Stiffness, k	$1/C$	Reluctance, $\mathbf{R} = 1/\mathbf{P}$

applied in the linear case. The magnetic circuit generally involves nonlinear C-elements with characteristics as sketched in Figure 11.2.

11.2 MAGNETIC ENERGY STORAGE AND LOSS

Figure 11.2 showed several ways to exhibit the properties of a nonlinear but lossless magnetic core. In this section the energy storage and loss will be investigated in more detail.

Figure 11.4 shows two small pieces of material of physical volume Al. On the left, a uniform stress σ is associated with a force F and causes an extension δ and strain ε in the isotropic material. The total energy stored would be

$$\mathbf{E} = \int F d\delta = \int \sigma Al\, d\varepsilon = Al \int \sigma d\varepsilon. \tag{11.8}$$

Thus, the energy per unit volume is the area shown shaded in the $\sigma - \varepsilon$ diagram when the material is deformed from a state of zero stress and strain. When the material is in the linear elastic range, $\sigma = E\varepsilon$ and the energy is $E\varepsilon^2/2$, but if σ versus ε is *any* single-valued curve, the material is obviously energy conservative.

On the other hand, if the material yields, the stress and strain follow a hysteresis loop such as the one shown dashed in the sketch, and the area enclosed by the loop represents the energy loss per cycle per unit volume. Mechanical elements that remain elastic are represented by C-elements, but when yielding occurs, some kind of R-element must be added to model the energy loss.

The material on the right of Figure 11.4 behaves under the action of a magnetic field in a manner similar to the material on the left under the action of a mechanical stress field. Assuming that the flux φ has a uniform density B and that the material is isotropic, the total stored energy would be computed according to the classifications of variables in Tables 11.1 and 11.2 as follows:

$$E = \int M d\varphi = \int lHA dB = Al \int H dB. \tag{11.9}$$

Thus, the magnetic energy per unit volume is represented by the hatched area in the $B-H$ plot on the right in Figure 11.4. In the linear range, $H = B/\mu$ and

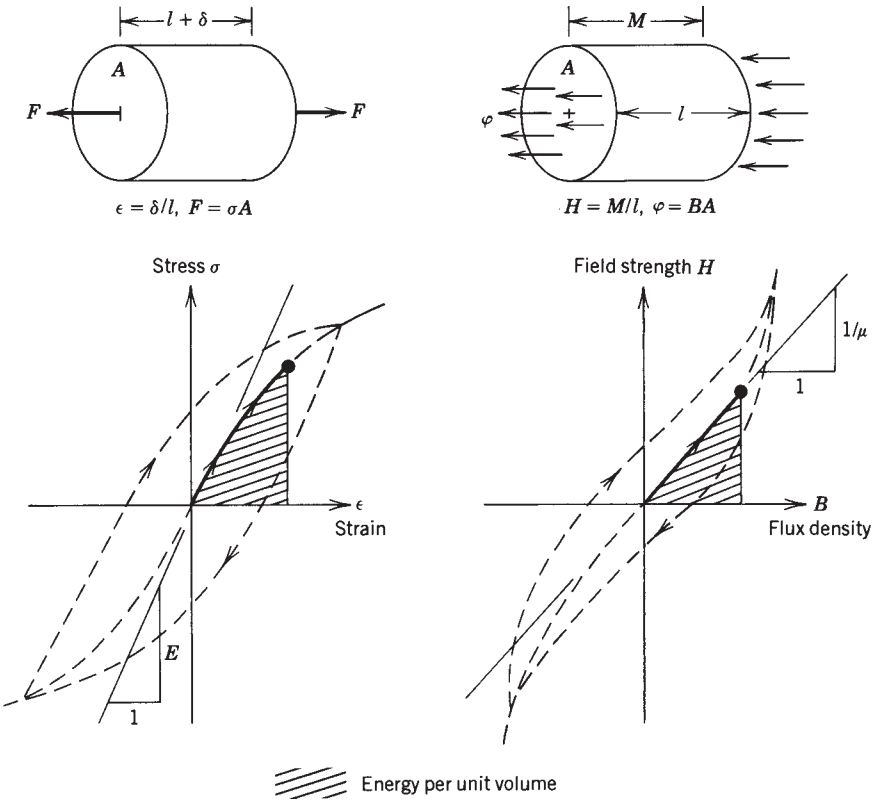


FIGURE 11.4. Mechanical and magnetic energy densities.

the energy is $B^2/2\mu$, but if B and H lie on *any* single-valued curve, energy is conserved.

When B and H are cycled, a hysteresis loop indicated by the dashed loop forms, and the area inside the loop represents the energy loss per cycle per unit volume. For soft magnetic materials, the loops are quite narrow and a C-element with a curved constitutive law can be used to model the saturation effect without loss. The hysteresis effect requires the addition of R-elements, since energy loss is involved.

The detailed modeling of hysteresis is generally quite complicated, since it involves a sort of memory effect. One way to handle this is to replace the single C-element, which has a single displacement state variable, with a combination of C- and R-elements, which then serves to remember aspects of the loading of the hysteretic system through the values of a number of hidden state variables. The procedure is discussed in Reference [3].

Permanent magnets are made of hard magnetic materials with very large hysteresis loops. These materials resemble elastic–plastic mechanical materials

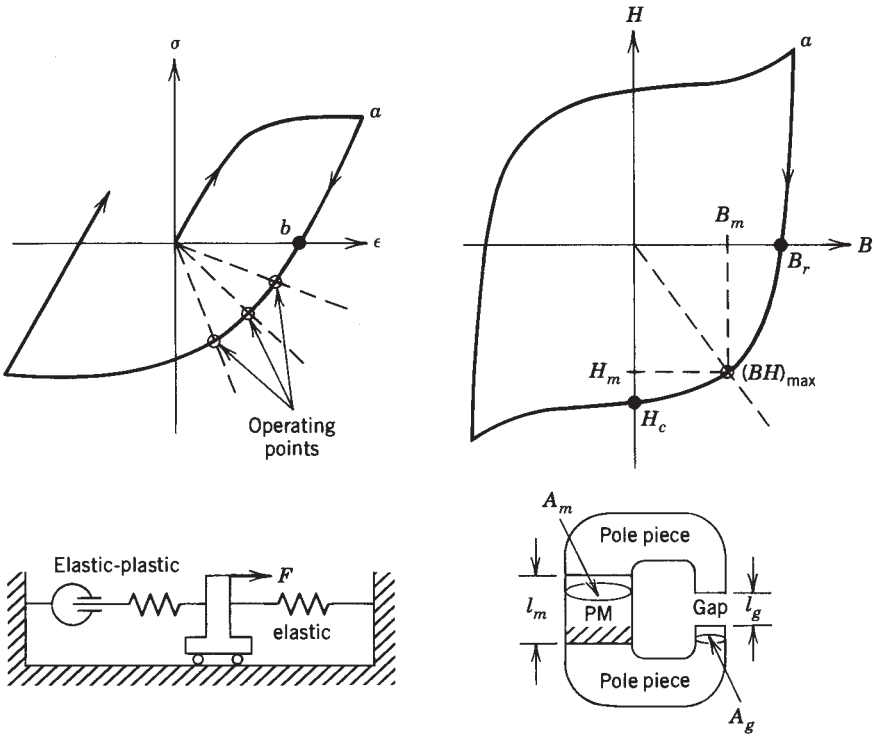


FIGURE 11.5. Hysteresis loops for an elastic–plastic material and a permanent-magnet material.

which can retain a large strain after having been yielded. A permanent magnet retains a large **B**-field after magnetization even in the presence of an **H** or magnetomotive force in a direction that would tend to demagnetize the material.

Figure 11.5 shows the stress–strain plot of an elastic–plastic material and the B–H plot for a typical permanent magnet. In most plots the B and H axes are interchanged from the way they are presented here. When this is done, the hysteresis loops are traversed in the opposite direction and the areas representing stored energy change places. The unconventional H–B plot is required to keep the displacement variable φ or B on the horizontal axis and the effort variable M or H on the vertical axis, as has been done consistently for all other energy domains in this book.

On the left of Figure 11.5 is a mechanical system which will help explain how permanent magnets can be used to produce flux in air gaps for use in motors, electrodynamic loudspeakers, and other electromechanical transducers instead of through the use of current-carrying coils. In the stress–strain diagram, if the material is strained beyond its yield point to point a, when the stress is relaxed, the material returns to point b with a permanent strain. A similar phenomenon

occurs for the magnetic material represented on the right. After being magnetized to point a on the H–B hysteresis loop, the material would retain a B-value of B_r, the *remanence* or *remanent field*, if H were returned to zero. To achieve $B = 0$, one would have to apply a negative field strength of magnitude H_c, the *coercivity*. The parameters B_r and H_c are often used to describe the properties of permanent magnets.

Now suppose that some elastic–plastic material is connected in mechanical series to some elastic material as shown schematically in the lower left of Figure 11.5. If the force F is used to stretch the elastic–plastic material beyond its yield point and then removed, the system will not return to point b on the stress–strain diagram, because the elastic element will tend to recompress the yielded element. The stress will be negative, and the strain will be less than the strain at point b. Several possible operating points are shown, depending on how stiff the elastic element happens to be.

An analogous situation arises when a permanent magnet is used in a magnetic circuit with an air gap as shown in the lower right of Figure 11.5. If we assume for simplicity that the pole pieces have negligible reluctance and we neglect any leakage flux, then all the flux from the magnet will also traverse the gap, and the total MMF from the magnet will appear across the gap. The magnet will have to work against a negative H caused by the MMF at the gap necessary to force the flux through the gap. We now investigate a means for choosing optimum magnet and gap dimensions.

A common principle is that a given magnet should be used so that the total energy in the air gap is maximum. The MMF–flux relation in the air gap is linear,

$$M_g = R_g \varphi_g, \tag{11.10}$$

so the total energy is

$$E = R_g \varphi_g^2 / 2 = M_g \varphi_g / 2, \tag{11.11}$$

and we wish to maximize $M_g \phi_g$. The assumptions mean that the magnet flux is equal to the gap flux and the magnet MMF is just the negative of the gap MMF:

$$\varphi_m = \varphi_g, \qquad M_m = -M_g. \tag{11.12}$$

Then, assuming a uniform flux density and field strength and using the magnet length l_m and area A_m, we can relate $M_g \varphi$ to the B and H values in the magnet:

$$|M_g \varphi_g| = |M_m \varphi_m| = (A_m B)(l_m H) = (A_m l_m)(BH), \tag{11.13}$$

where only the magnitude of H is considered, although H is actually negative. The point is that the gap energy is maximized by operating the magnet at that point in the B–H plane where the product BH is maximum. Let us now define the point B_m, H_m on the hysteresis curve to be the point at which BH is maximum,

$(BH)_{\max}$. We then dimension the gap in such a way that the magnet operates at B_m, H_m.

Three simple equations apply: (1) the magnet and gap fluxes are equal,

$$B_m A_m = B_g A_g; \tag{11.14}$$

(2) the magnitudes of the magnet and gap MMFs are equal,

$$H_m l_m = H_g l_g; \tag{11.15}$$

and (3) in the air gap

$$B_g = \mu_0 H_g. \tag{11.16}$$

Elimination of B_g and H_g yields the final result:

$$\frac{A_g l_m}{A_m l_g} = \frac{B_m}{\mu_0 H_m}. \tag{11.17}$$

The term on the right in Eq. (11.17) can be thought of as a normalized slope of the line passing through the $(BH)_{\max}$ point. Its value is often tabulated for magnet materials. The value of $(BH)_{\max}$ is a measure of the strength of the magnet material.

Note that by adjusting the magnet and gap areas, the B-value in the gap can be greater than in the magnet. If $A_g = A_m$, then the ratio of l_m to l_g is just $B_m/\mu_0 H_m$. For older metallic magnets this ratio was usually greater than unity, leading to designs with magnets much longer than the air gap length, but recent high-strength rare earth magnets have a ratio of about unity, which means that the magnet length is about the length of the air gap unless $B > B_m$ is required.

For more details on the design of magnetic circuits to produce constant fields either with permanent magnets or with current-carrying coils see References [4] and [5]. We now proceed to model circuits in which fields may vary dynamically.

11.3 MAGNETIC CIRCUIT ELEMENTS

Using the effort–flow identifications of Table 11.2, it is possible to begin a study of lumped-parameter elements for magnetic circuits. In Figure 11.6*a*, a coil around a length of core is shown contributing an increase in MMF. In this model, the length of core has no reluctance. In Figure 11.6*b*, an MMF drop is shown related to the flux φ. When two flux paths are joined as shown in Figure 11.6*c*, the result can be modeled with a 0-junction, since the three flux rates sum to zero and there is a single MMF. Finally, in Figure 11.6*d*, an air gap is modeled essentially in the same way as a length of core material. The air gap has high reluctance for its length or low permeability (essentially the permeability of free space, μ_0) and does not exhibit saturation as iron cores do. In computing reluctance or

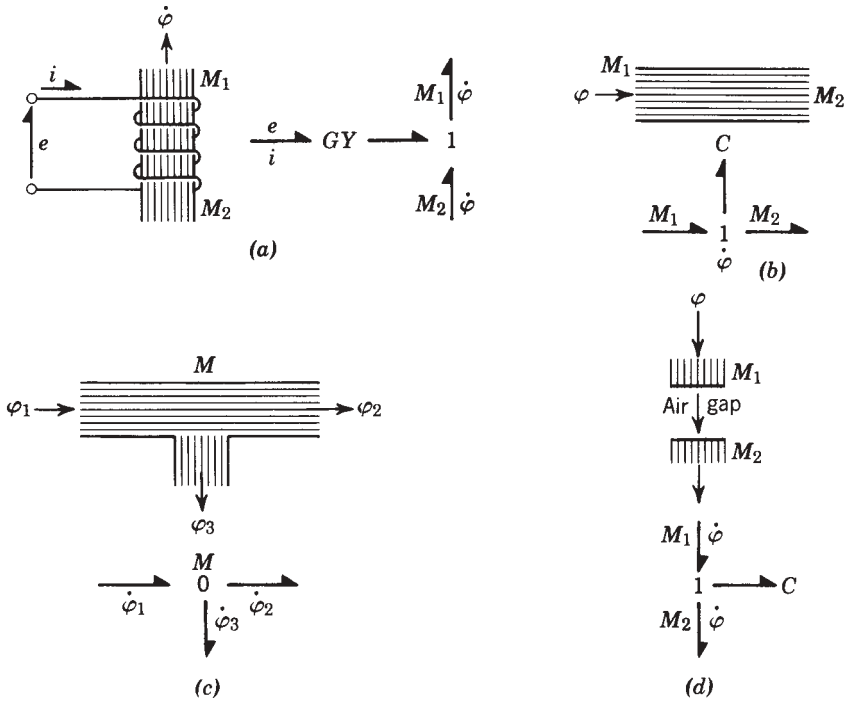


FIGURE 11.6. Bond graphs for magnetic circuit elements: (*a*) driver coil; (*b*) piece of ferromagnetic core; (*c*) core junction; (*d*) air gap.

permeance parameters for —*C* elements of flux paths, we use Eqs. (11.6) and (11.7), in which areas and lengths for the flux path elements must be estimated judiciously, with consideration given to the actual paths of the flux lines.

Using the elements of Figure 11.6, a bond graph for a simple magnetic circuit such as the one shown in Figure 11.7 may be constructed. The MMF can be treated just like voltage in an electric circuit. In Figure 11.7*b* the MMF variables have been each assigned to a 0-junction, and drops in MMF across —*C* elements and a rise due to the coil have been indicated using 1-junctions. Since only differences in MMF are significant ultimately, M_d is chosen as a zero MMF point, and the bond graph can be simplified. It may be helpful to consider the previously studied case of the analogous electric circuit shown in Figure 11.7*c*.

After the ground node has been eliminated and the 2-port 0-junctions removed, it is found that three —*C* elements are joined to a single 1-junction. This means that all *C*'s have the same flow variable and the effort drops add. In magnetic terms, this means that for the linear case, a single *C*-element with a reluctance equal to the sum of the three reluctances of the original *C*'s can be substituted as shown in Figure 11.7*d*. In the nonlinear case, an equivalent —*C* relation can be found by simply adding MMF drops for the three original *C*'s as the common flux is varied over the range of interest.

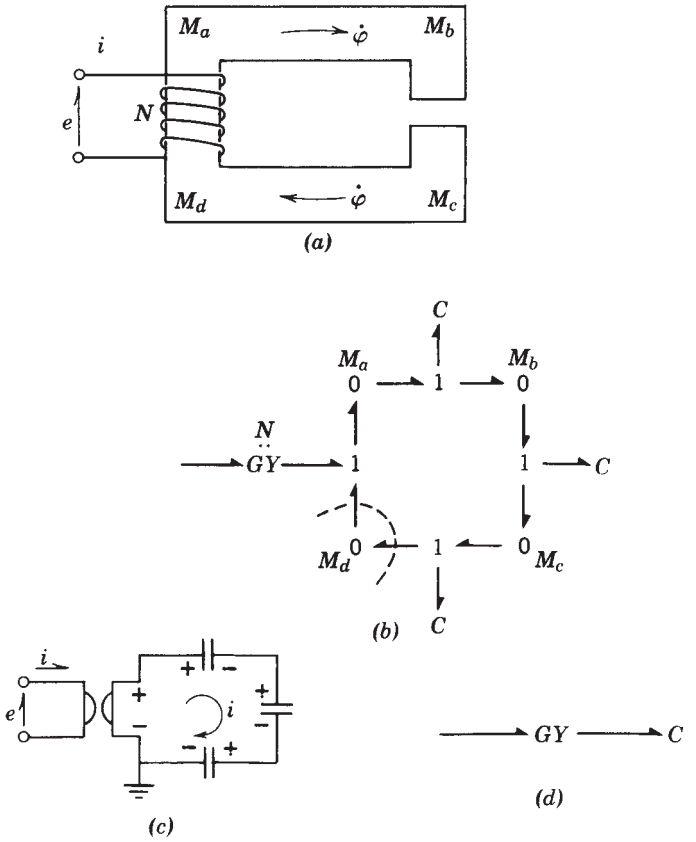


FIGURE 11.7. Bond graph model of magnetic circuit with air gap: (*a*) sketch of device; (*b*) bond graph; (*c*) analogous electric circuit; (*d*) simplified bond graph.

A more complicated magnetic circuit appears in Figure 11.8. The flux path associated with φ_3 could be physically present, or it could be a lumped representation of the path taken by leakage flux that bypasses the air gap. In fact, there always is some flux that escapes from the main path and spreads into the surrounding space. Although the real leakage paths are distributed in space, it is often sufficient to include a single equivalent leakage path and leakage flux. Since the leakage path reluctance is usually dominated by the high reluctance of the air portion, it is common to assume that the leakage path can be represented by a linear, high-reluctance —C. The basic bond graph of Figure 11.8*b* is simplified by the choice of a zero MMF point, and the bond graph of Figure 11.8*c* results. The loop may be eliminated using the bond graph identity discussed in Chapter 4 for the 0–1–0–1 ring, since the signs are proper for the reduction. The bond graphs of Figure 11.8*d* and *e* result. In the last form, it is clear that $C2$, $C4$, and $C6$ could be combined, as could $C1$ and $C5$. If these C's were linear, their

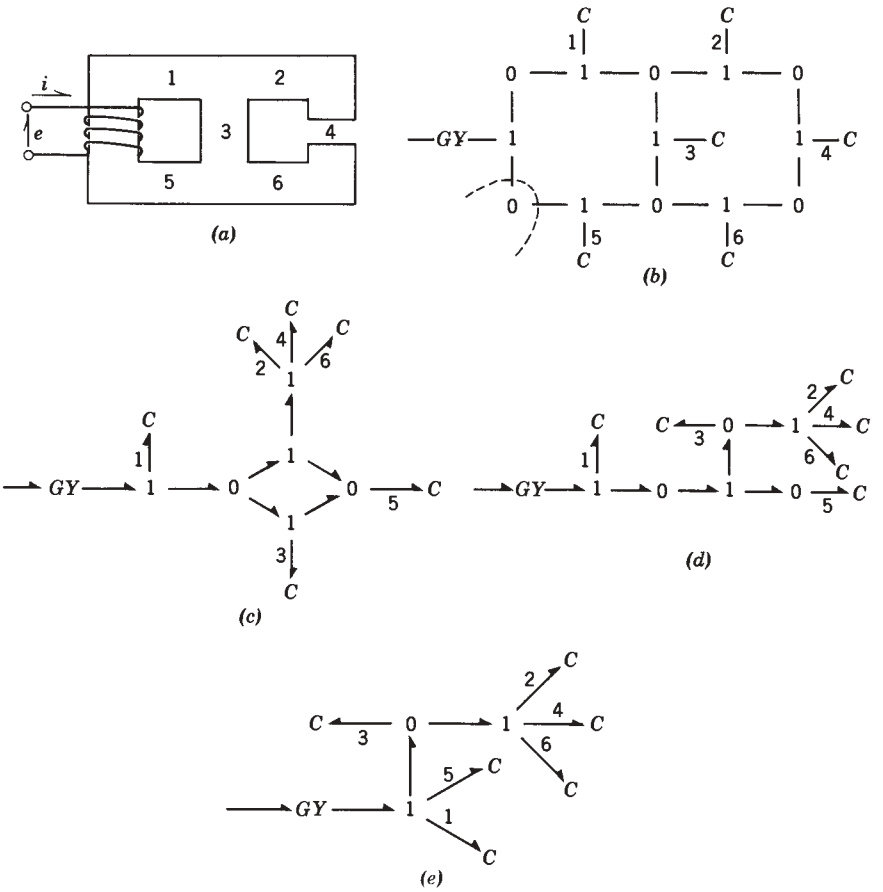


FIGURE 11.8. Circuit with extra flux path to model leakage flux: (*a*) sketch; (*b*) basic bond graph; (*c*–*e*) simplified bond graphs.

reluctances would be summed. Ultimately, a single equivalent C could be found for the entire collection of C's, 0's, and 1's, but then, of course, none of the internal MMFs or fluxes could be found.

When no loss mechanisms are included in magnetic circuits, then they appear internally as networks of Cs, and it is often convenient to reduce the system by finding equivalences among the subfields. At a more detailed level, however, there are energy losses associated with eddy currents in the core and other effects. Laminations in the core of a magnetic circuit help in reducing the loss associated with eddy currents, but accurate models require the insertion of loss elements. Analysis of a laminated core can show that the R–C transmission line model for the core may be used for the linear case [2]. A study of such detailed models would take too long for present purposes, and in practice a simpler approach

often suffices. It is often sufficient to append a magnetic resistance to the 1-junction of Figure 11.6*b*. This implies an extra MMF drop associated with a length of core that is a function of $\dot{\varphi}$, in addition to the drop in MMF that is associated with φ. As might be expected, it is not easy to predict the magnitude of the resistance in the absence of experimental evidence. On the other hand, particularly for a limited frequency range of excitation, it is often possible to adjust a linear resistance relation to provide a quite accurate representation of losses in the core.

11.4 MAGNETOMECHANICAL ELEMENTS

Before exhibiting magnetic circuits containing loss elements, it is important to show a method of handling magnetomechanical transducers. (The solenoid studied in Chapter 8 was treated in multiport fashion without analyzing the magnetic circuit in detail.) The basic idea behind many transducers is that mechanical motion can alter the relation between flux and MMF in a flux path. This type of transducer is often called a "variable-reluctance" transducer.

Consider the device of Figure 11.9, in which a force F is associated with a movable pole piece that communicates with a fixed pole through an air gap of variable length X. A flux φ passes through the gap and is associated with an MMF drop ΔM. The device can be represented by a C-field, as shown in Figure 11.9*b*, and is analogous to the movable-plate capacitor of Chapter 8.

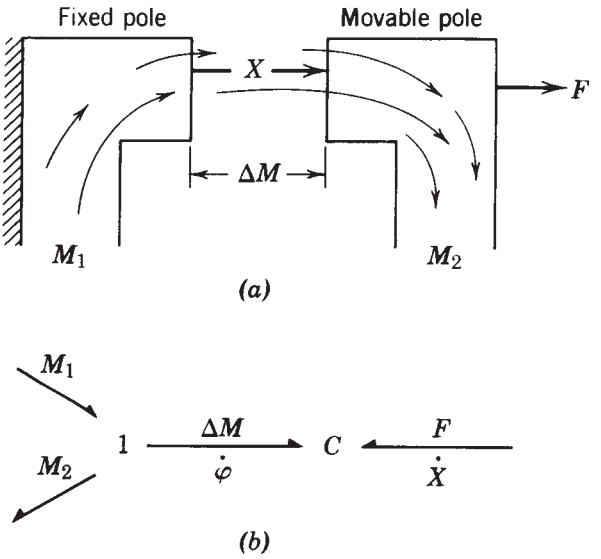


FIGURE 11.9. Magnetomechanical transduction: (*a*) sketch; (*b*) bond graph representation.

Since this particular device uses an air gap, it is reasonable to assume that the device is linear, but that the reluctance **R** is a function of *X;* that is,

$$\mathbf{R}(X)\varphi = \Delta M. \tag{11.18}$$

Using arguments similar to those in Chapters 7 and 8, we may write the energy **E** as

$$\mathbf{E}(X, \varphi) = \frac{1}{2}\mathbf{R}(X)\varphi^2. \tag{11.19}$$

Then

$$\Delta M = \frac{\partial E}{\partial \varphi} = R(X)\varphi \tag{11.20a}$$

and

$$F = \frac{\partial E}{\partial X} = \frac{1}{2}\frac{dR}{dX}\varphi^2. \tag{11.20b}$$

From Eq. (11.7), we might expect

$$\mathbf{R}(X) \cong \frac{X}{\mu_0 A}. \tag{11.21}$$

So

$$F \cong \frac{\varphi^2}{2\mu_0 A}, \tag{11.22}$$

indicating the F would be roughly constant for small changes in the value of X if φ were held constant. In most cases, however, φ is more nearly inversely proportional to **R** (when an MMF is held constant somewhere in the circuit), so F tends to vary as X^{-2} for small X. For very large values of X, the circuit model itself breaks down, and the flux does not follow the paths assumed. Any movement of core pieces will be associated with changes in stored energy and hence with forces or torques. The bond graph of Figure 11.9*b* may be used even when the magnetic relations are nonlinear due to saturation, but the computation of the energy is more complicated.

Figure 11.10 shows a ferromagnetic rotor that is subjected to an aligning torque τ when a field is induced in the stator. There are several ways to compute the torque, but perhaps the easiest involves energy considerations. We can again use Eq. (11.19) but with **R** a function of θ instead of X. Assuming that the flux is concentrated in the narrow gap of length l_g, Eq. (11.7) can be used with the area of the gap as a function of θ rather than the length. Letting the axial length of the rotor be l, the effective area of the two air gaps is $l(l_0 - r\theta)$, so the reluctance is

$$\mathbf{R} = \frac{2l_g}{\mu_0 l(l_0 - r\theta)}. \tag{11.23}$$

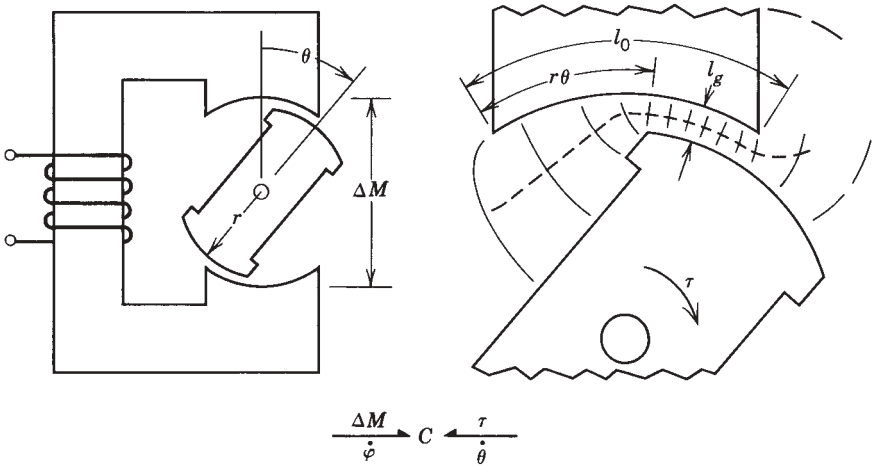


FIGURE 11.10. Rotary, variable-reluctance transducer.

Then

$$\mathbf{E}(\theta, \varphi) = \frac{l_g \varphi^2}{\mu_0 l (l_0 - r\theta)} \tag{11.24}$$

and

$$\tau = \frac{\partial E}{\partial \theta} = \frac{l_g r \varphi^2}{\mu_0 l (l_0 - r\theta)^2}, \tag{11.25}$$

$$\Delta M = \frac{\partial E}{\partial \varphi} = \frac{2 l_g \varphi}{\mu_0 l (l_0 - r\theta)}. \tag{11.26}$$

The last two relations describe the 2-port C-field in integral causality.

Other equivalent expressions may not be so useful for bond graph dynamic models. For example, defining

$$B = \frac{\varphi}{l(l_0 - r\theta)}, \tag{11.27}$$

Eq. (11.25) can be written

$$\tau = \frac{l r l_g B^2}{\mu_0}. \tag{11.28}$$

While correct, Eq. (11.28) may obscure the fact that B varies with θ and φ as the rotor turns and the excitation changes.

11.5 DEVICE MODELS

As an example of the utility of the bond graph models for magnetic circuits developed above, consider the relay of Figure 11.11. Bond graphs are an aid

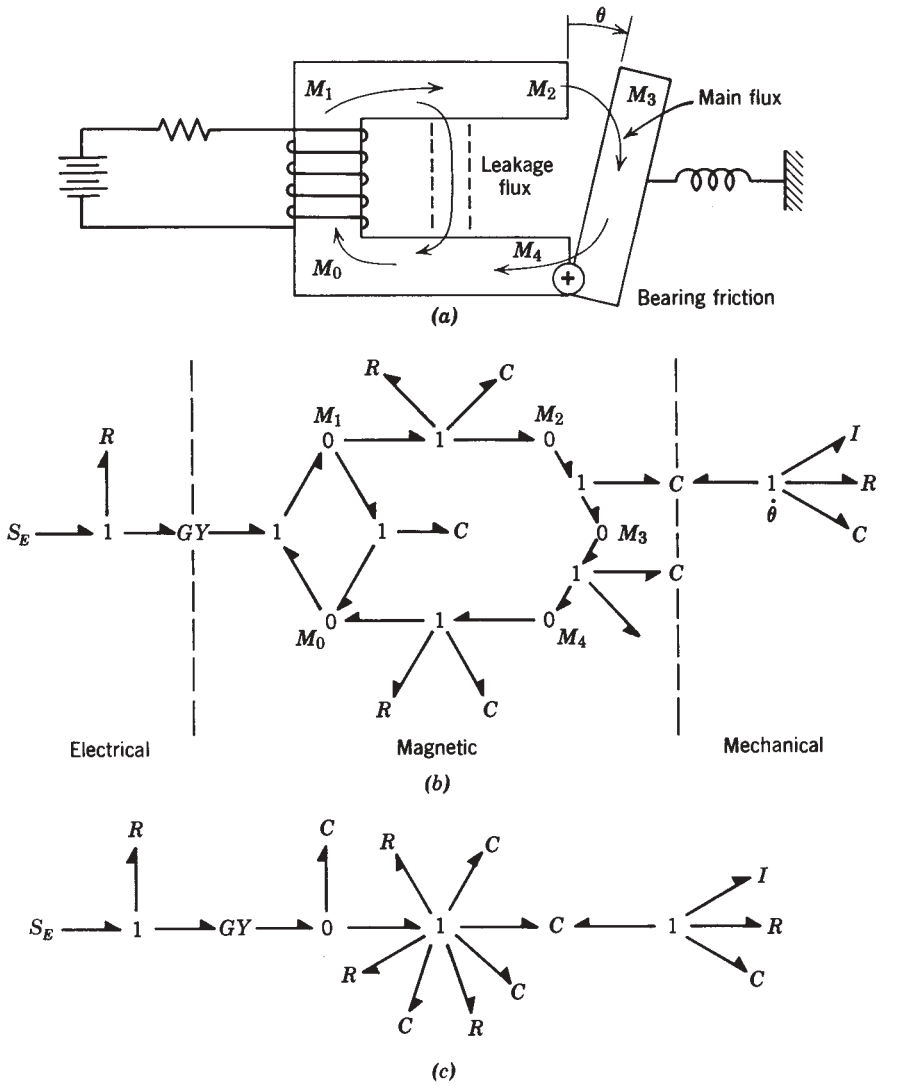


FIGURE 11.11. Relay: (*a*) schematic diagram; (*b*) basic bond graph; (*c*) simplified bond graph.

in the study of such systems, since three energy domains are involved. From the schematic diagram of the device, Figure 11.11*a*, one may begin assembling a bond graph model by indicating MMF values and then representing MMF drops with 1-junctions and C- and R-elements. At this point, some judgment is required, since it is not clear how many flux paths should be represented, nor is it clear where the eddy current losses will be most severe. When the bond graph is

simplified after choosing M_0 as the zero MMF, as in Figure 11.11*c*, it becomes clear that several C- and R-elements can be combined.

Discounting the loss elements, the transduction from electrical energy to mechanical energy is accomplished by a C-field addressed at one port through a gyrator or, in other words, by an IC-field transducer. This would have been predicted by the methods used in Chapter 8. However, the present approach using the detailed model of the magnetic circuit allows a designer to study the flux paths in detail and to model internal losses. From a designer's point of view, a single relay is a rather complex dynamic system. Conventional descriptions of models of such systems using equations are not particularly easy to follow. For an example of a real device modeled at about the level of complexity shown in Figure 11.11, see Reference [6]. The reader who takes the trouble to read this paper carefully may become convinced that the bond graph representation yields a compact and insightful means of displaying the physical assumptions used in creating a multiple-energy-domain model.

It is possible to create a very general bond graph model that describes a wide variety of electromagnetic–mechanical devices, including all the power-conserving and energy-storing devices discussed previously. The model is developed in Reference [7], and the devices are described in more detail in References [8–10].

We consider four devices and three forms of force or torque law, which will be incorporated into one general bond graph model. The derivation of the fundamental form of the force law appears in Reference [8] as applied to synchronous motors such as the salient pole version shown in Figure 11.12.

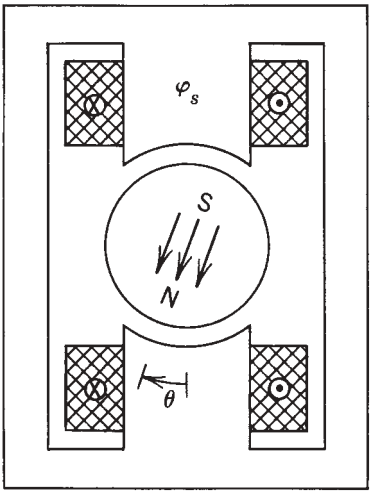


FIGURE 11.12. Synchronous motor with salient poles and permanent-magnet rotor discussed in Reference.

The flux linked by the stator windings is called φ_s in Reference [8] and is imagined split into a component due to the coil current i_s and a part due to the permanent magnet rotor position θ:

$$\varphi_s = L_s(\theta) i_s + \varphi_{sr}(\theta), \tag{11.29}$$

in which $L_s(\theta)$ is a position-dependent inductance. The model is therefore electrically linear, which means that saturation in the magnetic circuit is assumed not to occur. The magnetic energy $\mathbf{W}_m$ is also split into a part having to do with current in the stator windings and a part having to do with the permanent magnet position, $\mathbf{W}_{mr}$:

$$\mathbf{W}_m = \frac{1}{2} L_s(\theta) i_s^2 + W_{mr}(\theta). \tag{11.30}$$

An energy argument leads to the following expression for the electromagnetic torque T_e:

$$T_e = i_s \frac{d\varphi_{sr}}{d\theta} - \frac{dW_{mr}(\theta)}{d\theta} + \frac{1}{2} i_s^2 \frac{dL_s(\theta)}{d\theta}. \tag{11.31}$$

This expression, although derived for a fairly specific type of rotary device, actually can be interpreted as showing the correct form for the three basic types of forces present in electromagnetic actuators.

In Figure 11.13, two fundamental types of linear actuators discussed in Reference [9] are sketched. In previous bond graph literature, case *a* would be modeled either by an energy-conservative *IC*-field or by a *C*-field if the magnetic circuit variables were explicitly included, and case *b* would be modeled by a gyrator.

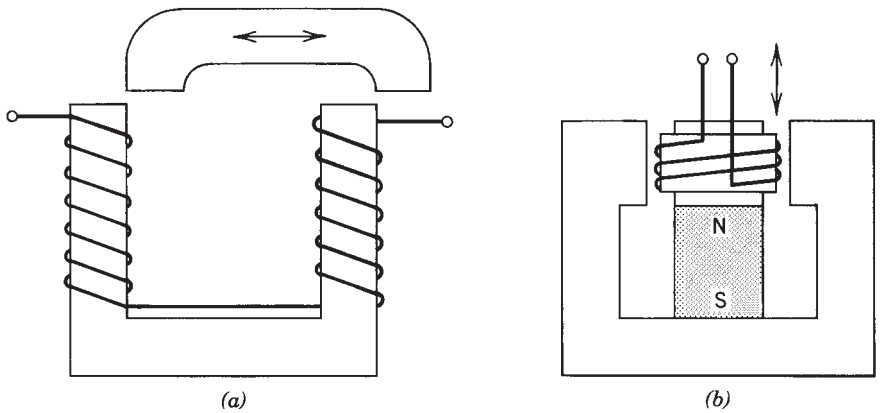


FIGURE 11.13. Linear motion actuators discussed in Reference [9]: (*a*) variable-reluctance device; (*b*) voice coil device.

In Reference [9], however, a general force law is stated that applies to both devices:

$$F_e = i\frac{d\varphi_0}{dx} - \frac{dW_0(x)}{dx} + \frac{\partial}{\partial x}\int_0^i \varphi_i(x, i')di'. \tag{11.32}$$

The terms in Eq. (11.32) are strictly analogous to those in Eq. (11.31). The rotor angular position θ is replaced by the linear position x of the moving element. The symbol φ_0 corresponds to φ_{sr}, and W_0 to W_{mr}. The integral term has to do with self-inductance and could also be represented as in Eq. (11.30) using a variable-inductance coefficient.

Finally, in Figure 11.14, a very different type of linear actuator is analyzed in which permanent magnets (PMs) form the moving element. The three iron cores are stationary with a simple winding distributed along the middle one. The force law used in Reference [10] is

$$F = I\frac{d\varphi_{cm}}{dp} - \frac{dW_m}{dp} + \frac{1}{2}I^2\frac{dL_c}{dp}. \tag{11.33}$$

The similarities to Eqs. (11.31) and (11.32) are evident if one notes that here p represents the magnet position.

We now create a bond graph capable of representing Eqs. (11.31)–(11.33) and thus capable of modeling any of the four devices in Figures 11.12–11.14, as well as many others. The bond graph will also predict the induced voltage

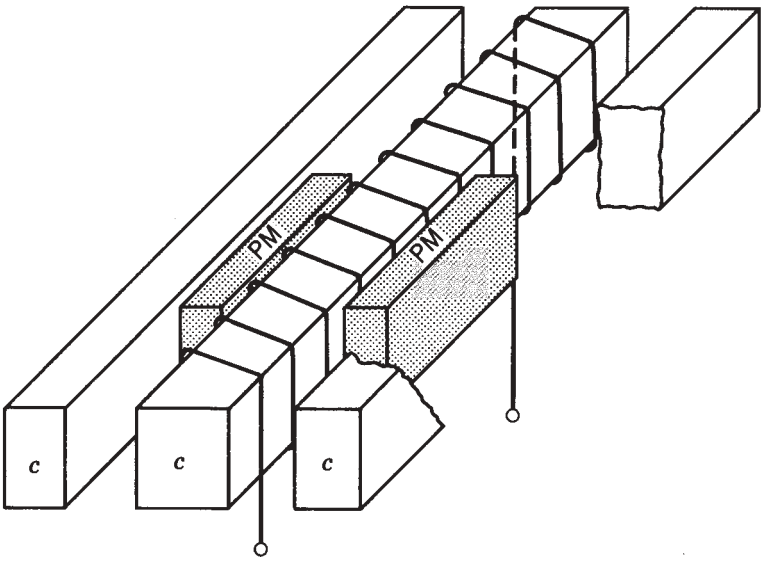


FIGURE 11.14. Linear motor with moving permanents as discussed in Reference [10].

effects consistent with the forces and will allow the simple addition of inertia, friction, and resistance effects. For simulation purposes, the forces will actually be expressed in terms of flux or flux linkage state variables rather than as functions of current.

The bond graph to be developed for the actuators is shown in Figure 11.15. The mechanical variables F and $\dot{x}$ stand for force and linear velocity, but for rotary actuators they could be interpreted as torque and angular velocity, respectively. The force law corresponding to Eqs. (11.31)–(11.33) is

$$F = i\frac{d\lambda_4}{dx} - \frac{dW_2}{dx} + \frac{1}{2}i^2\frac{dL}{dx}, \tag{11.34a}$$

$$F = F_1 - F_2 - F_3, \tag{11.34b}$$

where i is the coil current, λ_4 is the flux linkage due to the permanent magnet, W_2 is the magnetic energy in the magnetic circuit due to the permanent magnet, and the C-field or the IC-field represents the energy stored due to the coil self-inductance effect. Using a gyrator for the coil of N turns, we can conveniently relate the voltage component e_5 and current i to the rate of change of flux, $\dot{\varphi}$, in the coil and to the magnetomotive force M. Then a 2-port C-field can represent the magnetomechanical energy storage. For the electrically linear case, the position-dependent reluctance $R(x)$ is useful to express the stored energy in terms of x and φ. An alternative description uses as electric variables the current i and the voltage component $e_5 = \lambda_5$ and uses an IC-field. In this case, the position-dependent inductance $L(x)$ can be used to express the stored energy as a function of x and the self-flux linkage of the coil, λ_5. See Chapter 8.

The three force components in Eqs. (11.34a,b) should be compared directly with the corresponding terms in Eqs. (11.31)–(11.33). The components F_1, F_2, and F_3 will be discussed separately.

The force component F_1 can sometimes be computed directly from the Lorentz force. This would be the case for the voice coil of Figure 11.13*b*. A more general approach is to start from a flux linkage expression. Suppose $\lambda_4(x)$ is the coil flux linkage due to the permanent-magnet field (if one is present). In Figure 11.13*b*, the flux linkage varies with the position of the coil with respect to the magnetic circuit. In Figures 11.12 and 11.14 the configuration of the magnetic circuit changes as the magnet moves, and hence the flux linkage changes in the coils.

The voltage component e_4 is found from the rate of change of the flux linkage λ_4:

$$e_4 = \dot{\lambda}_4 = \frac{d\lambda_4}{dx}\frac{dx}{dt}. \tag{11.35}$$

By power conservation, we find

$$F_1 = \frac{d\lambda_4}{dx}i, \tag{11.36}$$

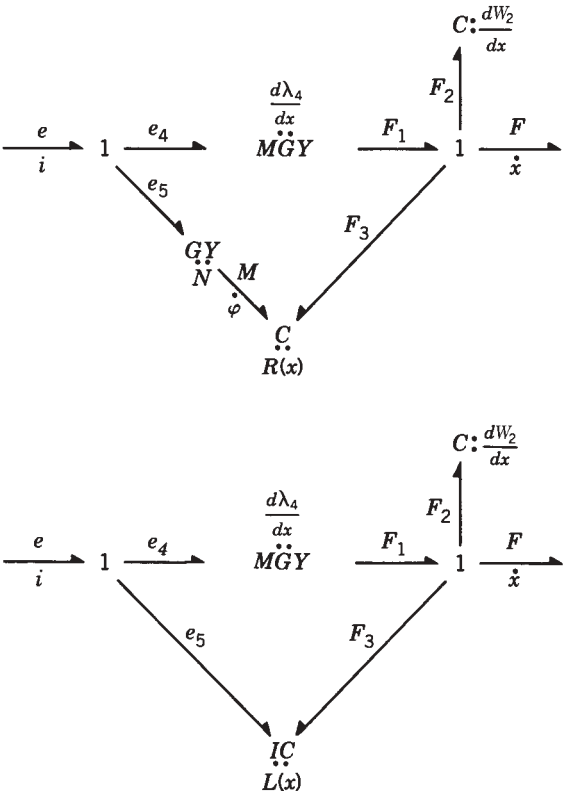


FIGURE 11.15. General bond graph for actuators in C-field form using magnetic circuit variables M and $\dot{\varphi}$ and in IC-field form using current and voltage variables.

which is represented by the modulated gyrator in Figure 11.15. For some devices, such as voice coils with modest excursions and, roughly, dc motors, $d\lambda_4/dx$ is nearly constant and the modulated gyrator can be replaced by a constant-parameter gyrator. On the other hand, if the field is produced not by a permanent magnet but rather by an externally excited field, the modulated gyrator parameter might vary with time or field current as well as with position.

The force component F_2 arises when a permanent magnet moves in a magnetic circuit. The magnet will have generally preferred positions corresponding to positions of minimum reluctance. Even without coil current there can be magnetic forces depending on position. Such forces are of the same nature as mechanical spring forces. If the energy in the magnetic circuit due to the permanent magnet, $\mathbf{W}_2(x)$, is given as a function of position, then energy conservation requires

$$F_2 = \frac{d\mathbf{W}_2(x)}{dx}. \tag{11.37}$$

Cases in which there is no permanent magnet (Figure 11.13*a*) or in which x does not influence $\mathbf{W}_2$ (Figures 11.13*b* and 11.14) have no force component F_2, so the C-element representing F_2 can be eliminated.

The final force component F_3 is due to the self-inductance effects of the coil current. For the electrically linear case, the energy $\mathbf{W}_3$ may be expressed as a function of the flux φ in the coil and the position variable x by using the reluctance $\mathbf{R}(x)$ in the form

$$W_3(x, \varphi) = \frac{\mathbf{R}(x)\varphi^2}{2}. \tag{11.38}$$

Differentiation of this expression yields the 2-port C-field laws for magnetomotive force M and force component F_2 in conservative form implied in the upper bond graph of Figure 11.15; that is,

$$M(x, \varphi) = \frac{\partial \mathbf{W}_3}{\partial \varphi} = \mathbf{R}(x)\varphi, \tag{11.39}$$

$$F_3(x, \varphi) = \frac{\partial \mathbf{W}_3}{\partial x} = \frac{\varphi^2}{2}\frac{d\mathbf{R}(x)}{dx}. \tag{11.40}$$

The gyrator with parameter N relates the voltage component e_5 to the rate of change of flux $\dot{\varphi}$ and the current i to the magnetomotive force M, yielding

$$e_5 = N\dot{\varphi}, \tag{11.41}$$

$$Ni = M. \tag{11.42}$$

Alternatively, the IC-field bond graph in the lower part of Figure 11.15 can be described by the flux linkage associated with the coil current λ_5 as the position-dependent self-inductance $L(x)$. The energy $\mathbf{W}_3$ is then

$$\mathbf{W}_3(x, \lambda_5) = \frac{\lambda_5^2}{2L(x)}, \tag{11.43}$$

which yields the IC-field laws for current force by differentiation:

$$i = \frac{\partial \mathbf{W}_3}{\partial \lambda_5} = \frac{\lambda_5}{L(x)}, \tag{11.44}$$

$$F_3(x, \lambda_5) = \frac{\partial \mathbf{W}_3}{\partial x} = -\frac{\lambda_5^2}{2}\frac{1}{L^2(x)}\frac{dL(x)}{dx}, \tag{11.45}$$

where

$$\frac{d\lambda_5}{dt} = e_5. \tag{11.46}$$

The reluctance and the inductance are related by the expression

$$L(x) = \frac{N^2}{\mathbf{R}(x)}. \tag{11.47}$$

The basic bond graph expressions for F_3 involve the state-variable flux φ or flux linkage λ_5, while the corresponding expressions in Eqs. (11.34) and the analogous expressions in Eqs. (11.31)–(11.33) use the current. To see that these expressions are exactly equivalent, we merely use Eq. (11.44) to eliminate λ_5 in terms of i in Eq. (11.45), obtaining

$$F_3(x, i) = -\frac{i^2}{2}\frac{dL(x)}{dx}. \tag{11.48}$$

The C-field laws (11.39), (11.40) and the IC-field laws (11.44), (11.45) are useful in simulation because they are in integral causality form.

The bond graph incorporates the force laws but has the additional advantage that it also automatically incorporates the voltage law for the electromagnetically induced voltage e in the form

$$e = e_4 + e_5, \tag{11.49}$$

with e_4 from Eq. (11.35) and e_5 from Eqs. (11.39)–(11.41) or from Eqs. (11.44) and (11.45). Thus, the complete actuator description is included in the bond graph in a power- and energy-consistent form.

Figure 11.16 shows the lower bond graph of Figure 11.15 supplemented by electrical resistance, mechanical inertia, and mechanical friction elements. Complete integral causality is shown: the input signals to the device are an external voltage e_{ext} and an external force F_{ext}, and the model responds with the current

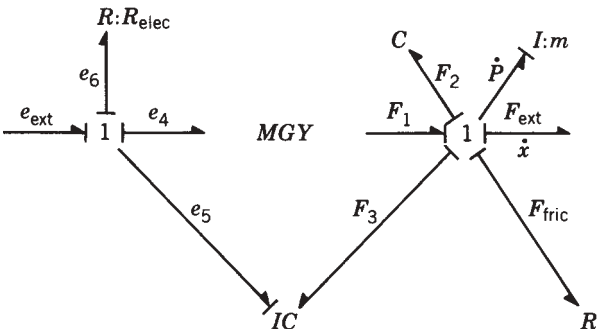


FIGURE 11.16. General bond graph with coil resistance and mechanical mass effects added. Integral causality shown.

i and velocity $\dot{x}$. Certain interconnections to other system components could, of course, induce derivative causality.

This bond graph can be simplified to reduce to known models of specific devices such as those studied in Chapter 8. A permanent-magnet dc motor, for example, does not need the IC-field, because the inductance is not a strong function of x or θ. Thus, a simple constant inductance I attached to bond 5 suffices. Also, the MGY can usually be simplified to a constant-parameter GY, and the C-element on bond 2 can be eliminated if the "cogging" torque is neglected. Analogous simplifications can be made for the devices of Figures 11.13*b* and 11.14. The device of Figure 11.13*a* has no permanent magnet, and if residual magnetic effects are neglected, both the MGY and the C on bond 2 may be removed. Then, only the energy-storing C- or IC-field remains, as in the models for relays or movable-core solenoids. Finally, there are cases in which the electrical time constants are short compared to the system time scales or the mechanical inertia and friction can be neglected. In such cases, the model may be drastically simplified and only very limited causal patterns may be possible.

Although this brief introduction to magnetic system bond graphs by no means exhausts the subject, the basic ideas have been set forth, and it is hoped that the interested reader can extend bond graph techniques to more complicated situations in magnetic systems.

REFERENCES

[1] E. Colin Cherry, "The duality between interlinked electric and magnetic circuits and the formation of transformer equivalent circuits," *Phys. Soc. London Proc.*, **62**, 101–111 (1949).

[2] R. W. Buntenbach, "Improved circuit models for inductors wound on dissipative magnet cores," *1968 Conference Record of Second Asilomar Conference on Circuits and Systems*, 68C64-ASIL, New York: IEEE, 1969, pp. 229–236.

[3] D. Karnopp, "Computer models of hysteresis in mechanical and magnetic components," *J. Franklin Inst.*, **316**, No. 5, 405–415 (1983).

[4] J. K. Watson, *Applications of Magnetism*, New York: Wiley, 1980.

[5] M. G. Say and E. O. Taylor, *Direct Current Machines*, 2nd ed., London: Pitman, 1986.

[6] P. G. Stohler and H. R. Christy, "Simulation and optimization of an electromechanical transducer," *Simulation*, **13**, No. 4, 202–210 (1969).

[7] D. Karnopp, "Bond graph models for electromagnetic actuators," *J. Franklin Inst.*, **319**, No. 1/2, 173–181 (1985).

[8] E. M. H. Kamerbeek, "Electric motors," *Philips Tech. Rev.*, **33**, No. 8/9, 215–234 (1973).

[9] J. Timmerman, "Two electromagnetic vibrators," *Philips Tech. Rev.*, **33**, No. 8/9, 249–259 (1973).

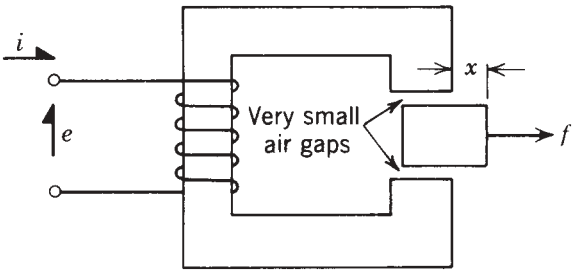
[10] L. Honds and K. H. Meyer, "A Linear C.D. motor with permanent magnets," *Philips Tech. Rev.*, **40**, No. 11/12, 329–337 (1982).

PROBLEMS

11-1. Consider the device of Figure 11.7. Suppose you are given the permeability of the core material and of air, the pertinent physical dimensions of the core, and the number of turns, N. Estimate the inductance of the device if the core material remains in the linear range of the $B-H$ curve.

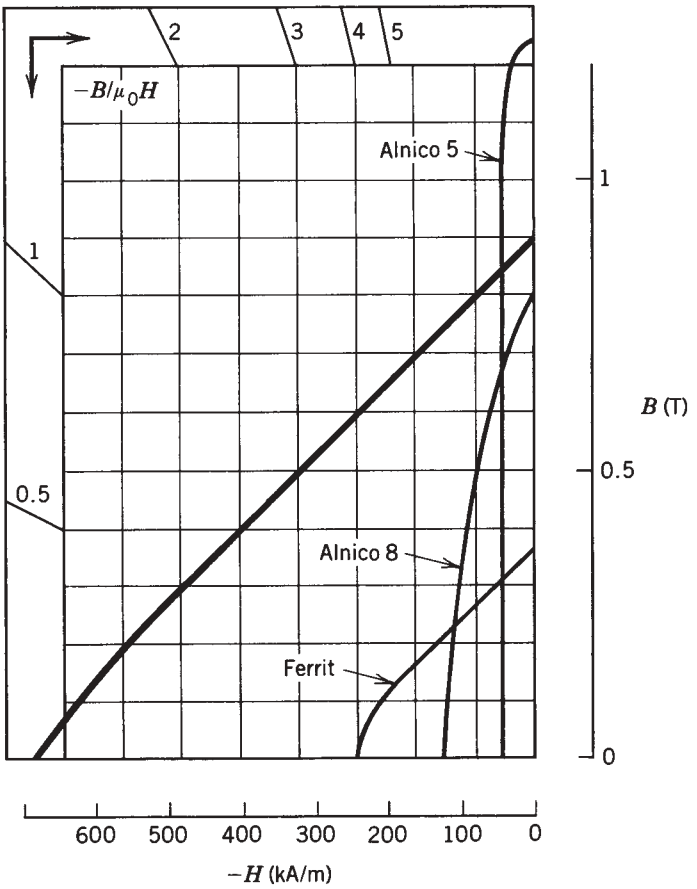
11-2. In Figure 11.8, find an expression for the capacitance parameters for the sections of core material in terms of permeability and physical dimensions in the linear case. Combine the capacitance parameters into a single equivalent capacitance, and show how to estimate the inductance using this equivalent capacitance.

11-3. In the transducer sketched below, a slug of magnetic material slides partly in and out of the flux path.



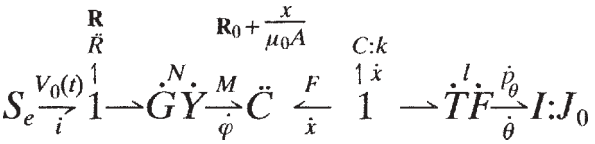
Make a simple bond graph model of the device, neglecting all loss effects and all leakage paths. By imagining how the reluctance at the slug would vary with displacement x, discuss qualitatively the difference you might expect between the f-versus-x relation at constant current for this device and for the device of Figure 11.11.

11-4. The figure shows typical demagnetization curves for several permanent-magnet materials. (This is a conventional replotting of the lower right-hand quadrant of the hysteresis curve in Figure 11.5.) The heavy curve represents a typical high-energy-product rare earth material. The $(BH)_{\max}$ point for this material has $B_m/\mu_0 H_m \cong 1$, and one can see that the Alnico materials have corresponding $B_m/\mu_0 H_m$ values greater than 1.



Consider the design of a loudspeaker voice coil as shown schematically in Figure 11.13*b*. The cylindrical coil is to have a diameter of 30 mm. The air gap is 2 mm with an axial length of 10 mm. The magnet is also cylindrical, with a diameter of 30 mm and unknown length. Assuming that the rare earth magnet material is represented by the heavy curve on the $B-H$ plot, find the optimum length of the magnet, l_n, as well as the flux density in the gap, B_g. Neglect any reluctance in the pole pieces that lead the flux to the radial gap.

11-5. The bond graph below represents a simplified version of the relay model of Figure 11.11:



The core loss elements have been eliminated, and the total core reluctance $\mathbf{R}_0$ has been incorporated in the 2-port C-field constitutive law. The air gap effective length x is related to the angle θ by the length l. The mechanical spring force is $k(x - x_0)$, which applies as long as the relay has not encountered a stop. The moment of inertia about the pivot is J_0, and the C-field reluctance is $R_0 + x/\mu_0 A$. (See Eqs. (11.18)–(11.22).) Apply causality and write the state equations for φ, x, and p_θ.

11-6. Consider the force associated with a flux φ across an air gap as illustrated in Figure 11.9. Consider only the gap reluctance approximated by Eq. (11.21) and that the MMF ΔM is generated by a coil of N turns carrying a constant current I. Express the force F as a function of x and I.

11-7. The bond graphs of Figure 11.15 apply to the devices of Figures 11.10, 11.12, 11.13, and 11.14. However, in several cases parts of the general bond graph are not needed, because they represent effects that are either totally absent or negligible in normal operation. Show simplified bond graphs for the five devices sketched, and note why some elements can be eliminated.

11-8. Suppose a permanent-magnet lifting system were capable of inducing a B-value of 0.5 T across a very narrow air gap between the magnet surface and a piece of steel. Develop an expression for the force per unit area in terms of B and μ_0, and find the force developed for an area 1 cm × 1 cm. How many kilograms could be lifted by this magnet?

11-9. Consider the design of a voice coil actuator of the type shown in Figure 11.13*b*. The design parameters are to be

$$B = 0.5\text{T}, \qquad R = 8\,\Omega, \qquad V_{\max} = 10\text{V},$$

$$\text{Allowable current density} = 20 \times 10^6 \text{A/m}^2 = \text{J},$$

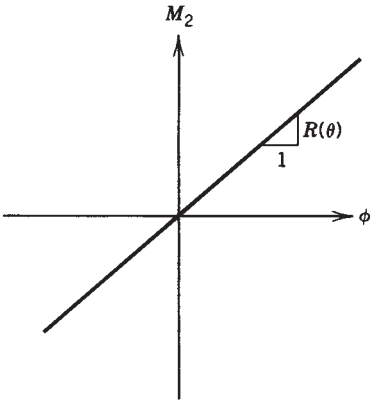
$$\text{Resistivity of copper wire} = 1.72 \times 10^{-8}\,\Omega \cdot \text{m} = \rho_{\text{Cu}}.$$

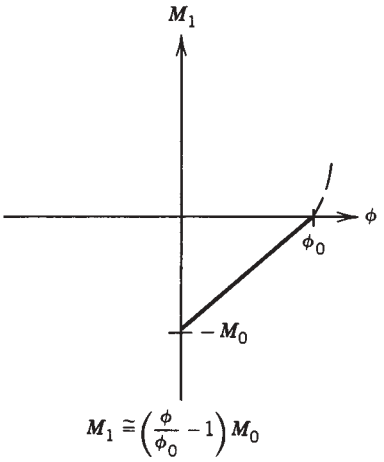
(The resistance of the coil is $R = \rho_{\text{Cu}} l/A_w$, where l is the wire length and A_w is the wire area.) What would the maximum force of this device be?

11-10. Consider a device similar to that shown in Figure 11.10 but with a permanent magnet replacing the coil. Suppose that the total reluctance of the magnetic circuit is $\mathbf{R}(\theta)$. (Equation (11.23) indicates how $\mathbf{R}(\theta)$ for the gap might vary for small angles, but $\mathbf{R}(\theta)$ would be valid for all θ and would include core reluctance.) The bond graph below shows a 2-port C based on $\mathbf{R}(\theta)$ as well as a 1-port C representing the demagnetization curve for the permanent magnet. The relevant parts of the constitutive laws are also sketched. (The $M_1 - \varphi$ sketch can be related to the lower right-hand quadrant of the B–H hysteresis loop shown in Figure 11.5.)

$$P.M.: C \overset{M_1}{\underset{\varphi}{\leftarrow}} 1 \overset{M_2}{\underset{\varphi}{\rightarrow}} \overset{R(\theta)}{C} \overset{\tau}{\underset{\dot{\theta}}{\leftarrow}}$$

Noting that $M_2 = -M_1$, the flux φ can be eliminated from the constitutive laws so that the "cogging torque" τ can be expressed in terms of the parameters M_0, φ_0 and the reluctance $R(\theta)$. Find the law that relates τ to θ.





12

THERMOFLUID SYSTEMS

This chapter discusses sophisticated modeling concepts for thermofluid systems. These systems involve more sophisticated versions of the fluid systems discussed in Chapters 3 and 4 as well as heat flow and temperature effects that have not been treated thus far in our system models.

12.1 PSEUDO-BOND GRAPHS FOR HEAT TRANSFER

We start by defining a so-called *pseudo-bond graph* for conduction heat transfer which uses effort and flow variables that do not quite fit the pattern we have used in modeling other physical systems. Traditionally, thermal systems have often been presented as analogous to electric circuits, usually with *temperature* as the *effort* variable analogous to voltage and *heat flow* as the *flow* variable analogous to current. With this analogy there are then thermal resistors, capacitors, and parallel and series connections (our 0- and 1-junctions) as well as effort and flow sources analogous to voltage and current sources. There are no thermal inertias, however.

Since this analogy in which temperature is an effort variable and heat energy flow is a flow variable has proven to be intuitively satisfying and useful in practical engineering, we present it here. There is one major hitch, however. The product of temperature (the effort) and heat flow (the flow) is *not* a power as effort times flow has been in what we now call a *true bond graph*. (Heat flow itself has the dimensions of power.) We choose to call any bond graph in which e and f are *not* power variables a *pseudo-bond graph*. In contrast, in a *true bond graph* the product of an effort and a flow on a bond represents the power flowing on the bond.

A pseudo-bond graph cannot be coupled to a true bond graph using power variables except by means of some ad hoc elements that do not obey all the rules

Copyright © 2012 John Wiley & Sons, Inc.

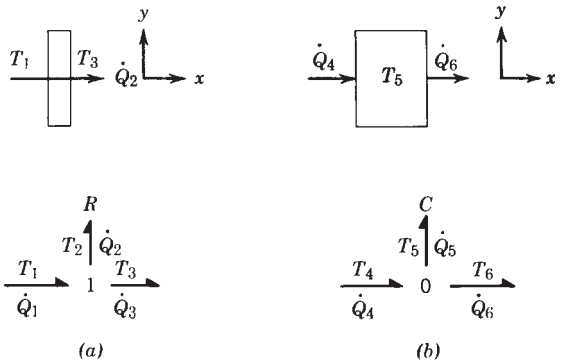


FIGURE 12.1. Basic elements for models of conduction heat transfer systems: (*a*) thermal resistor and 1-junction; (*b*) thermal capacitor and 0-junction.

of normal bond graph elements. However, bond graph techniques may be usefully applied to any pseudo-bond graph as long as the basic elements in the pseudo-bond graph correctly relate the new versions of the e, f, p, and q variables. Later, it will be shown that a *true* bond graph for heat transfer results if *temperature* and *entropy flow* are used for effort and flow variables, but the thermodynamic arguments are more complicated than those necessary to establish the usefulness of the heat transfer pseudo-bond graph. Later in the chapter, we will illustrate how true bond graph models of thermofluid systems can be constructed and finally how pseudo-bond graphs can be extended to model compressible gas dynamics in a simpler way than is possible using true bond graphs.

In Figure 12.1, two common situations that arise in the study of heat transfer systems are depicted. In both cases, we assume the temperature gradients and heat flow are present only in the x direction. The case of Figure 12.1*a* represents a pure pseudo-bond graph resistance. If T_1 and T_3 are the temperatures (on any convenient scale) on the two sides of a slab of material of area A, we assume that the flow of heat through the slab, $\dot{Q}_2$, is a function of $T_2 = T_1 - T_3$. Note that Q_2 is heat energy and $\dot{Q}_2$ has the dimensions of energy flow or power. In the linear case, we say,

$$R\dot{Q}_2 = T_1 - T_3 = T_2, \tag{12.1}$$

where $\dot{Q}_2$ may be measured in watts, Btu/s, cal/s, or any other power measure. (Because this is a pseudo-bond graph, and we will not be considering power and energy conservation in various energy domains, the choice of units is not critical as long as all element parameters are defined in a consistent manner.) For a material with thermal conductivity k, thickness l, and area A,

$$R = \frac{l}{kA}.$$

Note that in Figure 12.1*a*, the thermal resistor implies a relation between T_2 and $\dot{Q}_2$, and the 1-junction implies $\dot{Q}_1 = \dot{Q}_2 = \dot{Q}_3$ and $T_2 = T_1 - T_3 = 0$. We

have simply shown how the 1-port R and the 1-junction can be used together to constrain the common flow $\dot{Q}_2$ to be a function of the difference between two efforts, $T_1 - T_3$.

Figure 12.1*b* represents a lump of material that changes temperature as a function of the net heat energy stored in it. That is, T_5 is a function of

$$Q_5 = \int^t \dot{Q}_5 \, dt = \int^t (\dot{Q}_4 - \dot{Q}_6) \, dt$$

Since T_5 is an effort and Q_5, the integral of a flow, is a displacement, the element is a pseudo-bond graph capacitor. (The common use of the term *thermal capacitance* refers to the relation between temperature and heat energy, which is another good reason for the introduction of these pseudo-bond graph variables.)

The bond graph of Figure 12.1*b* combines a C— with a 0-junction to indicate that $T_4 = T_5 = T_6$ and $\dot{Q}_5 = \dot{Q}_4 - \dot{Q}_6$. In the linear case a thermal capacitance C can be defined such that

$$T_5 = T_{50} + \frac{1}{C} \int_{t_0}^t \dot{Q}_5 \, dt, \tag{12.2}$$

where T_{50} is the temperature at $t = t_0$. The capacitance C can be found by assuming that the element does negligible work by expanding or contracting, so that changes in its internal energy are only the result of $\dot{Q}_5$. Then, if c is a *specific heat*,

$$c = \frac{\partial u}{\partial T}, \tag{12.3}$$

where u is the internal energy per unit mass, and

$$C = mc,$$

where m is the mass of the material being considered. Strictly speaking, c is the specific heat *at constant volume*, but for most solids and liquids the work done by expansion is small, as compared with the internal energy Q_5, so that c can be used even if the material is allowed to expand. (The case in which gases do significant work as they are heated will be treated in Sections 12.4 and 12.5.)

Figure 12.2 shows a typical use of the elements shown in Figure 12.1 and introduces a thermal effort or temperature source. A unit area of the wall of a pipe carrying hot fluid is modeled. A lumped-parameter model consisting of three resistors and two capacitors is used, and the inside and outside temperatures are assumed to be determined by effort (temperature) sources.

If T_1 changes because the temperature of the fluid flowing through the pipe changes, the bond graph model can give a prediction of the dynamic changes in temperature in the pipe wall. If the pipe is really made of a uniform material,

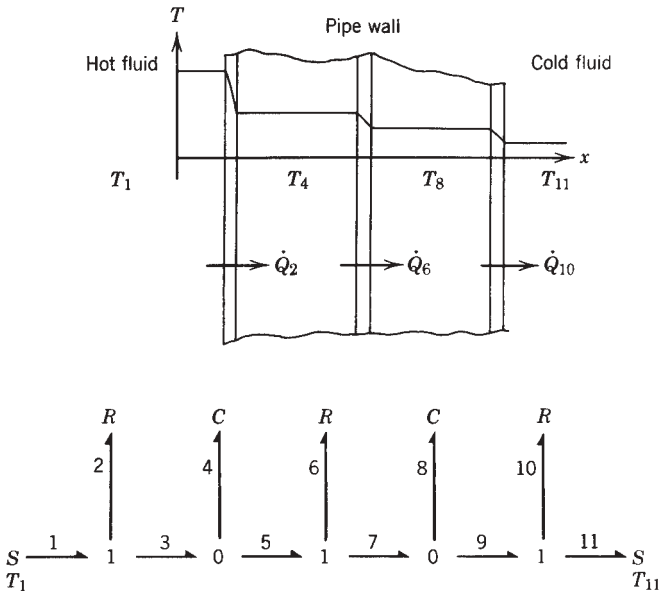


FIGURE 12.2. Heat transfer in pipe wall.

then the stepwise distribution of temperature with distance, x, that the model provides will be a fairly crude approximation of the actual temperature distribution (a continuous function of both time and distance, x). A better approximation could be obtained by splitting the pipe into a large number of resistor and capacitor layers, but the model would be more complicated to use.

As one can see from this simple discussion, the use of pseudo-bond graphs for conduction heat transfer problems leads to systems analogous to electrical RC circuits. However, general thermofluid systems can be considerably more complex than this heat transfer example so we will present a succession of true bond graph and pseudo-bond graph techniques to handle the various types of thermofluid systems of engineering significance.

12.2 BASIC THERMODYNAMICS IN TRUE BOND GRAPH FORM

Thermodynamics is essentially the universal science of physical processes. All of the models of physical systems that have been made so far may be studied from the point of view of thermodynamics, and indeed, when C- and I-fields were constrained to conserve energy or R-fields were arranged to dissipate power, essentially thermodynamic arguments were used. On the other hand, since heat flow and temperature dependence have not yet been discussed, most of our dynamic models can be characterized as *isothermal models*; that is, we treated our multiports as if they were immersed in an infinite reservoir that maintains

constant temperature even when energy is lost or supplied. In fact, we know that all elements change their constitutive laws with temperature to greater or lesser extent, and the transfer of heat and change in temperature are the special province of thermodynamics.

Much of engineering thermodynamics is not very dynamic at all. It deals with changes between equilibrium states but often is not concerned with the process of change itself. In many cases, useful information may be extracted from knowledge of the constraints that end-point equilibrium states of a system must satisfy. But if we are really interested in the path a system will follow in state space, then we must study what has come to be called *irreversible thermodynamics*, which involves non-equilibrium systems.

In recent years, there has been an upsurge of interest in making dynamic models for thermodynamic systems in which a variety of electrical, chemical, mechanical, fluid dynamic, and heat transfer effects are important. Here we can only give an introductory account in which the model systems discussed previously are shown to have thermodynamic implications. We do this by the device of considering systems of such physical extents and time scales that local pseudo-equilibrium conditions exist. In this way, we may modify some familiar relations from thermostatics into a slowly varying form of irreversible thermodynamics.

Another restriction for the present is to consider only closed systems, that is, systems in which no mass passes through the system boundaries. This restriction to a Lagrangian point of view will be lifted in the sections on fluid mechanics and compressible gas dynamics. The alternative to the Lagrangian point of view, the control volume or Eulerian point of view, which is often used for certain types of systems, introduces a set of modeling difficulties that are best left to be discussed in later sections.

Consider first, then, a fixed amount of a pure substance that is in at least pseudo-equilibrium, so that all parts of the substance are at essentially identical conditions of pressure, temperature, density, and so forth.

We consider a unit mass of the substance. Extensive quantities such as volume, internal energy, and entropy for a unit mass will be designated by the lowercase letters v, u, s. When a mass of the substance, m, is involved, the corresponding variables will be designated by uppercase letters: $V = mv$, $U = mu$, and $S = ms$. The substance is supposed to be of such small extent and the disturbances on such a slow time scale that wave motion, turbulence, and so forth are negligible. Then in the absence of motion and electromagnetic or surface-tension forces, we may assume that such a pure substance has only two independent properties. All other properties are related to any two independent properties by *equations of state* for the substance.

A useful, and logical, way to describe the constitutive relations for a pure substance begins with the statement that the internal energy per unit mass, u, of the substance is a function of the volume v and entropy s per unit mass:

$$u = u(s, v). \tag{12.4}$$

The well-known Gibbs equation relates changes in u to changes in v and s:

$$du = T\,ds - p\,dv, \tag{12.5}$$

where T is the thermodynamic temperature and p is the pressure. (In all that follows, T will indicate thermodynamic or absolute temperature, a positive quantity.)

The relation (12.5) involves energy or work. If the substance changes state slowly enough, we may treat it as if it were a lumped-parameter multiport with the relation

$$\frac{du}{dt} = T\frac{ds}{dt} - p\frac{dv}{dt}. \tag{12.6}$$

All three terms in Eq. (12.6) have the dimensions of power. The term $T\,ds/dt$ is associated with a flow of heat or, sometimes, with dissipative work such as might be done by a paddle wheel slowly stirring a fluid. The term $p\,dv/dt$ represents the sort of reversible power with which we have been dealing frequently in dynamic models. (Note that dv/dt is a volume flow rate for the unit mass; this is the type of variable used in Chapter 4 as a bond graph flow variable in hydraulic systems.)

Since

$$du = \frac{\partial u}{\partial s}ds + \frac{\partial u}{\partial v}dv,$$

it is clear that

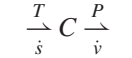
$$T = \frac{\partial u}{\partial s} \quad \text{and} \quad -p = \frac{\partial u}{\partial v}. \tag{12.7}$$

Also

$$\left.\frac{\partial T}{\partial v}\right|_s = \frac{\partial^2 u}{\partial v \partial s} = \left.\frac{\partial(-p)}{\partial s}\right|_v, \tag{12.8}$$

which we may recognize as a reciprocal relation similar to those found for other energy-storing fields as studied in Chapter 7. In thermodynamics, Eq. (12.8) is called a *Maxwell reciprocal relation*.

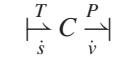
The constitutive laws of the pure substance may be represented by a true bond graph C-field; thus,



if we are willing to call T and p efforts, $\dot{s}$ and $\dot{v}$ flows, and s and v displacements.

This identification of T as an effort and $\dot{s}$ as a flow distinguishes a true bond graph (with $T\dot{S}$ being power) from the pseudo-bond graph of the previous section that used T as an effort and $\dot{Q}$ as a flow (for which the product $T\,\dot{Q}$ is temperature *multiplied by* power).

Integral causality,



implies

$$T = T(s, v), \quad p = p(s, v), \tag{12.9}$$

which are also implied by Eqs. (12.4) and (12.5). The only unusual feature of the pure-substance C-field is the sign convention on the $p\dot{v}$ port, which is positive outward corresponding to the traditional form of the Gibbs equation, Eq. (12.5).

The internal energy state function is associated with all integral causality and with constitutive laws in the form of (12.9). For mixed- and all-derivative causality, the constitutive laws are switched around. In thermodynamics this switching around of independent and dependent variables is often accomplished using the Legendre transformations of the internal energy u. The enthalpy h, Helmholtz free energy f, and Gibbs free energy ϕ are all Legendre transformations of the internal energy u and correspond to different causal patterns. The enthalpy is

$$h(s, p) \equiv u + pv, \tag{12.10}$$

and its derivatives are

$$\frac{\partial h}{\partial s} = T(s, p), \tag{12.11}$$

$$\frac{\partial h}{\partial p} = v(s, p). \tag{12.12}$$

Using $h(s, p)$, another Maxwell relation can be found:

$$\left.\frac{\partial T}{\partial p}\right|_s = \frac{\partial^2 h}{\partial s \partial p} = \left.\frac{\partial v}{\partial s}\right|_p. \tag{12.13}$$

The causal pattern for the C-field corresponding to h is shown below:

$$h = h(s, p) \Leftrightarrow \overset{T}{\underset{\dot{s}}{\vdash\!\!\rightarrow}} C \overset{P}{\underset{\dot{v}}{\vdash\!\!\rightarrow}}.$$

The Helmholtz free energy is defined by the transformation

$$f(T, v) \equiv u - Ts. \tag{12.14}$$

Its derivatives are

$$\frac{\partial f}{\partial T} = -s(T, v), \tag{12.15}$$

$$\frac{\partial f}{\partial v} = -p(T, v). \tag{12.16}$$

And yet another Maxwell relation is

$$\left.\frac{\partial(-s)}{\partial v}\right|_T = \frac{\partial^2 f}{\partial T \partial v} = \left.\frac{\partial(-p)}{\partial T}\right|_v. \tag{12.17}$$

The causal pattern corresponding to f is

$$f = f(T, v) \Leftrightarrow \overset{T}{\underset{\dot{s}}{\rightarrow}}|C\overset{P}{\underset{\dot{v}}{\rightarrow}}|.$$

Finally, the Gibbs free energy is defined by a double Legendre transformation on u:

$$\phi(T, p) = u + pv - Ts. \tag{12.18}$$

The derivatives of ϕ are

$$\frac{\partial \phi}{\partial T} = -s(T, p), \tag{12.19}$$

$$\frac{\partial \phi}{\partial p} = +v(T, p), \tag{12.20}$$

and the Maxwell relation is

$$\left.\frac{\partial(-s)}{\partial p}\right|_T = \frac{\partial^2 \phi}{\partial T \partial p} = \left.\frac{\partial v}{\partial T}\right|_v. \tag{12.21}$$

The Gibbs free energy corresponds to all-derivative causality for the C-field:

$$\phi = \phi(T, p) \Leftrightarrow \overset{T}{\underset{\dot{s}}{\rightarrow}}|C|\overset{P}{\underset{\dot{v}}{\rightarrow}}.$$

The energy function u and the co-energy functions h, f, and ϕ, which are Legendre transformations of u, have been expressed in terms of their own natural variables. When this is done, the constitutive functions for the substance may be found by differentiation, and such constitutive laws automatically ensure that the C-field will be conservative. On the other hand, one could write out constitutive laws in any form without the use of state functions such as u, h, f, or ϕ. The only difficulty is that arbitrary constitutive laws would not, in general, allow the existence of an internal energy function. In a cycle, a substance with arbitrary constitutive laws might allow net energy production and thus the construction of a perpetual motion machine of the first kind. In a sense, deriving constitutive laws from u, h, f, or ϕ is safer than not using state functions, since conservation of energy will be built into the relations.

Although there is elegance to the connection between thermodynamic C-fields and states and the C-fields discussed in Chapter 7 for other physical systems, the use of true bond graphs for thermofluid systems is not without its practical difficulties.

A classical example of a constitutive law not stated in terms of state function derivatives is the perfect gas law. This very useful law can be written

$$pv = RT, \tag{12.22}$$

for unit mass, in which R is a constant. It is often stated, in addition, that u is a function of temperature only. This means that if the constitutive laws for the gas were used to express $u(s, v)$ in terms of, say, T and v, then u would be a function only of T and not of v. As an exercise, we first we derive this fact using some of the relations shown above.

If Eq. (12.22) is solved for p in terms of T and v, the constitutive law in the form of Eq. (12.16) is obtained:

$$\left.\frac{\partial f}{\partial v}\right|_T = -p(T, v) = \frac{-RT}{v}. \tag{12.23}$$

Using this result, the Helmholtz free energy can be found by integration:

$$f(T, v) = -RT \ln v + \psi(T), \tag{12.24}$$

where $\psi(T)$ is some function of temperature alone. Then, using Eq. (12.15),

$$-s(T, v) = \frac{\partial f}{\partial T} = -R \ln v + \frac{d\psi(T)}{dT}. \tag{12.25}$$

Finally, substituting Eq. (12.25) into Eq. (12.14) and solving for u,

$$\begin{aligned} u(T, v) &= f(T, v) + Ts = -RT \ln v + \psi(T) + RT \ln v - T\frac{d\psi(T)}{dT} \\ &= \psi(T) - T\frac{d\psi(T)}{dT}, \end{aligned} \tag{12.26}$$

which demonstrates that u is only a function of T.

Although we have now shown that u can be represented as a function of T, we really need u in terms of s and v in order to evaluate the complete equations of state. There are two complete equations of state because the gas is a 2-port C-field. We need more information than just Eq. (12.22). (We could solve for the function ψ in Eq. (12.24) for example.) It is more common to assume that the two so-called specific heats are constant. The specific heat at a constant pressure, c_p, is defined as

$$c_p = \left.\frac{\partial h}{\partial T}\right|_p, \tag{12.27}$$

and the specific heat at constant volume, c_v, is

$$c_v = \left.\frac{\partial u}{\partial T}\right|_v. \tag{12.28}$$

By substituting Eq. (12.22) into Eq. (12.10) and differentiating as in Eq. (12.27), we find that

$$c_p = c_v + R. \tag{12.29}$$

If we now assume that c_v is constant, then Eq. (12.28) may be integrated to yield

$$u = c_v(T - T_0), \tag{12.30}$$

and, noting that Eqs. (12.10) and (12.22) imply that h is also a function only of T, Eq. (12.27) may be integrated similarly to yield

$$h = c_p(T - T_0), \tag{12.31}$$

in which the subscript 0 stands for a state in which it is assumed that $u = h = s = 0$.

Rearranging the basic equation (12.5) and using Eq. (12.30), we can find s:

$$ds = \frac{du}{T} + p\frac{dv}{T} = c_v\frac{dT}{T} + R\frac{dv}{v},$$

or

$$s = c_v \ln\frac{T}{T_0} + R\ln\frac{v}{v_0},$$

which may be solved for $T(s, v)$:

$$T = T_0 e^{s/c_v}\left(\frac{v}{v_0}\right)^{-R/c_v}. \tag{12.32}$$

Further manipulations of Eq. (12.5) yield

$$ds = c_p\frac{dv}{v} + c_v\frac{dp}{p},$$

$$s = c_p \ln\frac{v}{v_0} + c_v\ln\frac{p}{p_0},$$

which yield $p(s, v)$,

$$p = p_0 e^{s/c_v}\left(\frac{v}{v_0}\right)^{-c_p/c_v}. \tag{12.33}$$

Equations (12.32) and (12.33) are the complete relations for the C-field in the form (12.4). It is desirable, however, to show that these equations do indeed derive from $u(s, v)$. Using Eq. (12.30) and (12.32), we have

$$u(s, v) = c_v T_0\left[e^{s/c_v}\left(\frac{v}{v_0}\right)^{-R/c_v} - 1\right], \tag{12.34}$$

from which we find

$$T = \frac{\partial u}{\partial s} = T_0 e^{s/c_v}\left(\frac{v}{v_0}\right)^{-R/c_v},$$

which agrees with Eq. (12.32), and

$$-p = \frac{\partial u}{\partial v} = c_v \frac{T_0}{v_0} e^{s/c_v} \left(\frac{-R}{c_v}\right) \left(\frac{v}{v_0}\right)^{(-R/c_v)-1} = \frac{-RT_0}{v_0} e^{s/c_v} \left(\frac{v}{v_0}\right)^{-(R+c_v)/c_v}, \tag{12.35}$$

which agrees with Eq. (12.33) upon using Eqs. (12.22) and (12.29).

The reader probably feels that all of the above manipulations of the perfect gas law have only produced some complicated laws in Eqs. (12.32) and (12.33) in place of the simple law $pv = RT$. You are entirely correct in this view, but it is important to remember that (1) $pv = RT$ is not a complete characterization of the C-field of the gas and (2) $pv = RT$ is in a form that only allows differential causality on the $T\dot{S}$ port. Equations (12.32) and (12.33) are complete and in integral causality form at both C-field ports and hence are useful in making a dynamic model of a quantity of gas as a thermomechanical transducer using true bond graphs.

Later in this chapter we show that practical models of thermodynamic systems are often more easily made by defining pseudo-bond graphs rather than continuing to use true bond graphs.

12.3 TRUE BOND GRAPHS FOR HEAT TRANSFER

The $T\dot{S}$ port for the pure-substance C-field may be used to model all types of power flow that are associated with entropy increase, or, in other words, irreversible effects. Changes in entropy of a pure substance can be accomplished by a variety of dissipative effects, such as stirring with a paddle wheel or heating with an electrical resistor immersed in a fluid. In each case, one may identify the power dissipated with $T\dot{S}$ in order to find the change in state of the substance. An important use of $T\dot{S}$ ports is in modeling the effects of heat flow. Using $\dot{\mathbf{Q}}$ to stand for the rate of transfer of heat in power units, we may often identify $\dot{\mathbf{Q}}$ with $T\dot{S}$.

Consider first the simple case of conduction heat transfer shown in Figure 12.3*a*. This is the same type of system studied using pseudo-bond graphs in Section 12.1 but now we will use entropy flow rather than heat flow as the flow variable in our true bond graph.

The idea is that two reservoirs of thermal energy at absolute temperatures T_1 and T_2 are allowed to communicate through a thermal resistance but in no other way. Generally, a heat flow $\dot{\mathbf{Q}}$ will be set up between the two reservoirs. It is common experience that heat flows from higher toward lower temperature, and, indeed, this observation is behind the second law of thermodynamics. In Figure 12.3*b*, possible relations between $\dot{\mathbf{Q}}$ and $T_1 - T_2$ are sketched for the thermal resistance. All we need to assume is that $\dot{\mathbf{Q}}$ is related to $T_1 - T_2$ in such a way that any values of $\dot{\mathbf{Q}}$ and $T_1 - T_2$ would plot in the first and third quadrants of the $\dot{\mathbf{Q}}$ versus $T_1 - T_2$ plane. (Although it is common to assume that $\dot{\mathbf{Q}}$ is a monotonic function of $T_1 - T_2$ as sketched, the argument is true if $\dot{\mathbf{Q}}$ is positive when $T_1 - T_2$ is positive, negative when $T_1 - T_2$ is negative, and zero when $T_1 - T_2$ is zero.)

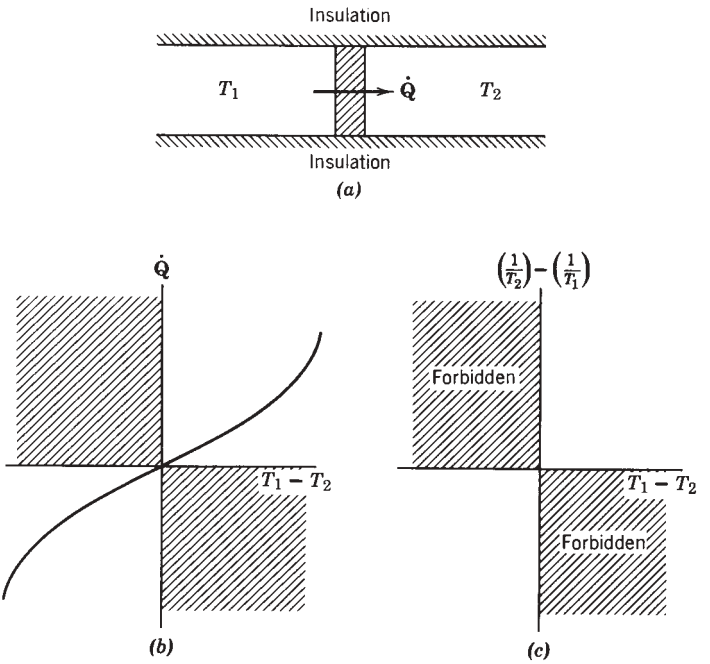


FIGURE 12.3. Conduction heat transfer: (a) thermal resistance; (b) allowed relations between Q and $T_1 - T_2$; (c) forbidden regions for $1/1T_2 - 1/1T_1$.

If we now write

$$T_1\dot{S}_1 = \dot{\mathbf{Q}} = T_2\dot{S}_2, \tag{12.36}$$

which implies that the heat flow out of one body is instantaneously equal to the heat flow into the other body, then the net entropy flow rate, $\dot{S}_2 - \dot{S}_1$, may be found:

$$\dot{S}_2 - \dot{S}_1 = \frac{\dot{\mathbf{Q}}}{T_2} - \frac{\dot{\mathbf{Q}}}{T_1} = \dot{\mathbf{Q}}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) = \dot{\mathbf{Q}}\frac{T_1 - T_2}{T_1 T_2}. \tag{12.37}$$

In Figure 12.3c, it is shown that $1/T_2 - 1/T_1$ is positive, negative, or zero as $T_1 - T_2$ is positive, negative, or zero. This is true because the absolute temperatures are inherently positive quantities. Thus, the net entropy production rate, which is the product of $\dot{\mathbf{Q}}$ and $1/T_2 - 1/T_1$, is positive for any finite value of $T_1 - T_2$ and vanishes only when $\dot{\mathbf{Q}}$ vanishes. The thermal resistor may be represented by the 2-port field shown below:

$$\underset{\dot{S}_1}{\overset{T_1}{\rightharpoonup}} R \underset{\dot{S}_2}{\overset{T_2}{\rightharpoonup}}.$$

It may come as a surprise that the true bond graph requires a 2-port R to represent the same system that the pseudo-bond graph represented as a 1-port R

attached to a 1-junction. This is an indication that the analogy between electrical RC systems and heat transfer systems is not as complete as is often assumed.

Although the 2-port resistive field is power conservative (Eq. (12.36)), it is neither a transformer nor a gyrator. It has the peculiar property that, with the sign convention shown above,

$$\dot{S}_2 - \dot{S}_1 \geq 0,$$

which implies that the entropy flow leaving the field is greater than the entropy flow entering the field no matter which way the heat is flowing. When the thermal resistor is part of a system, it will tend to increase the entropy of the system whenever any heat flows through the resistor.

A possible constitutive law for the resistor is

$$\dot{\mathbf{Q}} = H(T_1 - T_2),$$

or

$$\dot{S}_1 = \frac{H(T_1 - T_2)}{T_1}, \quad \dot{S}_2 = \frac{H(T_1 - T_2)}{T_2}, \tag{12.38}$$

where the heat transfer coefficient H is assumed to be constant or perhaps a function of the average temperature, $(T_1 + T_2)/2$. As a check, the net entropy production rate is

$$\dot{S}_2 - \dot{S}_1 = \frac{H(T_1 - T_2)}{T_2} - \frac{H(T_1 - T_2)}{T_1} = \frac{H(T_1 - T_2)^2}{T_1 T_2} > 0, \tag{12.39}$$

which is always positive as it should be.

Note that Eq. (12.38) is written in the causal form,

$$\overset{T_1}{\underset{\dot{S}_1}{\rightarrow|}} R \overset{T_2}{\underset{\dot{S}_2}{|\rightarrow}}.$$

Given any two temperatures $T_1 > 0$, $T_2 > 0$, it is possible to solve for $\dot{\mathbf{Q}}$, $\dot{S}_1$, and $\dot{S}_2$. The remaining causal forms are not so useful. For example,

$$\overset{T_1}{\underset{\dot{S}_1}{|\rightarrow}} R \overset{T_2}{\underset{\dot{S}_2}{\rightarrow|}}.$$

implies that given any $\dot{S}_1$ and $\dot{S}_2$, the resistance should provide T_1 and T_2. But we cannot actually impose arbitrary variables $\dot{S}_1$ and $\dot{S}_2$, since Eq. (12.37) or (12.39) must be obeyed and the temperatures can never be negative. This causality is therefore not useful for dynamic systems. Even mixed causalities are fraught with similar difficulties, so in what follows, the thermal resistor will be assumed to accept only the causality $-|R|-$.

12.3.1 A Simple Example of a True Bond Graph Model

Figure 12.4 shows a system incorporating two chambers containing compressible fluids. One chamber is of constant volume, so that the internal energy of the fluid contained in the chamber varies only because of heat flow through the partition, which is modeled as a thermal resistance. The volume of the other chamber is changed as a piston moves back and forth.

The bond graph for the system shown in Figure 12.4*b* incorporates two capacitance elements for the fluids and an R-field for the thermal resistance. The 0-junction is used simply to achieve inward sign conventions for both the $T_1\dot{S}_1$ port of the left-hand C and the resistive field, so that the sign conventions assumed previously are reflected in the model. The 0-junction assures that $T_1 = T_2$ but $\dot{S}_1 = -\dot{S}_2$.

In Figure 12.4*c* integral causality is shown for the system. The preferred R-field causality is compatible with integral causality on both C-elements. The state variables are S_1 (the entropy of the fixed-volume fluid), S_3 (the entropy of the variable-volume fluid), and V_4 (the volume of the variable chamber). The velocity $\dot{X}_5(t)$ plays the role of an input variable.

The constitutive laws for the fluids could be found by starting with the internal energy per unit mass, $u(s, v)$, a function of entropy per unit mass and specific volume, and then converting to the total energy U as a function of total entropy S and volume V. The perfect gas internal energy function of Eq. (12.34) could be used, for example. The evaluation of U requires a knowledge of the total mass as well as the initial state of the fluids. The pressure and temperature constitutive relations could then be found by differentiating U with respect to S and V. (In the left-hand chamber, the volume does not actually change during a simulation

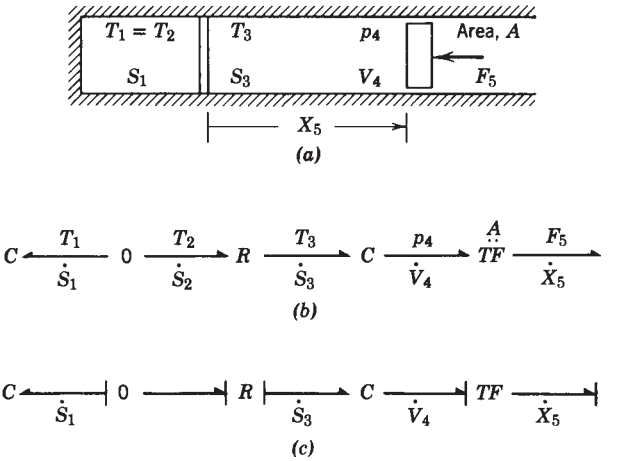


FIGURE 12.4. Example system: (*a*) schematic diagram; (*b*) bond graph; (*c*) bond graph with integral causality.

but the entropy does.) The constitutive relations for the thermal resistance have been discussed in the previous section.

Clearly, the model for the system represented by the bond graph can only be appropriate for slow changes. We assume, for example, that the fluids are in pseudo-equilibrium, so that the temperature and pressure are essentially uniform throughout the volumes at every instant. This would not be the case if one needed to worry about acoustic waves or the details of heat transmission in the fluids themselves. Also, although the entire system is power conservative, the entropy of the two C-elements will increase if any heat flows. From the point of view of the external port, this irreversibility makes it appear that energy is lost. For example, if X_5 is taken through a cycle starting from an equilibrium state at which $T_2 = T_3$, the fluid temperature in the variable volume will change, heat will flow, and, during the cycle, a net energy loss will be observed at the external port. What the bond graph shows is that energy is not lost, but rather converted back into thermal energy that cannot be completely converted back into mechanical energy except in the limiting case in which X_5 moves so slowly that T_2 and T_3 are virtually identical and the net entropy production almost vanishes.

In the model of Figure 12.4 no irreversible phenomena other than heat transfer have been included. It would not be hard to include other dissipative effects. For example, if the piston had Coulomb friction and if all the mechanical energy lost in friction were converted into thermal energy, a simple model would involve the creation of another equivalent entropy flow to the fluid C-field equal to the dissipated power divided by the fluid temperature. A bond graph for this case is shown in Figure 12.5.

The Coulomb friction resistor may be described by a relation such as

$$F_6 = A \operatorname{sgn} \dot{X}_5,$$

so that the power dissipated is

$$F_6 \dot{X}_5 = A \dot{X}_5 \operatorname{sgn} \dot{X}_5 \geq 0.$$

The entropy flow $\dot{S}_4$ is then just

$$\dot{S}_4 = \frac{1}{T_3} A \dot{X}_5 \operatorname{sgn} \dot{X}_5 \geq 0.$$

The 2-port R-field version of a mechanical resistor can accept either causality at the mechanical port but only the causality shown in Figure 12.5 at the thermal

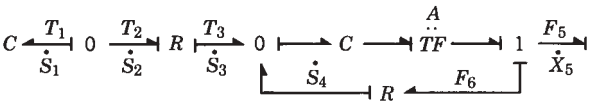


FIGURE 12.5. System of Figure 12.4 with piston friction.

port. Actually, the frictional energy may not reach the fluid instantaneously, as assumed in the model: It may heat the piston and wall material first and then warm the fluid. A more complex model with other thermal capacitances could show such effects.

12.3.2 An Electrothermal Resistor

Power dissipation always results in thermal effects, although these effects can sometimes be neglected or treated separately from a main dynamic model. For example, electrical resistors heat up as electrical power is dissipated, but in many cases the heating does not change the characteristics of the device very much if a means of cooling the resistor is provided. For example, typical circuit models often assume that the circuit components remain essentially at a constant temperature. Here, we construct a true bond graph model that contains both electrical and thermal effects.

The resistor is sketched in Figure 12.6*a*. The body of the resistor is assumed to have a fairly uniform temperature T and to be immersed in an atmosphere at temperature T_0. The resistor is assumed to dissipate electrical power:

$$ei \geq 0. \tag{12.40}$$

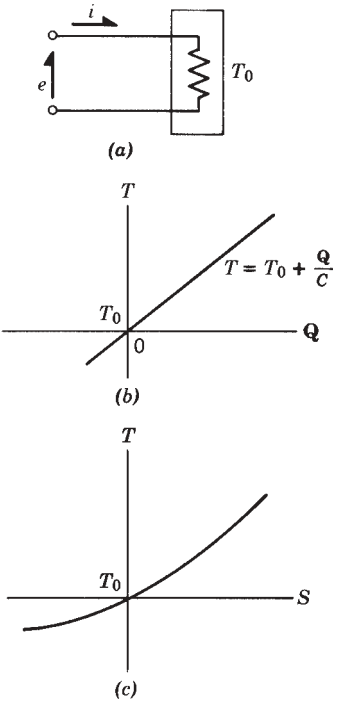


FIGURE 12.6. The electrothermal resistor: (*a*) schematic diagram; (*b*) definition of thermal capacity; (*c*) capacity relationship in terms of entropy.

As long as Eq. (12.40) is obeyed, we may assume that the relation between e and i for the resistor has some temperature dependence.

The power lost electrically is often spoken of as being converted to heat, so we will assume that

$$ei = \dot{\mathbf{Q}} = T\dot{S}, \tag{12.41}$$

in which it is a moot point whether $\dot{\mathbf{Q}}$ should be regarded as a heat flow in the usual sense. Two causal forms for an R-field representation are useful:

$$\overset{e}{\underset{i}{\vdash\!\!\rightarrow}} R \overset{T}{\underset{\dot{S}}{\vdash\!\!\rightarrow}}.$$

$$e = e(i, T), \quad \dot{S} = \frac{ie(iT)}{T}, \tag{12.42}$$

and

$$\overset{e}{\underset{i}{\leftarrow\!\!\vdash}} R \overset{T}{\underset{\dot{S}}{\vdash\!\!\rightarrow}},$$

$$i = i(e, T), \quad \dot{S} = \frac{ei(e\,T)}{T}. \tag{12.43}$$

As in previous examples, there is a preferred causality at the thermal port.

The temperature of the resistor body depends on how much thermal energy has been stored up in the resistor body. In what follows, we neglect any work done against the pressure of the atmosphere by expansion of the resistor, that is, the assumption is that no significant power flows at a $P\dot{V}$ port. The temperature, then, will depend on the amount of energy in the form $T\dot{S}$ that has been absorbed and a 1-port C element will suffice as it did for the constant volume chamber in Figure 12.4.

It is common to assume that T is a function of the stored energy, $\mathbf{Q}$, but now we must express T as a function of the entropy, S. These two points of view are readily reconciled, since $\dot{\mathbf{Q}}$ and $\dot{S}$ are related by Eq. (12.41). Suppose, for example, that a thermal capacity C (approximately constant) has been defined as in Figure 12.6*b* and in the equation

$$T = T_0 + \frac{\mathbf{Q}}{C}, \tag{12.44}$$

in which we have integrated $\dot{\mathbf{Q}}$ to find $\mathbf{Q}$ at a time when $T = T_0$. Using Eqs. (12.41) and (12.44), S may be found:

$$S = \int_0^S dS = \int_0^{\mathbf{Q}} \frac{d\mathbf{Q}}{T_0 + \mathbf{Q}/C} = \frac{C \ln(T_0 + \mathbf{Q}/C)}{T_0}, \tag{12.45}$$

where S and $\mathbf{Q}$ are both assumed to vanish at the initial time. Solving Eq. (12.45) for $\mathbf{Q}$ in terms of S and then substituting the result into Eq. (12.44), we find

$$T = T_0 e^{S/C}, \tag{12.46}$$

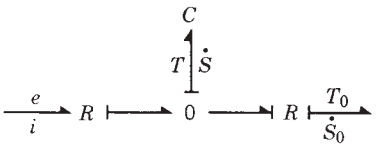


FIGURE 12.7. Bond graph for electrothermal resistor.

which is the relationship for the thermal capacitance when S is used instead of Q. (In reality, since T is an absolute temperature, $Q \cong T_0 S$ and $T \cong T_0(1 + S/C)$ for modest excursions of T from T_0.)

The bond graph of Figure 12.7 shows the complete model for the resistor. The entire system is conservative, but power is lost from the electrical port, since the power on some heat flow bonds cannot reverse. Both the electrothermal and heat transfer R-fields represent irreversible effects and generate entropy. As long as the temperature dependence of the electrical resistance is not strong, one may avoid the complicated model of Figure 12.7 in favor of a simple 1-port electrical resistance. However, in principle, the electrical and thermal systems are always coupled bilaterally, so that models of this type may sometimes be required.

The reader will have noticed in this brief introduction to true bond graph models for heat transfer systems that only closed systems were modeled; that is; no mass crossed the boundary of the system. It is also probably clear that the introduction of entropy flow as a bond graph flow variable in order to generate a true bond graph complicates the bond graph modeling process compared to the pseudo-bond graph introduced at in Section 12.1.

Much of the science of thermodynamics is concerned with flow processes in which not only power, but also mass, flows into and out of control volumes in the system being modeled. While such systems can often be modeled with true bond graphs, the convection of energy as mass moves through the boundary of a control volume complicates the modeling process. Some flow-process models, in which the distinction between Lagrangian and Eulerian descriptions of fluid motion parallels that between closed and open thermodynamic systems, will be discussed in the following sections before we define a class of pseudo-bond graphs particularly useful for open thermofluid systems.

12.4 FLUID DYNAMIC SYSTEMS REVISITED

In Chapter 4, we established effort, flow, displacement, and momentum variables for fluid systems of the closed-circuit hydraulic type, in which the static pressure times the volume flow rate represented most of the transmitted power. (In this section, as in Chapters 3 and 4, Q represents a volume flow rate—a bond graph flow variable—not heat energy as in the previous section.) We now take a deeper look at fluid systems in general and point out some of the limitations of the modeling done in Chapter 4.

Much of fluid mechanics work is concerned with field problems in which a flow field in two or three dimensions is to be determined. Such problems are usually described by partial differential equations. Unless an analytic solution to such a problem happens to be known, one must resort to finite-difference or finite-element techniques. Such techniques represent the continuum with a large number of similar lumps. Although bond graphs can be made for the "micro-lumps" involved in partial differential equations (or their finite approximations), there often is little to be gained by such a representation. The micro-lumps are all very similar, and they interact with their fellows only in standardized ways which don't benefit very much from bond graph techniques.

In this section, the main concern will be with the gross type of lumping commonly done when a fluid system is a small part of a larger system and therefore cannot practically be represented in great detail. Typical examples of such cases occur when hydraulic or pneumatic elements form part of a control system. In what follows, the main concern will be with interior flows and with fluid dynamic interaction with solid mechanical elements.

As a simple example of the sort of approximations often made in the analysis of fluid systems, consider the line element shown in Figure 12.8. The model, which is analogous to electrical-transmission-line models, attempts to treat a variety of

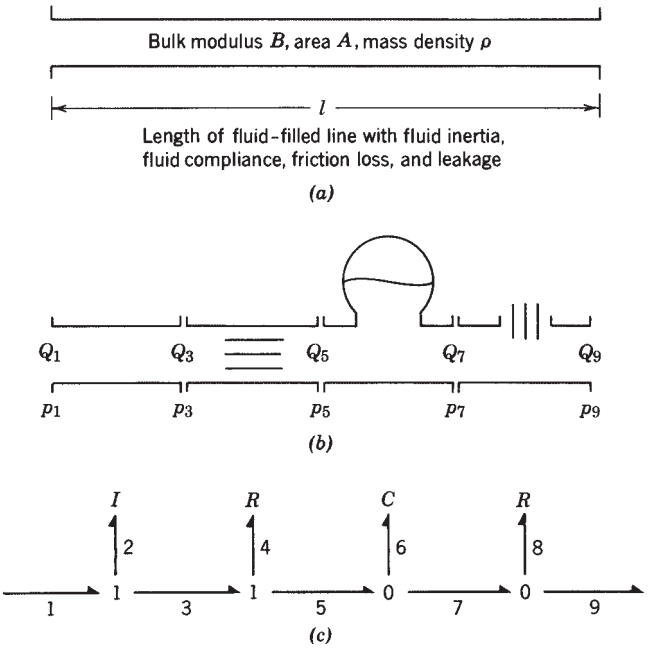


FIGURE 12.8. Lumped representation of fluid-filled lines: (*a*) sketch of line with some parameters; (*b*) schematic diagram; (*c*) bond graph.

effects clearly present in real lines. The model is plausible for short segments of line and especially for linear elements—yet it is often used for long lengths l and for nonlinear elements. As will be seen, there are some philosophical snags with this model if it is used without careful thought.

First of all, it is clear that the model is one dimensional. The volume flow rate past a unit cross-sectional area, Q, and the pressure, p, clearly must be thought of as quantities averaged over a field of velocities and pressures that vary over the line cross-sectional area. Since the velocity profile of the flow may vary widely during transient conditions, it is clear that the variables p and Q and the parameters of the —R, —C, and —I elements in Figure 12.8 cannot be evaluated exactly unless the nature of the flow over the cross-sectional area is known in advance. Nevertheless, in system design studies, one rarely is sure in advance how the flow fields in various parts of the system will behave, so that one must make at least a preliminary estimate of the system parameters before the dynamics of the flow can even be estimated.

Consider first the problem of estimating the inertia of the fluid in the line. An elementary derivation of the inertia coefficient as outlined in Section 4.3.3 proceeds as follows from Figure 12.8: The mass of the fluid in the line is ρAl, the force tending to accelerate the fluid is $A(p_1 - p_3)$, and the velocity of the fluid is Q_2/A. Thus,

$$\rho Al \frac{d}{dt}\left(\frac{Q_2}{A}\right) = A(p_1 - p_3) = Ap_2, \quad \text{or} \quad Q_2 = \frac{1}{\rho l/A} p_{p2}, \tag{12.47}$$

where p_{p_2} is the pressure momentum of bond 2 (or the time integral of p_2). This result shows that the mass-like inertia coefficient is $I = \rho l/A$ when p, Q variables are used. That the inertia of small-area lines is larger than the inertia of larger area lines of the same length was something of a surprise.

There are several problems with the simple derivation given above. First, in keeping with the assumption of one-dimensionality, the fluid in the pipe was treated as if it moved as a rigid body. It is difficult to improve on this assumption until it is known how the velocity profile of the fluid changes in space and time. The types of flows being considered here are often called *quasi-one-dimensional flows*, and in the steady state, when a well-developed velocity profile is known at each cross section, one may identify Q/A as an *average* velocity. In later calculations, the average of the square of the velocity over the cross section will be needed, and this quantity may be taken as $\beta Q^2/A^2$, where the correction factor $\beta \geq 1$ can be calculated if the velocity profile is known (see Reference [2]). The factor β is unity only for a uniform velocity distribution that would occur only for frictionless, irrotational flow. For transient conditions the velocity distribution is often very hard to estimate, so in the remainder of this section we will assume a uniform distribution even in cases in which this cannot strictly be true. The accuracy of some results may be slightly improved by the introduction of factors

such as β, although the estimation of the factors in the absence of experimental data may be difficult.

Second, unless one is willing to assume that the fluid is incompressible, the proper value of the density ρ is open to question. If $l = dx \to 0$, then $\rho(x, t)$ might represent an instantaneous value of ρ at position x, but if l is finite, then existence of capacitance in Figure 12.8 is not compatible with using ρ as a constant. If ρ varies, then, of course, the two ends of the slug of fluid do not necessarily move with the same velocity. Later on, a more sophisticated look at the problem will be taken. For now, let it simply be noted that for hydraulic systems in which the density changes are small, an average density for ρ is often sufficient, despite the contradiction implied by the use of a compliance together with a constant inertia.

One defect of this elementary derivation is that the slug of fluid is treated as a rigid body, even though it is clear that a control volume comprising a length of line through which mass and momentum flow is being considered. As will be demonstrated, momentum flows at the line ends cancel out only when the two ends of the pipe have identical cross-sectional areas. Further, the elementary derivation does not generalize readily when pipes of varying area are encountered. It is surprising that few authors of elementary system dynamics texts even mention the control volume basis for their fluid mechanical system dynamics.

The element $R4$ in Figure 12.8 represents a loss in pressure beyond that required to accelerate the fluid. The relation between p_4 and Q_4 would be easy to specify as a generally nonlinear relationship if one could use the data for flow in pipes that has been determined experimentally. But almost all the data on friction factors is for fully developed steady flow and would not apply to any but the slowest transient conditions. For this reason, the analyst must be prepared to experiment with the friction loss law until the model response matches experimental data sufficiently well.

The next element in Figure 12.8 is intended to model the compliance of the fluid and pipe walls. The flow Q_6, which is the difference between Q_5 and Q_7, represents a loss in flow between the ends of the pipe. The pressure p_6 (which equals p_5 and p_7) can be determined from the integral of Q_6 when one can define a bulk modulus B:

$$p_6(t) = p_6(0) + \frac{B}{V_0}\int_0^t Q_6\, dt = p_6(0) + \frac{B}{V_0}V_6(t), \tag{12.48}$$

where V_0 is the volume of the fluid in the pipe at $t = 0$ when the pressure is $p_6(0)$. One might think that, simply by replacing the linear relation (12.48) with some nonlinear relation between p_6 and V_6 to model compressibility effects of a gas, the same model would serve in the case in which large density changes occur, but the situation is not quite so simple. As is demonstrated in Section 12.5, when density variations are significant, the thermodynamics of the situation must be studied.

Finally, R_8 represents a loss of flow in the pipe section due to leakage. As the bond graph stands, the pressure on bond 8 represents gage pressure, i.e., the

fluid pressure with respect to atmospheric pressure. Clearly, if the p's are meant to represent absolute pressure rather than gage pressure, R_8 should react to the difference between p_8 and the pressure external to the pipe, rather than just to p_8 itself. This would require that R_8 be attached to a 1-junction that computes the difference between the internal and external pressure and imposes the condition that flow out of the pipe is the same as flow into the external atmosphere.

From the brief introduction to the lumped line elements given above, it is clear that there are some fundamental difficulties in justifying the lumping process. Many difficulties disappear when only a linearized lumped model is desired, since it then is plausible to neglect small changes in quantities that have finite mean values. To see more clearly the nature of some of the more commonly used fluid dynamic models, one needs to consider the partial differential equations of fluid mechanics and nonlinear effects.

12.4.1 One-Dimensional Incompressible Flow

To illuminate the connection between standard techniques in fluid mechanics and a lumped representation using bond graph elements, a derivation of Bernoulli's equation is given first. This is a straightforward extension of the true bond graph models considered in Chapter 4 and does not involve complicated thermodynamic effects.

Consider the problem of Figure 12.9*a*. Let s represent distance along the centerline of a curved rigid pipe of length l. Then $A(s)$ is the pipe cross-sectional area, and $x(s)$ and $z(s)$ describe the horizontal and vertical positions of the pipe centerline, respectively. Using the Eulerian method to describe the fluid motion, $v(s, t)$ represents the (average) fluid velocity at position s and time t, and $p(s, t)$ represents the pressure. After expressing the acceleration as a function of space and time, see Reference [2], Newton's law yields

$$\rho \frac{Dv}{Dt} = \rho \frac{\partial v}{\partial t} + \rho v \frac{\partial v}{\partial s} = -\frac{\partial p}{\partial s} - \rho g \frac{dz}{ds}. \tag{12.49}$$

Noting that, because of incompressibility, the volume flow rate $Q(t)$ is independent of s, we see that v is given by

$$v(s, t) = \frac{Q(t)}{A(s)}. \tag{12.50}$$

Substituting Eq. (12.50) into Eq. (12.49), we have

$$\rho \left(\frac{\dot{Q}(t)}{A(s)} + \frac{Q^2(t)}{A(s)} \frac{\partial 1/A(s)}{\partial s} \right) = -\frac{\partial p}{\partial s} - \rho g \frac{dz}{ds}, \tag{12.51}$$

which may be integrated in s from $s = 0$ to $s = l$:

$$I\dot{Q} + \frac{\rho Q^2}{2} \left(\frac{1}{A_l^2} - \frac{1}{A_0^2} \right) = p_0(t) - p_l(t) - \rho g(z_l - z_0), \tag{12.52}$$

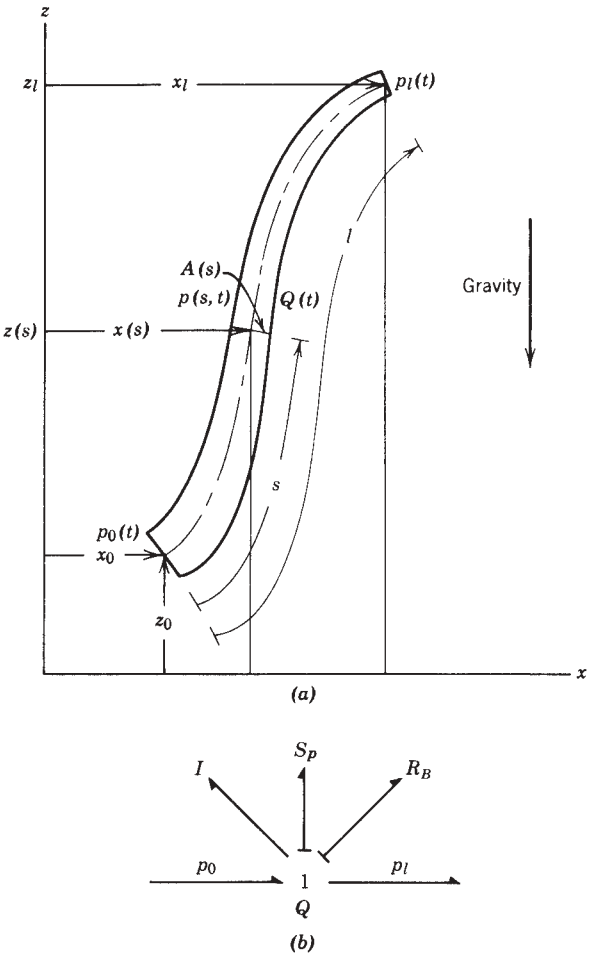


FIGURE 12.9. Constrained motion of an incompressible fluid: (*a*) sketch of system; (*b*) bond graph.

where

$$I = \rho \int_o^l \frac{ds}{A(s)}, \tag{12.53}$$

and one could substitute $\beta Q^2/A^2$ for the Q^2/A^2 terms as discussed previously.

It is interesting to note that Bernoulli's equation (essentially Eq. (12.52)) can be represented exactly by the lumped elements shown in Figure 12.9*b*.

The linear inertia coefficient defined in Eq. (12.53) reduces to that found by the nonrigorous method of Eq. (12.47) if $A(s)$ is constant, but several other terms appear in Eq. (12.52) that did not appear in Eq. (12.47). The pressure due to gravity that is represented by a constant-pressure source in Figure 12.9*b* is easily understood, but the term, involving Q^2 requires explanation.

First of all, if $A(s)$ is constant, then $A_1 = A_0$ and the Q^2 term vanishes, thus showing that Eq. (12.47) is essentially correct for the constant-area case. However, in deriving Eq. (12.47), the momentum flow of fluid through the ends of the pipe section was not properly considered. To use Newton's law in its simplest form, one must follow the flow of the fluid; that is, one must use a Lagrangian description. However, if one really wants to treat the pipe as a control volume through which fluid passes, then an Eulerian description is required. The Q^2 term in Eq. (12.52) may be thought of as a dynamic pressure-correction term, which was fortuitously absent from the constant area case of Eq. (12.47).

Another way to interpret the result of Eq. (12.52) is to note that the power flow $p(t)Q(t)$ does not represent the total flow of energy past a stationary point of the pipe. Terms of the form $\rho Q^2/2A^2$ represent *dynamic pressure* associated with the kinetic energy of the fluid. In the system under study, p and Q contain all the information required to find the dynamic pressure, since the velocity is Q/A because of the incompressibility assumtion.

The Q^2 term in Eq. (12.52) may be represented as a sort of Bernoulli resistance, R_B, in Figure 12.9*b*, since it is a relation between Q and a pressure (the dynamic pressure). When the resistance characteristic is plotted as in Figure 12.10, it is clear that if one keeps track only of the pQ power, then the dynamic pressure may be converted back and forth into static pressures as in nozzles and diffusers.

The resistance used to represent the Q^2 terms in Bernoulli's equation is unusual in two ways. First, no matter whether $A_0 > A_1$ or $A_0 < A_1$, there are regimes of operation in which the resistance appears to supply power rather than dissipate power. Physically, this means that dynamic pressure is being converted

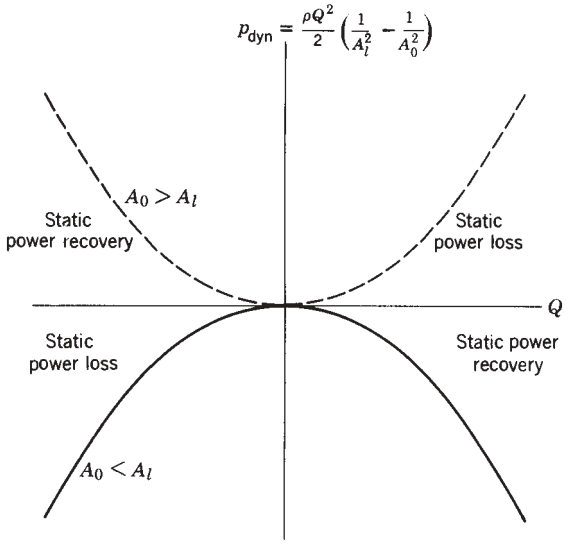


FIGURE 12.10. Constitutive law for the dynamic pressure resistor.

to equivalent static pressure and hence to an apparent power flow in the form pQ. Second, although the pressure in the resistance is uniquely determined for any given flow, given the pressure, then there is either no corresponding flow or two of them. Thus, there is a very strong preference for the causality of the resistance shown in Figure 12.9*b*.

One should also remember that the basis for our derivation was the assumption that the fluid filled the pipe, and hence $v = Q/A$. For some conditions this may not be true—a jet may form in the pipe. Also, the equations derived may be valid only for one direction of flow. In studying the flow of water through a nozzle into the atmosphere, for example, it may make no sense to consider reverse flow, because the nozzle would fill with air and the equations would no longer be meaningful. Suffice it to say that in any application of Bernoulli's equation including the bond graph representation shown in Figure 12.9, one must use care to avoid nonsensical results. Such difficulties with resistance relations are rare in other types of physical systems.

As an illustration of the utility of the bond graph representation, consider the classical elementary problem of estimating the time required for a tank to empty through a pipe. The system is shown in Figure 12.11. For simplicity, no friction losses will be considered, but the Bernoulli resistor may be thought of as indicating the loss in kinetic energy of the fluid that leaves the system. The tank has a capacitance of $A_T/\rho g$, where A_T, the tank area, is large compared to the pipe area. As in the classical analysis, we imagine the fluid entering the end

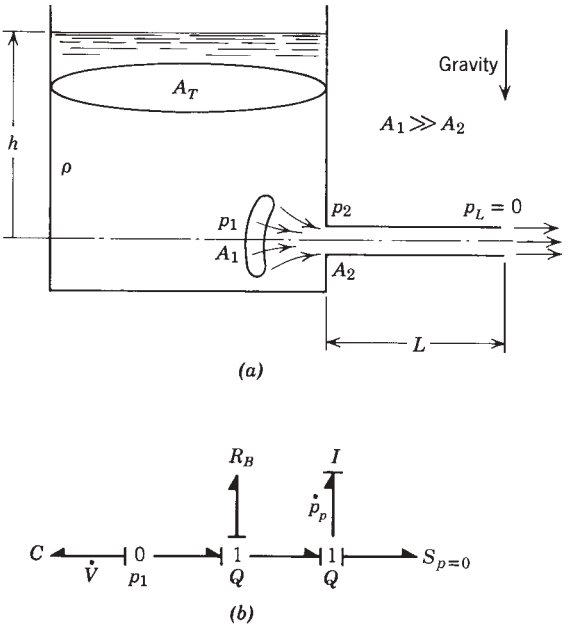


FIGURE 12.11. Tank-emptying problem: (*a*) sketch of system; (*b*) bond graph.

of the pipe from a large area, A_1, at pressure p_1 and essentially zero velocity. The Bernoulli resistor then gives a pressure $\rho Q/2A_2^2$, where Q is related to the pressure momentum by a relation of the form of Eq. (12.47). Using the notation shown in Figure 12.11 and noting from the bond graph that two state variables—$V(t)$, the volume, and $p_p(t)$, the pressure momentum—are required, we find

$$\dot{V} = -\frac{A_2}{\rho L} p_p, \tag{12.54}$$

$$\dot{p}_p = \frac{-\rho}{2A_2^2}\left(\frac{A_2}{\rho L} p_p\right)^2 + \frac{\rho g V}{A_T}, \tag{12.55}$$

or, combining Eqs. (12.54) and (12.55),

$$-\frac{\rho L}{A_2}\ddot{V} = \frac{-\rho(\dot{V})^2}{2A_2^2} + \frac{\rho g}{A_T} V \tag{12.56}$$

as long as $V \geq 0$. Note that, although no static pressure power flows into the zero-pressure source representing the atmosphere, there is a flow, and with a diffuser one could recover power from this flow.

Control volume problems can become quite complex if the control volume itself is not fixed in inertial space. Bond graph methods can clarify the modeling of such systems, but space limitations prohibit a full discussion here. An example of such a system appears in Reference [6]. We now go on to discuss the implications of dropping the incompressibility assumption, and we find that further difficulties await us and that more approximations are involved in the hydraulic system bond graphs used previously than at first might have been imagined.

12.4.2 Representation of Compressibility Effects in True Bond Graphs

For small changes in density, it is easy to see that compressibility effects are readily modeled by means of linear capacitors, as discussed previously. For large changes in density, however, it is best to begin with a study of the thermodynamics of a pure substance. In this section, we see how the attempt to maintain a true bond graph representation in the case of compressible fluids leads to further complications. This is one motivation for consideration of another form of pseudo-bond graph in Section 12.5. As we have seen in Section 12.2, the internal energy per unit mass of fluid, u, depends on two independent properties: the density ρ (or its reciprocal, specific volume, v) and the entropy s. Here we use $1/\rho$ instead of v to highlight the role of variable density and to avoid confusion with the use of v for velocity in the previous section. Changes in these quantities are related by the Gibbs equation:

$$du = T\,ds - p\,d\left(\frac{1}{\rho}\right), \tag{12.5 Repeated}$$

where T is the thermodynamic temperature, p is the pressure, $1/\rho$ is the specific volume, and

$$u = u\left(s, \frac{1}{\rho}\right). \qquad \text{(12.4) Repeated}$$

The characteristics of the fluid are conveniently summed up by Eq. (12.4). The Gibbs equation for a unit mass or a fixed amount of matter can be represented by a C-field, as we have seen previously.

Using the C-field representation, it is straightforward to model systems in which fixed amounts of a fluid are compressed and expanded and heated. When the fluid passes into and out of a control volume, however, the situation is somewhat more complex. In Figure 12.12, for example, fluid is compressed in a fixed volume by allowing fluid to pass slowly in and out of a port. In this case, the total energy contained in the volume V_0, denoted by U, depends on the mass m contained in the volume and the internal energy per unit mass, u. Changes in U occur not only because of the flow work pQ but also because of the convection of energy. If one uses the mass flow rate $\dot{m} = \rho Q$ instead of Q

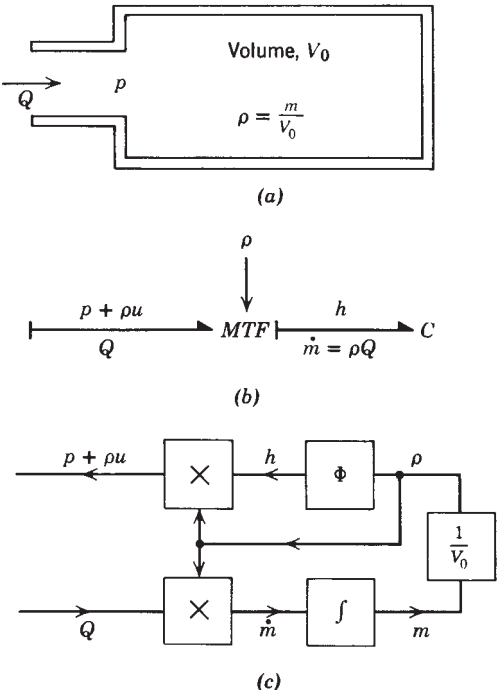


FIGURE 12.12. Isentropic compression of a fluid: (*a*) fluid flow into a rigid volume; (*b*) bond graph; (*c*) block diagram.

as a flow variable, then

$$dU = u\ dm + \frac{p}{\rho} dm \tag{12.57}$$

in the isentropic case. Defining the enthalpy h by the relation

$$h = u + \frac{p}{\rho}, \tag{12.58}$$

it now appears that the power flow past the port is given by

$$h\dot{m} = u\dot{m} + \frac{p}{\rho}\dot{m} = u\rho Q + \frac{p}{\rho}\rho Q = (\rho u + p)Q, \tag{12.59}$$

which shows that pQ is only the hydrostatic power. In this true bond graph model, if the volume flow rate Q is the flow variable, then the corresponding effort is not just the pressure but $\rho u + p = \rho h$, the density times the enthalpy. As in the previous section, in which it was found necessary to supplement the static pressure with a dynamic pressure term to account properly for real power flow, so here it is necessary to supplement p with an extra term, ρu, to account for convected internal energy. Figure 12.12*b* shows a bond graph based on Eqs. (12.58) and (12.59), and a block diagram corresponding to integral causality is given in Figure 12.12*c*. Note that the dynamic pressure term has not been incorporated, so this model is valid only for low flow rates (as indeed is the C-field representation for the Gibbs equation).

Also, note that the pressure and internal energy depend on density alone in the isentropic case, so the enthalpy also is determined by conditions inside the vessel. In more complicated cases, the enthalpy at the port is determined by conditions inside the vessel for outflow, $\dot{m} < 0$, but is determined by external conditions for inflow, $\dot{m} > 0$. This kind of causal switching with changes in direction of flow is not readily apparent in the true bond graph we have just created but will be explicitly considered in the pseudo-bond graphs developed in Section 12.5.

Figure 12.13 shows how the bond graph is modified when isentropic compression occurs partly due to variable volume and partly due to inflow and outflow, as in the power cylinder of a pneumatic servomechanism. The C-field in Figure 12.11*b* has one $h\dot{m}$ port and one $p\dot{V}$ port. The implication is that the total energy U depends on m and V with

$$h = \frac{\partial U}{\partial m}, \quad -p = \frac{\partial U}{\partial V}. \tag{12.60}$$

The bond graph of Figure 12.13 and the equations it represents are a sophisticated version of models of compressibility effects in fluid servomechanisms in which the volume of fluid concerned is variable. In practice, linearized compressibility parameters are commonly incorporated into the system equations to account for compressibility effects. But, as is often the case, it is virtually impossible to proceed from a linearized model of an effect to a more accurate nonlinear model without starting again from the basic physics of the situation.

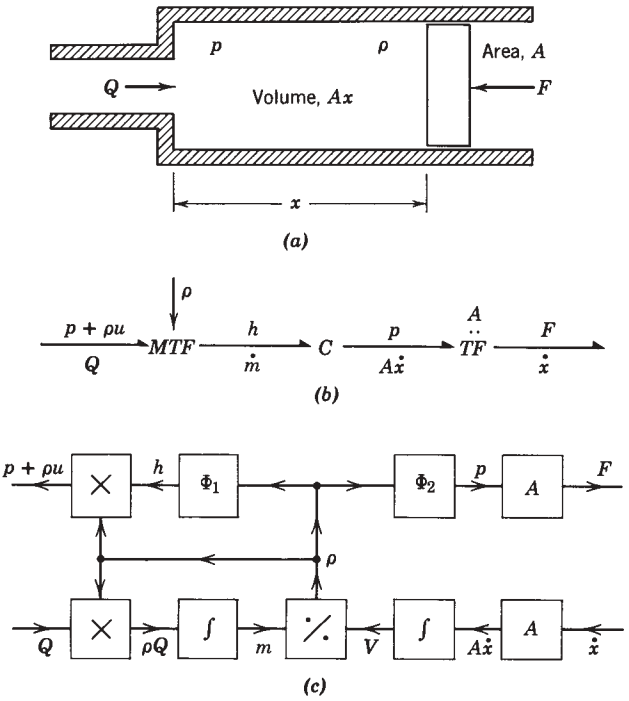


FIGURE 12.13. Compression in a ram: (*a*) sketch of system; (*b*) bond graph; (*c*) block diagram.

12.4.3 Inertial and Compressibility Effects in One-Dimensional Flow

The simple model of Figure 12.8, neglecting for the moment the loss elements, is a lumped approximation of the inertial effects and compressibility effects in a length of line. When a large number of such models with parameters appropriate to a length Δl are cascaded and when $\Delta l \to 0$, the state equations for the model form an approximation to the partial differential equation called the *one-dimensional wave equation*. In the sections above, models for the inertia effect in incompressible flow and the compliance effect when no inertia is considered were developed. The question now arises whether one can simply cascade a series of lumps that alternately account for inertia and compressibility effects in the manner of Figure 12.8 and thereby construct a model of a distributed line.

The answer is in the affirmative in several important cases, namely:

1. *The acoustic approximation.* In this case, small deviations in pressure and density are modeled with linear equations. The inertial and compliance coefficients are calculated at the mean pressure and density. See Reference [5]. Such a model is useful in studying compressibility effects in oil hydraulics and water-hammer problems, as well as in acoustics.

2. *Lagrangian descriptions.* When one is willing to follow the motion of a group of particles, then the inertia of the group is constant. As long as Newtonian mechanics is used, the system of equations can be nonlinear only due to nonlinear elastic or dissipative effects. For the vibration of bars and strings, for example, a Lagrangian description is natural, since the particles never move very far anyway. This was the type of system discussed in Chapter 10. For beams and plates, the differential equations are more complex than the wave equation, but the Lagrangian description still allows one to use finite models with I's and C's representing inertia and compliance effects. The junction structure is more complex than the simple 0–1–0–1 string of Figure 12.8, however. For fluid systems in which the particles move through fixed boundaries, the Lagrangian descriptions are used rather rarely.

For the cases above one may easily construct finite-element models from normal bond graph components that, when reduced to differential volumes, mirror the usual derivations of the partial differential equations of motion for the continuum model. Essentially, the inertial, compliance, and resistance aspects of the system may be treated separately in the finite elements, even though these aspects refer to a single point when the volume of the element is made to approach zero in the continuum description. The Eulerian description common to fluid mechanics entangles the inertial and compliance aspects so thoroughly that it is difficult to construct a series of finite elements converging to the continuum description without using a large number of active bonds, except in the acoustic approximation.

Consider the isentropic, one-dimensional flow of fluid through a tube of unit cross-sectional area. Using s for the space coordinate, $v(s, t)$ for velocity, $\rho(s, t)$ for density, and $p(s, t)$ for pressure, the equations describing the system are

$$\rho \frac{\partial v}{\partial t} + \rho v \frac{\partial v}{\partial s} = -\frac{\partial p}{\partial s}, \tag{12.61}$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial s} = 0, \tag{12.62}$$

$$p = p(\rho). \tag{12.63}$$

Equation (12.61) is just Newton's law for a length of fluid with the acceleration expressed in Eulerian form. Equation (12.62) is a statement of conservation of mass that, with the constitutive law (12.63) for the gas, can be used to define the compressibility effect. The density, $\rho(s, t)$, enters both the inertia law and the compressibility laws. In contrast, a Lagrangian description would express the inertia in terms of a constant mass of fluid in a reference state.

It is probably not sensible to attempt to construct a bond graph for Eqs. (12.61)–(12.63) unless ρ in Eq. (12.61) is nearly constant (in Eqs. (12.62) and (12.63) ρ *must* be allowed to vary) and $v\ \partial v/\partial s$ is a second-order small quantity. Basically, Eq. (12.61) is Newton's law following the flow infinitesimally.

At each succeeding instant of time, the law refers to *different* infinitesimal slices of fluid. Thus, the time-varying inertial effect will not conserve energy as normal bond graph inertial elements do. Also, the $\partial v/\partial s$ term implies that the velocity is different at the two ends of the differential control volume. Thus, the inertial effect and the compressional effect both appear in Eq. (12.61). Finally, the factors appearing in these equations (pressure and velocity for volume flow, since we are considering unit area) do not multiply to give the true power flow at position s and time t. As we have seen, the convected internal energy has been left out of the equations. One should not, therefore, expect that these (correct) equations can be represented by a bond graph in which power conservation and energy conservation are built in. It seems that the alternatives, if one wishes to represent fluid dynamic lines in terms of bond graphs, are to (1) use a Lagrangian description; (2) use an Eulerian description, but with a restriction essentially to the acoustic approximation; and (3) consider only the incompressible or noninertial cases, as was done in Sections 12.4.1 and 12.4.2. Thus, the intuitively appealing scheme of Figure 12.8 appears to be more restrictive than one might think, even without consideration of shear stresses, boundary layers, and so forth, which also complicate the modeling of fluid transmission lines.

In the next section a more practical approach to a number of thermofluid situations of engineering importance is developed using a pseudo-bond graph approach that uses efforts and flow variables commonly used in thermodynamics and fluid dynamics but without requiring the effort flow product to always represent power on a bond.

12.5 PSEUDO-BOND GRAPHS FOR COMPRESSIBLE GAS DYNAMICS

It has previously been seen that constructing bond graph models for fluid dynamic systems is straightforward when the Lagrangian approach of tracking a fixed packet of particles is used. However, the Lagrangian description is often not particularly convenient for computational purposes, and it is common to describe fluid systems with respect to an Eulerian control volume, in which observations are made at a fixed point in space as different fluid particles pass by. Unfortunately, some convective terms appear in the energy and momentum equations due to the fixed reference in the Eulerian approach, and the convective terms do not conveniently lend themselves to a true bond graph representation in terms of power variables.

A major emphasis of this book is the construction of low-order, accurate, understandable models of all types of interacting physical engineering systems. Thermodynamic and nonlinear gas dynamic system parts are another type of interaction we would like to include in our unified approach. After all, the gas dynamics is only part of the physics of an internal combustion engine, a Wankel compressor, a helical rotor expander, or an air cushion vehicle. There are many other dynamic elements that must interact with the gas dynamics to produce an overall system model.

This section develops a pseudo-bond graph representation for gas dynamics. The pseudo-bond graph elements rely on causality in the usual way and connect to other pseudo-bond graph fragments just as true bond graph elements do. Most of the virtues of bond graph modeling are preserved, including the a priori identification of state variables, the ease of deriving state equations, and setting up of a computational model. Pseudo-bond graph elements do not, however, conserve power and energy in the same way as some of the elements of true bond graphs do. Because of this, care must be used when connecting a pseudo-bond graph to a true bond graph to arrive at a complete system model. This turns out to be a small price to pay in exchange for the enormous modeling flexibility the pseudo-bond graph technique affords.

12.5.1 The Thermodynamic Accumulator—A Pseudo-Bond Graph Element

In this section, we define a pseudo-bond graph *accumulator*, which is similar to a true bond graph multiport *capacitor* or *C-field*. It is called an *accumulator* to emphasize the fact that its constitutive laws do not have all the properties connected with energy conservation that C-fields studied in Chapter 7 were found to have. It turns out that the variables chosen for efforts and flows for this pseudo-bond graph multiport are familiar variables commonly used in describing thermofluid systems. These variables allow a simpler description of compressible fluids using a pseudo-bond graph than the true bond graph equivalent.

Consider the control volume in Figure 12.14 with the internal gas at instantaneous pressure p, absolute temperature T, density ρ, and volume V, and containing mass m and energy E. (Note that in the theoretical parts of this section, we will use lower case p to represent pressure to be consistent with previous sections but in several figures, pressure is designated with uppercase P's. The bond graphs used for simulation would avoid this notation difficulty by representing all pressures as efforts, e's with bond numbers.)

The control volume can transport mass through the *in* port and the *out* port, but the inflow and outflow as functions of time can actually be positive or negative at either port. Finally, we can express work done by the gas in the control volume by the volume expansion indicated in the figure. Notice that we assume that a single pressure and temperature characterize the entire control volume. The mass, energy, and (variable) volume can be used to find variables such as density or specific volume and internal energy.

The one-dimensional energy, mass, and momentum equations can be written in the following forms (the subscripts i and o refer to the *in* and *out* ports, respectively):

$$\frac{d}{dt}E = \left(h_i + \frac{v_i^2}{2}\right)\dot{m}_i - \left(h_0 + \frac{v_o^2}{2}\right)\dot{m}_o - p\frac{dV}{dt} \quad \text{(energy)}, \tag{12.64}$$

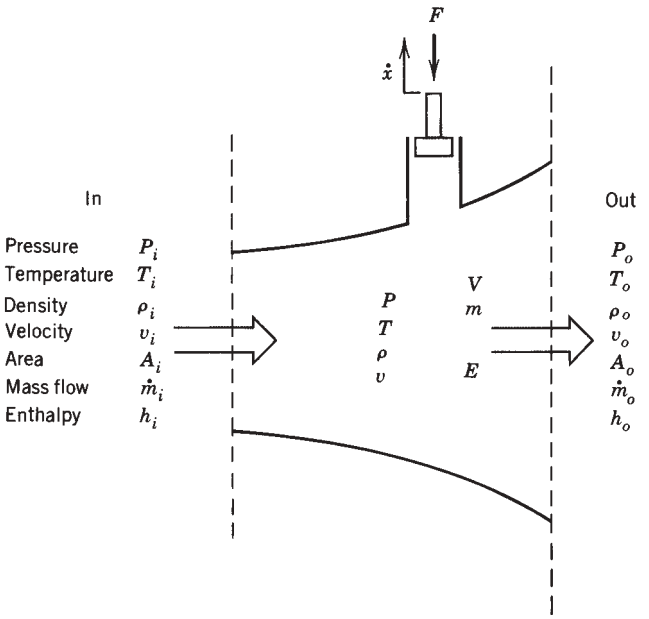


FIGURE 12.14. Control volume for gas dynamics.

$$\frac{d}{dt}m = \dot{m}_i - \dot{m}_o \quad \text{(mass)}, \tag{12.65}$$

$$\frac{d}{dt}(mv) = v_i\dot{m}_i - v_o\dot{m}_o + p_iA_i - p_oA_o + \frac{1}{2}(p_i + p_o)(A_o - A_i) + \mathcal{R} \quad \text{(momentum)}, \tag{12.66}$$

where $\mathcal{R}$ is the resultant external force acting on the control volume.

The internal energy contained in the control volume E is

$$E = mc_vT, \tag{12.67}$$

and the enthalpy, h, is given by

$$h_i = c_pT_i, \quad h_o = c_pT_o. \tag{12.68}$$

See Eqs. (12.27) and (12.28) for definitions of the specific heats.

We assume that the gas obeys the gas law for mass m, see Eq. (12.22),

$$pV = mRT. \tag{12.69}$$

For example, for air, the gas parameters are

$$\begin{aligned} R &\equiv \text{gas constant} \\ &= 287 \text{ N-m/kg-K}, \\ c_p &\equiv \text{specific heat at constant pressure} \\ &= 1005 \text{ N-m/kg-K}, \\ c_v &\equiv \text{specific heat at constant volume} \\ &= 718 \text{ N-m/kg-K}. \end{aligned} \tag{12.70}$$

Also (from Eq. (12.29),

$$R = c_p - c_v, \tag{12.71}$$

and the specific heat ratio γ is

$$\gamma = \frac{c_p}{c_v} = 1.4. \tag{12.72}$$

For now, we will not consider momentum variation within the control volume, and we will neglect the convected kinetic energy in the energy equation. This kinetic energy is typically small compared to the convected enthalpy, as becomes apparent if we consider that at room temperature of 20°C, or 293 K, the enthalpy is $h = c_p$ (293 K) $= 2.94 \times 10^5$ N-m/kg, while for a gas velocity of 100 m/sec, the unit mass kinetic energy contribution is only $v^2/2 = 5 \times 10^3$ N-m/kg.

We will work with the equations

$$\frac{d}{dt}E = h_i\dot{m}_i - h_o\dot{m}_o - p\frac{dV}{dt}, \tag{12.73}$$

$$\frac{d}{dt}m = \dot{m}_i - \dot{m}_o, \tag{12.74}$$

$$\frac{d}{dt}V = \dot{V}, \tag{12.75}$$

$$E = mc_vT, \tag{12.76}$$

$$pV = mRT. \tag{12.77}$$

Equations (12.73), (12.74), and (12.75) look like first-order state equations, and they motivate the construction of the bond graph of Figure 12.15. The 3-port C-field in this figure has one true bond and two pseudo-bonds. The true bond has pressure p as its effort variable and volume rate $\dot{V}$ as its flow variable. The effort–flow product, $p\dot{V}$, represents power on this true bond. These variables have been used previously for lumped modeling of incompressible fluid systems and appeared on the true bond graph C-field of Section 12.2.

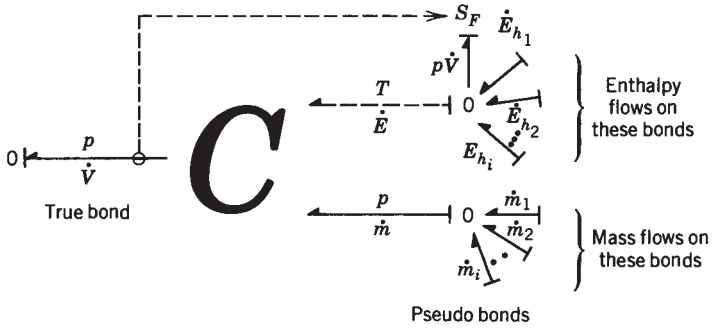


FIGURE 12.15. Bond graph of the thermodynamic accumulator.

One of the pseudo-bonds has energy flow $\dot{E}$ as the flow variable and temperature T as the effort. This effort flow variable pair is similar to that used in heat transfer pseudo-bond graphs in Section 12.1. You will note that the $T\dot{E}$ pseudo-bond has been shown as a *dashed line*. This has no particular significance except to distinguish it from the $p\dot{m}$ pseudo-bond to make understanding the graph somewhat easier.

The other pseudo-bond again uses pressure p as the effort just as the true bond does, but mass flow $\dot{m}$ as the flow in place of volume flow rate as in true bond graphs. (Note that $\dot{m}$ appeared in some of the true bond graphs in Section 12.3.2 but it was paired with enthalpy as an effort rather than pressure.) For both pseudo-bonds, the product of their effort and flow variables is *not* power. In fact, $\dot{E}$ is already power on the dashed upper pseudo-bond.

Even though the accumulator is not a true bond graph C-element, nevertheless, the operations associated with the effort and flow variables are the same as for a true C-field. As the causality indicates in Figure 12.15, the accumulator is in all integral causality and accepts flow inputs on all three bonds ($\dot{E}$, $\dot{m}$, and $\dot{V}$). It then integrates these flows to produce the state variables E, m, and V. And, finally, the accumulator operates on these state variables through appropriate constitutive laws to produce the outputs p and T. The constitutive relations come from Eqs. (12.76) and (12.77) and are

$$T = \frac{1}{c_v}\frac{E}{m} \tag{12.78}$$

and

$$p = \frac{mRT}{V} = \frac{mR}{V}\frac{1}{c_v}\frac{E}{m} = \frac{R}{c_v}\left(\frac{E}{V}\right). \tag{12.79}$$

Thus, if $\dot{E}$, $\dot{m}$, and $\dot{V}$ are inputs to the accumulator, then it will output T and p using the constitutive laws (12.78) and (12.79). Thus, the accumulator pseudo-bond graph element makes compressible gas dynamics (subject to the operational

assumptions that got us to this point) fit into our general unified approach for modeling systems.

An unusual feature is the flow source, SF, associated with the $T\dot{E}$ bond. It is necessary to properly account for the work done by the fluid in our control volume as the volume changes. The work removes some energy from the fluid in the control volume according to Eq. 12.73. The power removed from the fluid is $\dot{E} = p\dot{V}$, which accounts for the half-arrow sign convention shown. When the volume increases, energy is removed from the gas.

The remaining external bonds associated with the T, $\dot{E}$ bond tell the accumulator the power flows that are associated with the product of enthalpy and mass flow rate $\dot{E}_{h_i} = h_i\dot{m}_i$, from whatever systems are attached to the accumulator. The 0-junction shows that all the attached systems are subjected to the temperature of the control volume as determined by the accumulator.

Similarly, the $p\dot{m}$ pseudo-bond has external bonds that causally prescribe the mass flows from the attached systems. Another 0-junction indicates that the control volume pressure is an input to each of the attached systems.

As a next step, we need to describe the origin of the transported energy flows (enthalpy flows)

$$\dot{E}_h = \sum_i \dot{E}_{h_i} = \sum_i h_i\dot{m}_i = \sum_i c_p T_i \dot{m}_i \tag{12.80}$$

and the mass flows $\dot{m}_i$, where

$$\dot{m} = \sum_i \dot{m}_i, \tag{12.81}$$

but first we demonstrate the use of the thermodynamic accumulator for a simple case in which the mass flows are zero so that Eqs. (12.80 and 12.81) are not necesary.

Figure 12.16 shows a cylinder with a trapped mass of air under a moving piston with prescribed motion, $v_i(t)$. Without making any assumptions about the thermodynamic process, we now show that a model using the thermodynamic accumulator predicts a reversible adiabatic or isentropic process. The bond graph for this simple closed system is also shown in Figure 12.16.

The state equations are

$$\begin{aligned} \dot{E} &= -p\dot{V}, \\ \dot{V} &= A_p v_i \quad \text{(prescribed)}, \\ \dot{m} &= 0 \quad \text{(closed system)}. \end{aligned} \tag{12.82}$$

From the constitutive law of Eq. (12.76) we can write

$$\dot{E} = -\frac{R}{c_v} E \frac{\dot{V}}{V}. \tag{12.83}$$

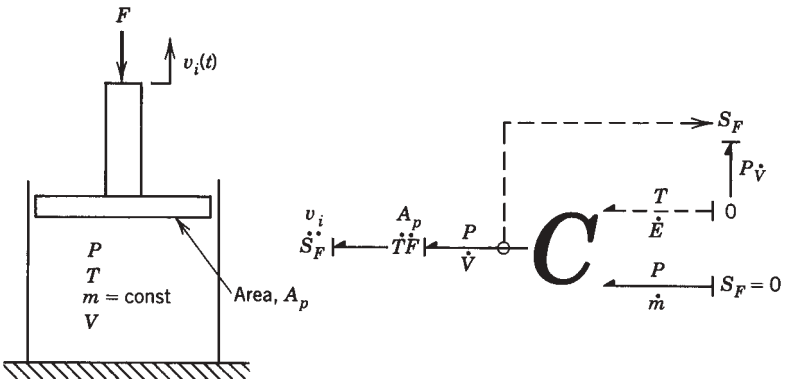


FIGURE 12.16. Nonlinear air spring and its bond graph model.

If we were pursuing a numerical solution, we would numerically integrate this equation along with the equation for $\dot{V}$ and then output pressure and temperature from (12.78) and (12.79). Instead, we recognize that (12.83) has an analytical solution,

$$EV^{R/c_v} = \text{const.} \tag{12.84}$$

If we reintroduce the pressure from Eq. (12.79), then

$$\frac{c_v}{R} pV^{(R/c_v+1)} = \text{const}, \tag{12.85}$$

or

$$pV^{\gamma} = \text{const}, \tag{12.86}$$

using Eqs. (12.71 and 12.72)

$$\gamma = \frac{c_p}{c_v} = \frac{R}{c_v} + 1. \tag{12.87}$$

The reader should recognize that (12.86) is the *p, V* relationship for an isentropic process.

12.5.2 The Thermodynamic Restrictor—A Pseudo-Bond Graph Element

The next pseudo-bond graph element necessary to model compressible gas dynamics is the *restrictor*, which resembles the multiport *resistors* or *R-fields* studied in Chapter 7. Again, the special name *restrictor* for this element is chosen to indicate that the effort and flow variables chosen for the pseudo-bond graph do not allow all the power dissipation interpretations that were possible for true bond graph elements. To show the type of constitutive laws defining a restrictor and to illustrate the utility of considering the use of effort and flow

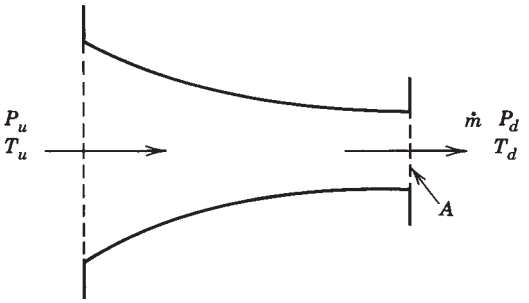


FIGURE 12.17. Isentropic nozzle.

variables that are not true bond graph variables, we discuss a very useful model pertaining to flows in gas dynamics—the isentropic nozzle.

Consider the energy flows $\dot{E}_{h_i}$ and mass flows $\dot{m}_i$ required as inputs to the thermodynamic accumulator. Figure 12.17 shows a nozzle with upstream pressure and temperature p_u, T_u; downstream pressure and temperature p_d, T_d; exit area A; and mass flow $\dot{m}$. If we assume that isentropic flow exists, then it turns out that $\dot{m}$ depends on the pressure ratio

$$p_r \equiv \frac{p_d}{p_u} \tag{12.88}$$

rather than just the pressure drop $p_u - p_d$ across the nozzle. A standard derivation (see Reference [7] for example) shows that $\dot{m}$ is given by the expression

$$\dot{m} = A\frac{p_u}{\sqrt{T_u}}\sqrt{\frac{2\gamma}{R(\gamma-1)}}\sqrt{p_r^{2/\gamma} - p_r^{(\gamma+1)/\gamma}}, \tag{12.89}$$

as long as the pressure ratio p_r is larger than the critical pressure ratio as defined by

$$p_{r\,\text{crit}} \equiv \left(\frac{2}{\gamma+1}\right)^{\gamma/(\gamma-1)}. \tag{12.90}$$

When the pressure ratio is less than the critical ratio, the flow is "choked," and $\dot{m}$ remains constant, even if the downstream pressure, p_d, is reduced.

The transported energy $\dot{E}_h$ associated with this mass flow $\dot{m}$ is

$$\dot{E}_h = c_p T_u \dot{m}. \tag{12.91}$$

We will consider the mass flows and energy transport flows into and out of the control volumes represented by the thermodynamic accumulator as isentropic nozzle flows given by Eqs. (12.89) and (12.91). Which end of the nozzle is considered the upstream side of the nozzle as well as the sign of $\dot{m}$ are dictated by the sign of the pressure drop Δp across the nozzle. If Δp changes sign at

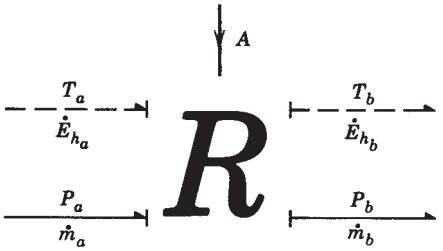


FIGURE 12.18. Bond graph for the isentropic nozzle.

some time, then the sign of $\dot{m}$ will change, as will our identification of the nozzle ends considered upsteam and downstream in the formulas.

Consider the 4-port R-element now called a pseudo-bond graph restrictor of Figure 12.18. As before the $T\dot{E}$ bonds are shown dashed to distinguish them from the $p\dot{m}$ bonds. Note that all bonds have *effort in* causality. This causality is the one most useful for the restrictor. A parallel causal restriction was found for the true bond graph of heat transfer studied in Section 12.3 and represented in Eq. (12.38).

Although these are all pseudo-bonds, the 4-port restrictor is functionally equivalent to an R-field in true bond graphs. This R-element operates on the input efforts p_a, p_b, T_a, T_b and delivers the output flows $\dot{m}_a$, $\dot{m}_b$, $\dot{E}_{h_a}$, $\dot{E}_{h_b}$. This can be expressed functionally as

$$\begin{aligned} \dot{m}_a &= \dot{m}_a(p_a, p_b, T_a, T_b), \\ \dot{m}_b &= \dot{m}_b(p_a, p_b, T_a, T_b), \\ \dot{E}_{h_a} &= \dot{E}_{h_a}(p_a, p_b, T_a, T_b), \\ \dot{E}_{h_b} &= \dot{E}_{h_b}(p_a, p_b, T_a, T_b). \end{aligned} \tag{12.92}$$

We now define the computational procedures that allow the restrictor of Figure 12.18 to represent the isentropic nozzle. First, the designation of upstream and downstream depends on the pressures at the two ends of the restrictor.

$$\begin{aligned} &\text{if} \quad p_a > p_b, \qquad \text{then} \quad p_u = p_a, \quad T_u = T_a, \quad p_d = p_b; \\ &\text{if} \quad p_a < p_b, \qquad \text{then} \quad p_u = p_b, \quad T_u = T_b, \quad p_d = p_a; \end{aligned} \tag{12.93}$$

and

$$p_r = p_d / p_u. \tag{12.94}$$

Furthermore,

$$\begin{aligned} &\text{if} \quad p_r > p_{r\ \text{crit}}, \qquad \text{then} \quad p_r = p_d / p_u; \\ &\text{if} \quad p_r \leq p_{r\ \text{crit}}, \qquad \text{then} \quad p_r = p_{r\ \text{crit}}. \end{aligned} \tag{12.95}$$

Now calculate

$$\dot{m} = A\frac{p_u}{\sqrt{T_u}}\sqrt{\frac{2\gamma}{R(\gamma-1)}}\sqrt{p_r^{2/\gamma} - p_r^{(\gamma+1)/\gamma}}, \tag{12.96}$$

and obtain outputs

$$\begin{aligned} &\text{if} \quad p_a > p_b, \quad \text{then} \quad \dot{m}_a = \dot{m}_b = \dot{m}, \\ &\text{if} \quad p_a < p_b, \quad \text{then} \quad \dot{m}_a = \dot{m}_b = -\dot{m}, \end{aligned} \tag{12.97}$$

in either case,

$$\dot{E}_{h_a} = \dot{E}_{h_b} = c_p T_u \dot{m}_a. \tag{12.98}$$

The relationships (12.93)–(12.98) fully represent the input–output behavior of the isentropic nozzle. And as long as we maintain the causality for which these relationships were derived, we can construct this computational procedure once and for all, to be used over and over when needed in a system model.

12.5.3 Constructing Models with Accumulators and Restrictors

In this section we show how pseudo-bond graph accumulators and restrictors can be connected together and to true to true bond graph models to create models for complex systems in which thermofluid elements play a role in the overall system dynamics.

Figure 12.19 shows a cylinder with gas trapped above and below the piston. If no leakage exists between top and bottom chambers, then the model is simply two independent accumulators identical to the one in Figure 12.16. When leakage past the piston is included, the model is shown in Figure 12.19. Notice how the 4-port restrictor interacts with the pseudo-bonds of the accumulators. The accumulator outputs, P_1, T_1, P_2, T_2 are the inputs to the restrictor, which, through the computational procedure (12.93)–(12.98), outputs the mass and transported energy flows for use as inputs to the accumulators. On the true bond side of the accumulators, all is normal, in that the pressure outputs are converted by transformers into forces on the piston, and the piston inertia, in turn, outputs the velocity $\dot{x}$, which then is converted by the same transformers into volume velocities for the two chambers.

Figure 12.19 demonstrates a very elegant coupling of mechanical and thermodynamic energy domains. Since the accumulators are in integral causality, setting up state equations for computational solution is straightforward.

Pseudo-bond graphs have been used to model many types of systems, such as internal combustion engines. A model of a multicylinder engine can be mechanically complex and can have a number of control volumes, some with fixed volumes

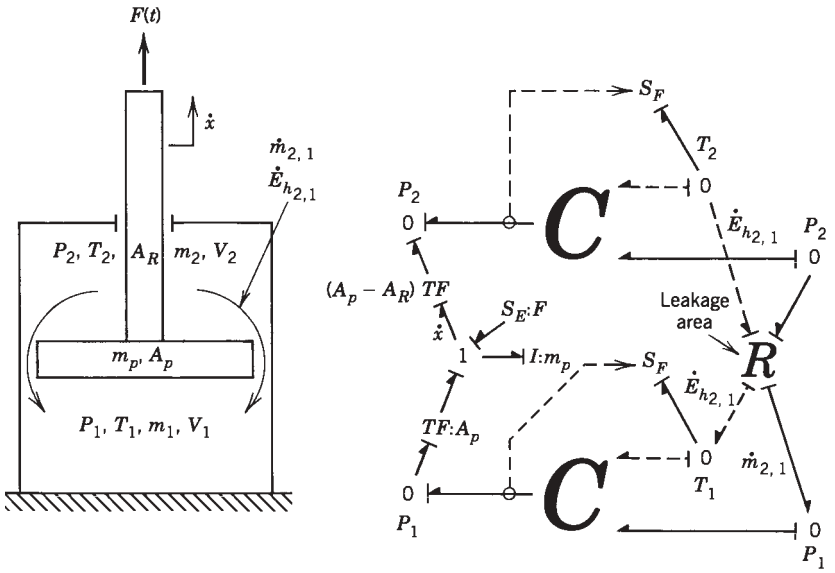


FIGURE 12.19. Two-sided air cylinder with leakage.

and some with variable volumes. The result is often a model with hundreds of state equations. On the other hand, a bond graph model can be constructed of standard true and pseudo-bond graph elements that are repeated many times. Using the bond graph, one can easily keep track of which variables are which in a simulation, and a bond graph processor can help set up the element equations in the correct causal form. Instead of using just equations, this process can be characterized as graph- and computer-aided modeling.

As a relatively simple example, Figure 12.20 shows a two-stroke, piston-ported internal combustion engine and its bond graph model. This model has been used to aid in the understanding of the interacting dynamics of two-stroke engines, and it is shown here as example of using the 3-port accumulators and 4-port restrictors to construct rather complex thermodynamic models. Readers should see Reference [8] for more details of this model.

The two-stroke engine relies heavily on fluid dynamics for its operation. As the piston is driven downward from the pressure due to combustion of an air–fuel charge, it first opens the exhaust port as the piston moves past the exhaust opening. As exhaust gases escape into the exhaust system (not shown), the piston uncovers the transfer port, and the compressed air–fuel in the crankcase moves through the transfer passage to supply the head with the next charge for combustion. As the piston moves upward after passing through bottom dead center, the intake port is opened, and fresh air–fuel enters the expanding crankcase, ready for the next crankcase compression. The air–fuel in the head is compressed and ignited, initiating another cycle.

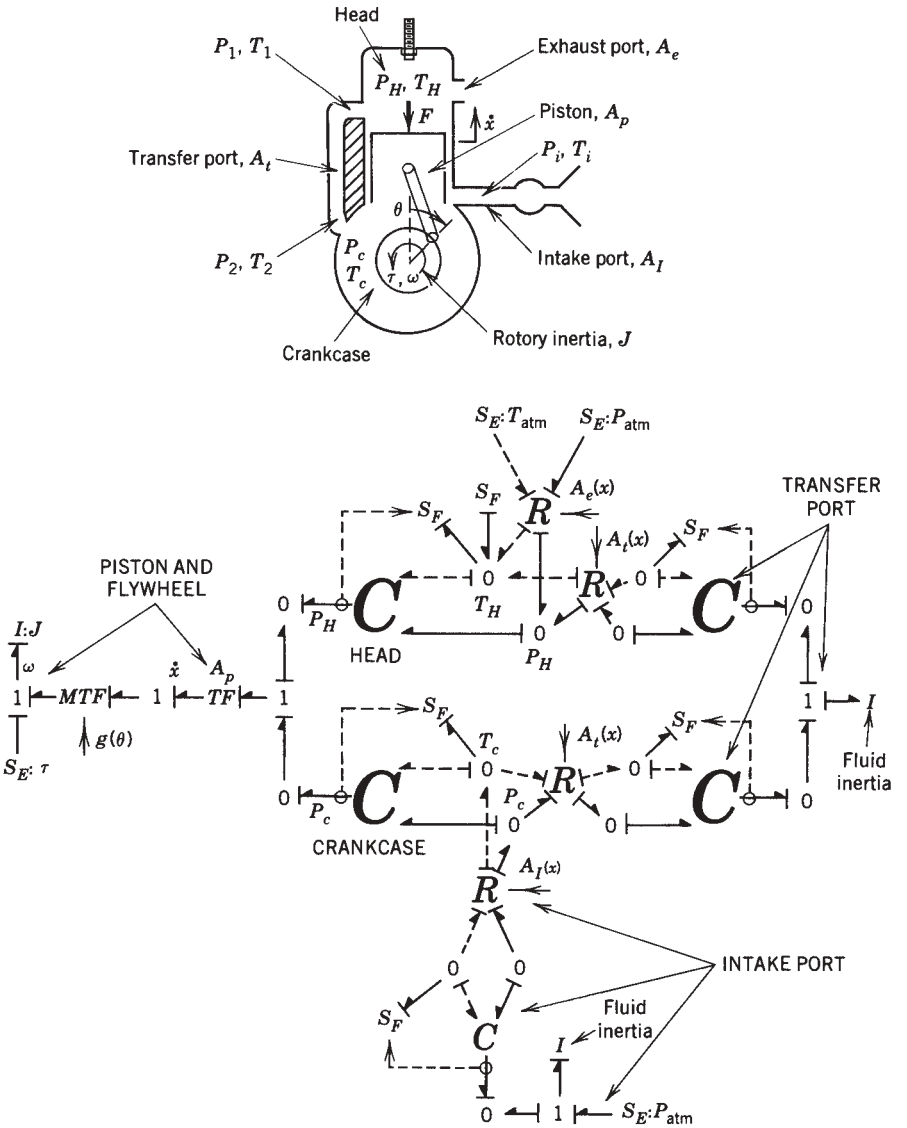


FIGURE 12.20. Bond graph for a two-stroke, piston-ported engine.

The inertia of the fluid in the intake and transfer passages, as well as the dynamics of the exhaust system, is essential to the high-performance operation of the engine. Once fluid is moving in the intake and transfer passages, its momentum continues to charge the crankcase and cylinder head against adverse pressure gradients even as the piston is closing the ports. It is this dynamic

behavior that makes the engine so interesting and dynamic modeling so essential for design purposes.

The inertia effects of the transfer and intake are handled in a functional manner, and the transported momentum from Eq. (12.66) is neglected. One can, in fact, construct a pseudo-bond graph for the momentum equation, but this is beyond the scope of the present discussion. For the model shown here, it is assumed that the fluid inertia of the intake and transfer passages behaves like a trapped mass of fluid in the passages. The density of the fluid is set by the accumulators at the ends of the passage, depending on the instantaneous flow direction. This has been shown experimentally to work well. The complete momentum equation must be used for the exhaust system.

Notice that integral causality exists throughout Figure 12.20; thus state-variable selection, equation formulation, and computer coding will be straightforward.

As a final example, consider Figure 12.21, which shows a Wankel compressor and its bond graph. For someone who has not studied the ideas of the accumulator and the restrictor, this may be a frighteningly complex figure. However, each element of the figure is itself relatively simple and it is just a matter of identifying control volume and flow path to understand the model in detail. The reader is urged to study this model and appreciate the elegance and utility of the representation.

12.5.4 Summary

In order to include compressible gas dynamic elements in an overall system model in a computationally convenient form, it was necessary to introduce the thermodynamic accumulator and associated pseudo-bonds. The 3-port accumulator possesses almost all of the characteristics of a true representation, and through causal considerations, the accumulator is a great aid in putting models together.

It is true that the true bond on the mechanical port requires an active bond and source to complete the energy balance in Eq. (12.73) as shown in Fig. 12.15. This seems a small price to pay for the ability to use of variables commonly used in thermofluid system analysis even though they are not classical power effort and flow variables of true bond graphs. The mechanical port on the accumulator provides the connection between the pseudo-bond graph and true bond graphs of any other mechatronic systems that form a model of a complete system.

The isentropic nozzle and its 4-port restrictor representation are required to interact with the accumulator to account for transported energy. As long as the causality presented in this section is not violated, then including even complex, nonlinear thermodynamics and gas dynamics in overall system models follows the same process as constructing any bond graph model for interactive components crossing many energy domains. In essence, the pseudo-bond graphs for thermofluid systems sacrifice some of the elegance of true bond graphs for practicality.

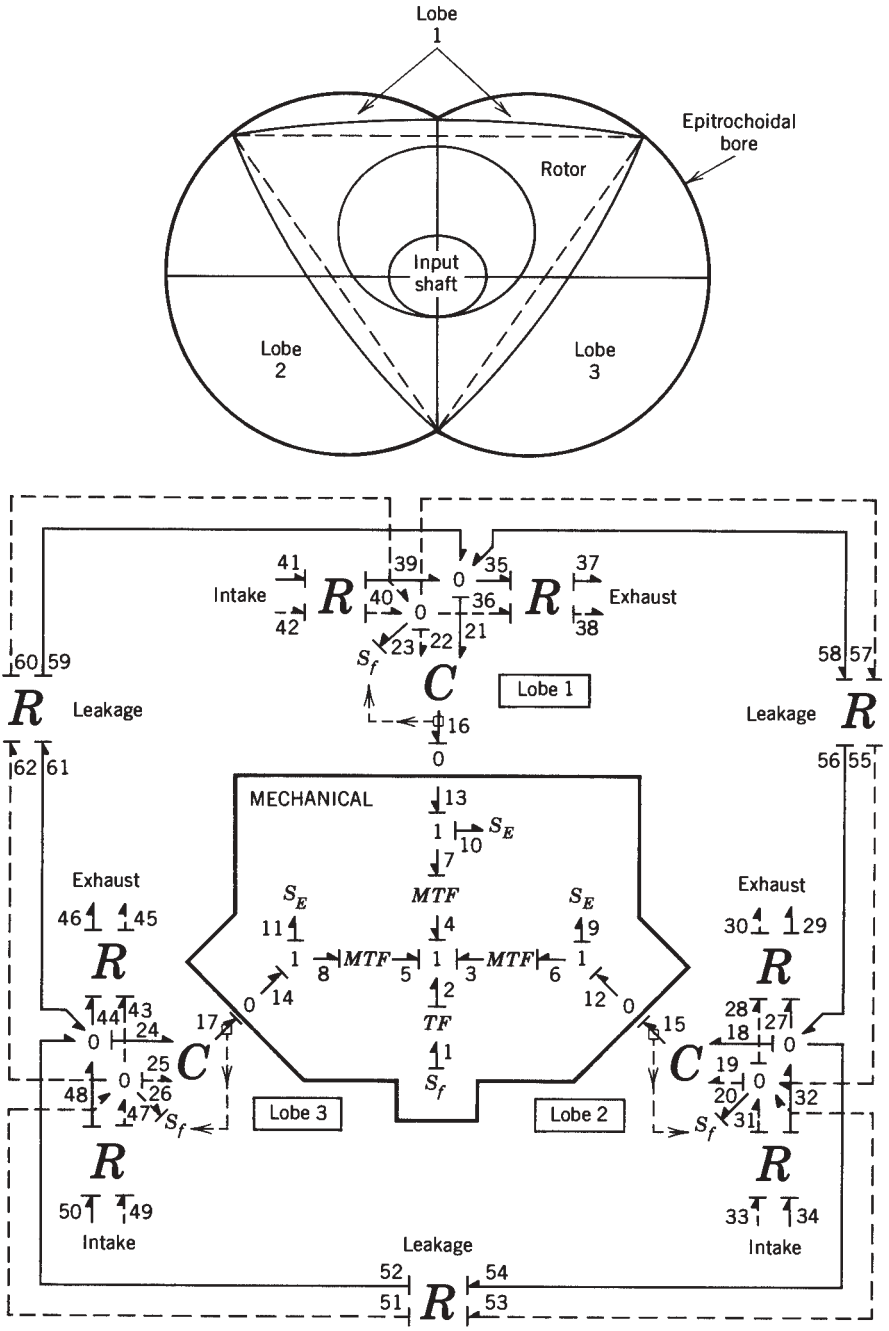


FIGURE 12.21. Wankel compressor and its bond graph.

REFERENCES

[1] R. H. Cannon, *Dynamics of Physical Systems*, New York: McGraw-Hill, 1967.

[2] W. M. Swanson, *Fluid Mechanics*, New York: Holt, Rinehart & Winston, 1970.

[3] C. J. Radcliffe and D. Karnopp, "Simulation of nonlinear air cushion vehicle dynamics using bond graph techniques," *Proceedings 1971 Summar Computer Simulation Conference*, Boston, MA, July 19–21, 1971.

[4] L. Tisza, *Generalized Thermodynamics*, Cambridge, MA: MIT Press, 1966.

[5] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, New York: McGraw-Hill, 1968.

[6] D. C. Karnopp, "Bond graph models for fluid dynamics systems," *Trans. ASME J. Dyn. Syst. Meas. Control*, **94**, Ser. G, No. 3, 222–229 (Sept. 1973).

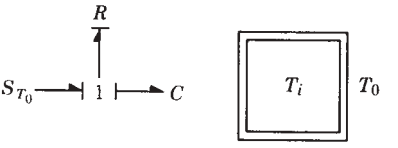
[7] R. H. Sabersky, A. J. Acosta, and E. Hauptmann, *Fluid Flow*, New York: Macmillan, 1971.

[8] D. L. Margolis, "Modeling of two-stroke ICE dynamics using the bond graph technique," *SAE Trans.*, pp. 2263–2275 (Sept. 1975).

[9] D. L. Margolis, "Bond graph fluid line models for inclusion with dynamic systems simulation," *J. Franklin Inst.*, **308**, No. 3, 255–268 (Sept. 1979).

PROBLEMS

12-1. A block of material with total heat capacitance C is covered with insulating material with total thermal resistance R and surrounded by an atmosphere of temperature T_0. The bond graph is supposed to aid in estimating how fast the block will heat up if it starts at some temperature $T_i(t_0) < T_0$. Draw a block diagram from the bond graph.

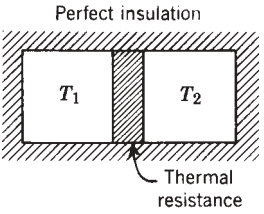


12-2. Consider a given mass of water as a thermal capacitor. Sketch the constitutive law for this element for a range of temperatures including the freezing point. Indicate the effect of the latent heat of freezing or melting.

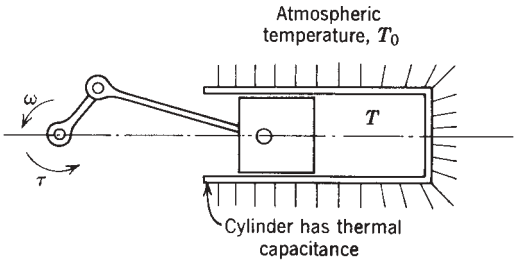
12-3. A heat transfer coefficient α for a surface is given in W/m^2-$^\circ$C. If an area A of this surface is involved in a problem, show how the effect of the

surface would be represented in a bond graph, and write the constitutive law for the element in terms of A and α.

12-4. Two bodies that do not expand or contract are in thermal contact through a thermal resistance. Make two bond graphs for the system using both $T\dot{Q}$ and $T\dot{S}$ variables. Explain the different relations for the —R— and —C elements when the two different variable sets are used. Using the $T\dot{S}$ bond graph and Figure 12.1, show that if $T_2 \neq T_1$ initially, the entropy of the system can only increase.

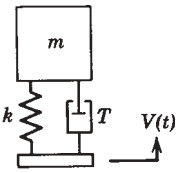


12-5. Consider the compression of a gas by means of a crank–piston arrangement. Let the cylinder have a single average temperature, and define thermal resistances between the gas temperature T and the cylinder temperature and between the cylinder temperature and the atmospheric temperature T_0. Make a bond graph that would allow you to predict the crank torque τ for low speeds of rotation, ω. (Note that you do not have enough information to evaluate all the system parameters.)

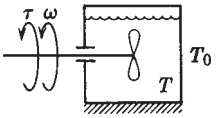


12-6. The dashpot in the suspension system has a force–velocity constitutive law that varies with the average temperature T of the dashpot, since it contains oil that changes viscosity with temperature. Make a simple model of the system that would predict how the dashpot heats up when the input base velocity $V(t)$ is given. Discuss your assumptions and how

you might estimate the system parameters you need for the thermal part of your model.



12-7. Suppose a shaft is connected to a paint stirrer such that, for low enough angular rates ω, all the power $\tau\omega$ goes to heating up the paint. Make a bond graph relating the mechanical power to an entropy flow $\dot{S}$ and including possible heat transfer to the atmosphere. Discuss the simplifying assumptions you have used.



12-8. In the acoustic approximation, the bulk modulus B is given by

$$B = \rho_0 c^2,$$

where ρ_0 is the mean mass density of the fluid and c is the speed of sound. If Δp represents a small increase in pressure over the mean pressure p_0, then

$$\Delta p = \rho_0 c^2 \frac{\Delta\rho}{\rho_0},$$

where $\Delta\rho$ represents the change in density.

(a) Considering a fixed mass of fluid that occupies volume V when the pressure is p_0, show that

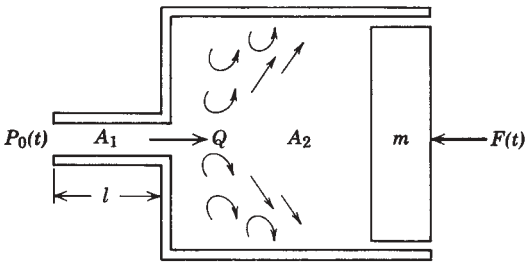
$$\frac{\Delta\rho}{\rho} = \frac{\Delta V}{V}, \qquad \Delta p = -B\frac{\Delta V}{V},$$

where $-\Delta V$ represents a *decrease* in volume.

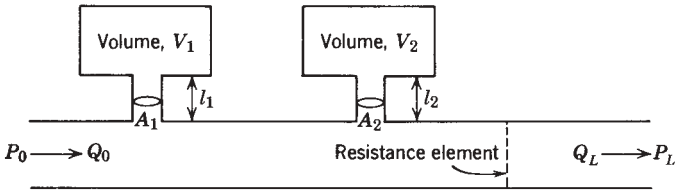
(b) Evaluate the inertia and capacitance parameters for the length of pipe shown in Figure 12.8.

(c) Using $\lambda = c/f$, where λ is wavelength, f is frequency, and c is sound speed, relate the length l for the pipe segment to the highest circular frequency ω of interest, so that even the shortest wavelength will span several "lumps" if one uses Figure 12.8 to make a model of a long pipe by cascading many segments.

12-9. A high-speed hydraulic ram is forced by a pressure source, and we desire to predict how fast it can be stroked. Let the inlet pipe have length l and area A, and consider only inflow Q. The ram has mass m and area A_2 and a force $F(t)$ applied to it. To be conservative, assume that only the static pressure in the ram acts on the piston, that is, all the dynamic pressure is assumed to be lost. Make a bond graph for this system using a Bernoulli resistor to model the dynamic pressure loss. Apply causality and write equations of motion for the system.



12-10. A schematic diagram of a muffler system employing two Helmholtz resonators and a resistive element is shown.

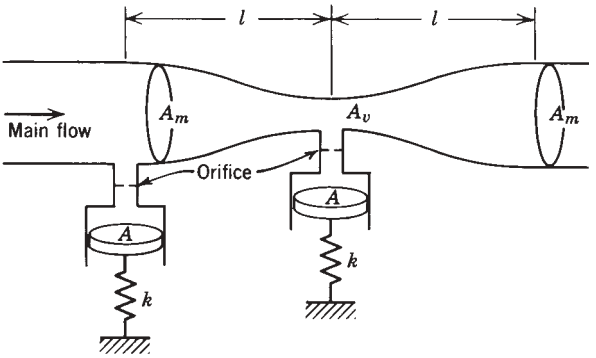


Assume all dimensions are less than one wavelength of the highest frequency of interest. Make a bond graph for the system in which the inertias of the necks of the resonators of effective lengths l_1, l_2, and areas A_1, A_2, the capacitances of the volumes V_1, V_2, and the resistance R are all represented. Using the results of Problem 12-5, list the capacitance and inertia parameters in terms of the density ρ_0 and the sound speed c.

12-11. Part of an automatic flow-metering system is shown. The main flow goes through a smooth venturi that reduces the area from A_m to A_v and in which we may assume that all dynamic pressure is fully recovered. The spring-mounted pistons deflect due to the pressure in the main pipe and at the throat. Orifices in the pipes connecting the piston chambers to the main pipe restrict the flows to the pistons to low values.

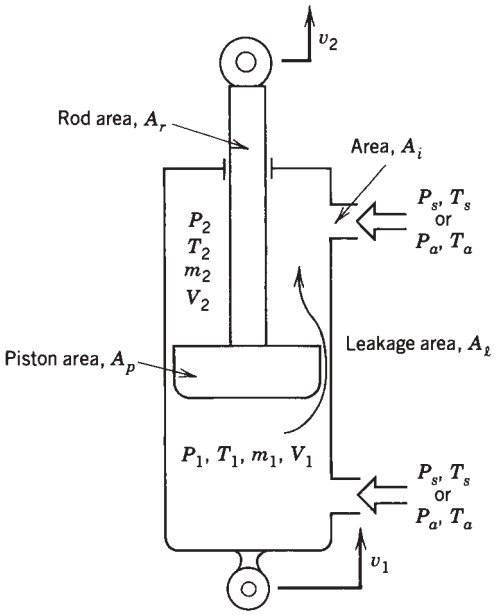
Construct a bond graph for this subsystem, and indicate inertia elements and Bernoulli resistors for the two area change sections of length l. Why

should we expect the two pistons to deflect differing amounts? Would the difference between the two deflections serve to measure the main flow?

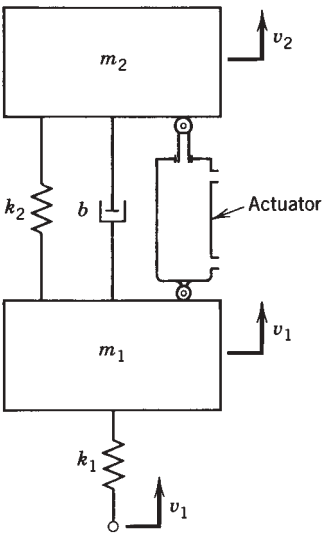


12-12. The figure shows an air cylinder that could be used as an actuator in a control system. The inlet ports at the top and bottom can be exposed to a supply at pressure P_s and temperature T_s or to exhaust at P_a and T_a. This switching would be accomplished through some valving, which is not shown.

Construct a bond graph model of this component that is ready for installation into a system. Assign causality so we know what causality is acceptable from any attached system.

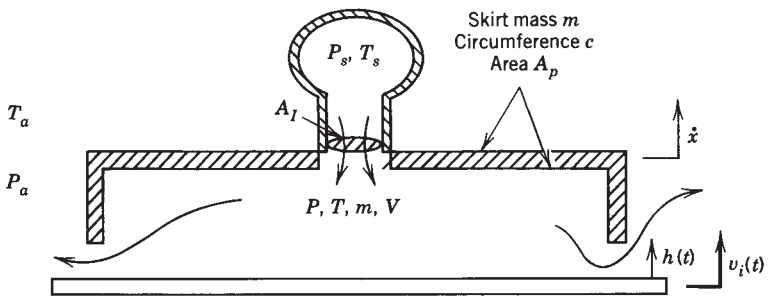


12-13. Install the actuator of problem 12.12 in the quarter-car model shown. Using words and equations, set up the state equations so that you understand how a computational model will result.



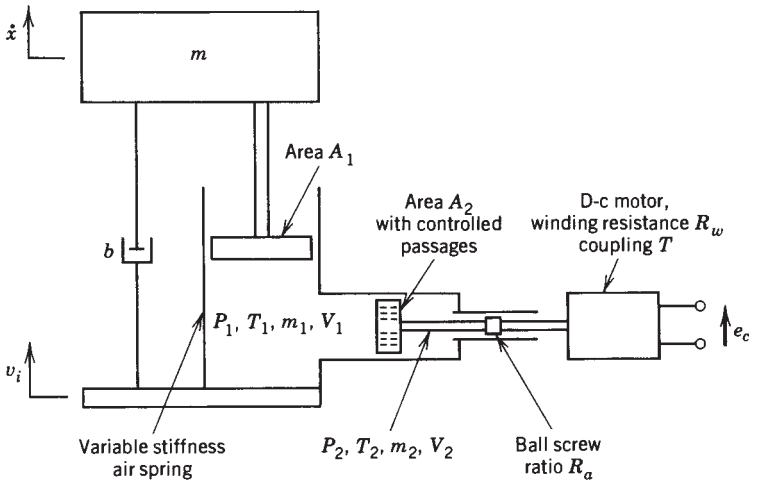
12-14. The figure shows the skirt of an air cushion vehicle. A supply at pressure P_s and temperature T_s fills the skirt volume, so that the vehicle "floats" on a cushion of air, while air escapes under the skirt to atmosphere. The height of the skirt above ground is called h and is time varying as the vehicle negotiates an uneven terrain. The ground input is the velocity $v_i(t)$

Construct a model of this system that, if solved, would predict the vertical motion time history of the vehicle. Since $h(t)$ is very small, it is all right to consider the relative motion between base and skirt as creating a volume change in the skirt.



12-15. The figure shows a schematic of a proposed variable-rate air spring capable of continuous variation. It is installed as a suspension element in the figure. It consists of a primary air cylinder coupled to a secondary cylinder with a motion-driven piston. If the piston is moved forward and mass is transferred from the front chamber to the rear chamber, then the vehicle mass will be suspended on a stiffer air spring, and the ride height will not have changed.

Construct a bond graph model of this system, including all the dynamics suggested in the figure. Assign causality to show that formulation will be straightforward. This model could be used to test control strategies for an adaptive suspension.



12-16. Consider the system from Problem 12-5. Neglect thermal capacitance of the piston walls but include leakage past the piston to the atmosphere. Using one thermodynamic accumulator and one 4-port isentropic nozzle, construct a bond graph model of the system. The crank radius is R, crank length is L, piston area is A_p, leakage area is A_L, and piston mass is m.

12-17. Modify the model from problem 12-16 to include the thermal capacitance of the cylinder as well as convection from the cylinder surface to the atmosphere. The heat transfer coefficient for the cylinder inside wall is h_1, and for the outside wall it is h_2. The thermal capacitance of the wall is C_W.

12-18. The system from Figure 12.19 is an interesting one. Since work is done from the external port to drive gas through leakage area A_L, the gas must continuously get hotter to account for this input energy. If there were no leakage, then all work would be reversible and the gas would heat and

cool as the piston cycled. Derive the equations for this system from the bond graph in Figure 12.19.

12-19. In Problem 12-18, assume the cylinder walls have thermal capacitance C_W and respective heat transfer coefficients h_1 and h_2 for the inside and outside cylinder wall. Account for convection to the cylinder wall on the inside and from the cylinder to the atmosphere, T_0, on the outside. Modify the bond graph and derive state equations for the system.

13

NONLINEAR SYSTEM SIMULATION

In Chapter 5, the bond graph sequential causality assignment procedure was shown to indicate, prior to any equation formulation, the state variables for the model, the number of needed state equations, as well as any possible equation formulation problems. If the procedure resulted in causal assignments on all bonds and all energy storage elements possessed integral causality, then equation formulation would be straightforward. If causality on some bonds was not assigned or some energy storage elements were forced into derivative causality, then algebraic problems exist in the formulation of explicit state equations. This was illustrated in Chapters 4 and 5, and some simulations of simple systems were demonstrated. Most of these systems were linear, and no intractable formulation problems existed.

For strictly linear systems, it is not absolutely necessary to use numerical integration, stepping along in time, to predict system response. Analytical solutions are known for sets of first-order, linear state equations, and there are excellent commercial, linear analysis programs that can present solutions for linear systems in a variety of forms without resorting to numerical integration.

In this chapter we look at the problems of generating solutions to complex, nonlinear systems where linear analysis tools fail. For almost all nonlinear models, system response can be obtained only through numerical simulation. There are several excellent commercial equation solvers that will march out a step-by-step response for nonlinear systems. All these programs require equations to be in some particular form in order to use them. This chapter discusses the various forms that equations can take as a result of modeling decisions as well as ways to solve algebraic problems that may arise.

Copyright © 2012 John Wiley & Sons, Inc.

13.1 EXPLICIT FIRST-ORDER DIFFERENTIAL EQUATIONS

A bond graph model with all integral causality and no unassigned bonds after completion of the sequential assignment procedure (Section 5.2) yields explicit first-order differential equations of the general form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{t}), \tag{13.1}$$

where $\mathbf{x}$ is a vector of state variables, $\mathbf{u}$ is a vector of system inputs, and $\mathbf{f}(\cdot)$ is a vector of functions. Individual equations, as would be derived using methods from Chapter 5, would appear as

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r, t), \\ \dot{x}_2 &= f_2(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r, t), \\ &\vdots \\ \dot{x}_n &= f_n(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r, t). \end{aligned} \tag{13.2}$$

Equations of this form are the most straightforward to solve computationally. The individual equations are written as difference equations, a reasonable time step, Δt, is selected, and the solution is marched out one time step at a time. There are countless algorithms available to solve explicit equations, but they all do basically the same thing [1].

The problem for the modeler is not so much involved with solving equations like (13.2), but much more involved with delivering the right-hand side of Eq. (13.2) to the computer equation solver. A nonlinear system may have nonlinearities that cannot be explicitly written as analytical functions. A 1-port resistance characterized by an effort–flow relationship might be derived from first principles as

$$e = gf^3, \tag{13.3}$$

where g is some constant. Such a relationship is easily incorporated into the right-hand side of Eqs. (13.2). However, the effects being modeled may not easily be represented as analytical functions. For example, if one has only a table of e and f values, perhaps generated experimentally, then one must incorporate the data in the simulation program in a form *equivalent* to the functions in Eqs. (13.2). Many physical effects cannot be realistically represented in equation form; rather, they must be described with logic statements or subroutines. Thus, even though Eqs. (13.2) are in principle the most convenient form for computation of time domain solutions, there is still a lot of work to be done to properly incorporate nonanalytical system constitutive laws in a simulation program.

Fortunately for us, there are many excellent commercial packages available to march out solutions to explicit equations like (13.2). These software packages

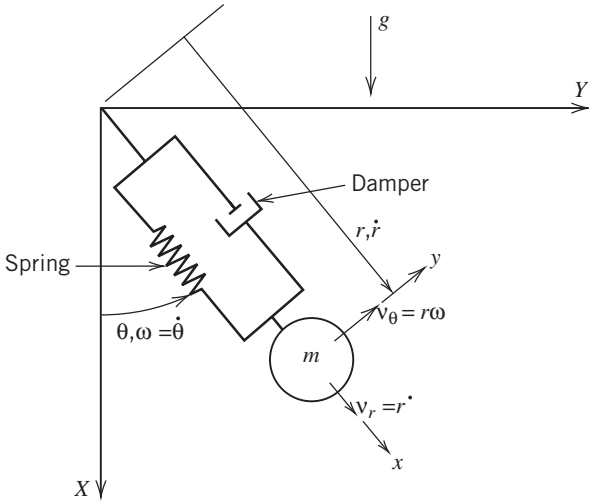


FIGURE 13.1. Nonlinear system exhibiting simultaneous pendulous and spring–mass oscillatory motions.

have built-in functions that handle many commonly encountered nonlinear relationships. They also allow tables of data to be incorporated with automatic interpolation between data points. These packages have excellent built-in plotting capabilities, which make the job of the user almost pleasant.

As an example of reasonably straightforward, nonlinear simulation, consider the system in Figure 13.1. It consists of mass, *m,* in a gravity field, *g,* attached to an inertial frame with a spring and damper. The spring and damper both could be nonlinear so no spring constant or damping constant is shown in the figure. The pendulous motion is not restricted to small values of the angle, θ. Body-fixed coordinates *x, y* are used, and the radial and tangential velocities are indicated. The astute reader will recognize this example as a slightly modified version of Figure 9.8 in Chapter 9.

The bond graph for this system is shown in Figure 13.2. The modulated gyrator (—*MGY*—) with the modulus $m\omega$ is explained in Chapter 9 and is of the type required when using body-fixed coordinates. Causality has been assigned to the bond graph, and the causality completed with three elements in integral causality. The state variables are the r-direction momentum, p_r, the θ-direction momentum, p_θ, and the spring displacement, q. Thus, three state equations are needed to characterize the motion of this system. Note that the spring force, F_s, and damper force, F_d, have been indicated in the bond graph.

Automated equation derivation was demonstrated in Chapter 4 for some simple linear systems. This topic will be addressed again in this chapter. But, for now, the state equations are derived, using the procedure from Chapter 5, as

$$\dot{p}_r = -F_s - F_d + m\omega\frac{p_\theta}{m} + mg\cos\theta, \tag{13.4}$$

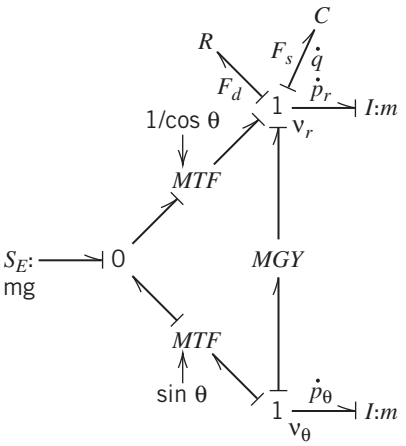


FIGURE 13.2. Bond graphs for the system from Figure 13.1.

$$\dot{p}_\theta = -mg \ \sin\theta - m\omega\frac{p_r}{m}, \tag{13.5}$$

$$\dot{q} = \frac{p_r}{m}, \tag{13.6}$$

where

$$F_s = F_s(q), \tag{13.7}$$

with

$$F_d = F_d\left(\frac{p_r}{m}\right), \tag{13.8}$$

and the kinematic relationship

$$\omega = v_\theta/r = (p_\theta/m)/r = \frac{p_\theta}{mr}. \tag{13.9}$$

Equations (13.4)–(13.9) are almost explicit equations like Eq. (13.2). These equations are ready to be solved once the constitutive behaviors of the spring and damper are specified. The nonlinear constitutive relationships for the spring and damper might be prescribed analytical functions or they may come from experimental data. These would be included straightforwardly in a simulation package.

Notice, however, that the angular displacement θ appears in the differential equations, but since no energy storage depends on θ, it is not automatically a state variable with its own differential equation. Thus, the state space must be expanded by the user such that θ is a state variable and there is an equation for $\dot{\theta}$. This is done by writing the *free integrator* equation

$$\frac{d\theta}{dt} = \omega = \frac{p_\theta}{mr}, \tag{13.10}$$

using Eq. (13.9). Since the radius r is needed in Eq. (13.10) but is also not a state variable in Eqs. (13.4)–(13.6), we write another extra state equation:

$$\frac{dr}{dt} = v_r = \frac{p_r}{m}. \tag{13.11}$$

There must be, in fact, a relationship between the displacement q of the spring (already a state variable) and the radius r, but adding this last extra state variable allows us the convenience of setting the free length of the spring and the initial length of the pendulum separately. Adding extra state variables and state equations is often convenient and causes no significant problems in simulation. A missing equation for a necessary state variable, however, will always cause a fatal error in any simulation program.

These free-integrator equations can now be solved numerically along with the original state equations, Eqs. (13.4)–(13.6), including the relations Eqs. (13.8)–(13.9). This expanding of the state space to make available modulating variables such as θ in this example is often necessary when dealing with mechanical systems that can move through large angular displacements. The displacement modulated transformers that arrise commonly is mechanical systems were discussed in Section 7.3. Fortunately, rates of change of these modulating variables are always available directly from the bond graph, so writing an extra state equation is no problem.

13.2 DIFFERENTIAL ALGEBRAIC EQUATIONS CAUSED BY ALGEBRAIC LOOPS

When causality does not complete on a bond graph after assigning causality by the rules of Chapter 5, an algebraic loop exists among some of the effort and flow variables of the model. This occurrence does not suggest anything about the quality of one's modeling decisions, but those decisions did lead to an algebraic problem. And this algebra must be dealt with prior to simulating the system.

When an algebraic loop exists, derivation of equations yields state equations like

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{z}, t), \tag{13.12}$$

where

$$\mathbf{z} = \mathbf{g}(\mathbf{x}, \mathbf{u}, \mathbf{z}, t), \tag{13.13}$$

and $\mathbf{x}$ is the vector of state variables chosen from the energy storage elements in integral causality, and $\mathbf{u}$ is the vector of inputs. The vector $\mathbf{g}$ contains e and f variables that are part of the algebraic loop, and, unfortunately, these variables depend on themselves, as shown by the right-hand side of Eq. (13.13).

If it is possible to solve Eq. (13.13) for $\mathbf{z}$ as it depends on $\mathbf{x}$, $\mathbf{u}$, and t, then substitution into (13.12) yields explicit equations as described previously. This

procedure was demonstrated in Section 5.4 for the case when the model is linear. When the model is nonlinear, it may not be possible to do the required algebra, especially if **z** contains several e and f variables.

Considering Eqs. (13.12) and (13.13), one can imagine an iterative computational procedure that, starting with the initial values of **x** and **u**, would manipulate the components of **z** until the right-hand side of Eq. (13.13) reproduces the values of the **z** variables on the left side to within some prescribed error bounds. Then this value of **z** substituted into Eq. (13.12) would allow one time step to be taken. This procedure could be repeated every time step. Clearly, this iteration at every time step slows down the simulation. There are computer algorithms that do this procedure more or less automatically, and they will be discussed further in the section on automated simulation.

Another, and perhaps more satisfying, approach to dealing with the algebra is to revisit the modeling assumptions that led to the algebraic problem in the first place. By modifying one's assumptions to include some inertia or compliance at locations where it was previously ignored, the algebraic loops can be made to disappear, and explicit equations like (13.2) will result. This procedure was coined the *Karnopp–Margolis method* in Reference [2], was demonstrated for a complex system in [3], and is shown for a simple system here.

Figure 13.3 shows a mechanical schematic for a system consisting of one mass, one spring, and two damping elements. The spring and dampers are all nonlinear. Assigning causality by the rules of Chapter 5 yields the incomplete bond graph of Figure 13.4*a*. Arbitrarily assigning F_1 into the 0-junction results

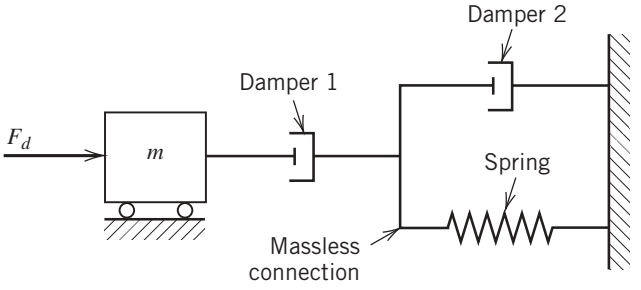


FIGURE 13.3. A simple system with an algebraic loop.

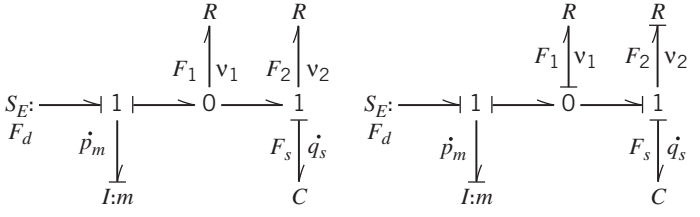


FIGURE 13.4. Bond graph for the system of Figure 13.3.

in the causally complete bond graph of Figure 13.4*b*. Since one arbitrary causal assignment was necessary, we know that one algebraic loop is present.

Assume the R- and C-elements are characterized functionally as

$$F_1 = \Phi_{R_1}(v_1), \tag{13.14}$$

$$v_2 = \Phi_{R_1}^{-1}(F_2), \tag{13.15}$$

$$F_s = \Phi_c(q_s). \tag{13.16}$$

Choosing F_1 as the auxiliary variable in the algebraic loop and using the procedure developed in Chapter 5, we can derive

$$F_1 = \Phi_{R_1}\left(\frac{p_m}{m} - v_2\right), \tag{13.17}$$

$$v_2 = \Phi_{R_2}^{-1}(F_1 - F_s). \tag{13.18}$$

Combining (13.16), (13.17), and (13.18) yields

$$F_1 = \Phi_{R_1}\left\{\frac{p_m}{m} - \Phi_{R_2}^{-1}[F_1 - \Phi_c(q_s)]\right\}. \tag{13.19}$$

This relationship is the example equivalent to Eq. (13.13) for the general case with F_1 playing the role of $\mathbf{z}$. Here we have F_1 related to state variables p_m and q_s and to itself.

The state equations are straightforwardly derived as

$$\dot{p} = F_d - F_1, \tag{13.20}$$

$$\dot{q}_s = \Phi_{R_2}^{-1}[F_1 - \Phi_c(q_s)]. \tag{13.21}$$

These equations are an example of the general Eqs. (13.12). Here the state derivatives depend on themselves and the auxiliary variable, F_1.

If the functions $\Phi_{R_1}(\cdot)$, $\Phi_{R_2}^{-1}(\cdot)$, and $\Phi_c(\cdot)$ are such that Eq. (13.19) could be solved for F_1, then that result can be used in Eqs. (13.20) and (13.21) to yield explicit equations, and the solution is straightforward. However, it should be appreciated that, even when only one variable is involved with the algebraic loop, the nonlinear functions may not allow solution of Eq. (13.19) for F_1, explicitly. If some of the functions are not analytical expressions, but rather tables of data, explicit solution of Eq. (13.19) would be impossible. Generally, if multiple auxiliary variables are involved in the algebra, explicit formulation is out of the question.

Returning to the original model of Figure 13.3, it was assumed that the connection in the middle is massless. This is a modeling decision based on reasonable assessment of the requirements of the model. We know that the connection is not massless, just like we know the proposed constitutive behavior of the elements is

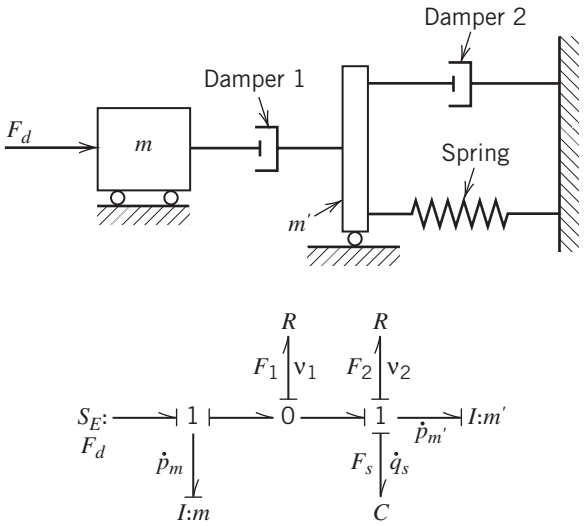


FIGURE 13.5. System from Figure 13.3 with mass at connection.

not exact. In Figure 13.5, the massless connection is changed to a connection has some mass. Also shown in Figure 13.5 is the bond graph with causality assigned. This time the causality completes, and there is no algebraic loop. Derivation of state equations is guaranteed to be in explicit form.

Directly from the bond graph, these equations are

$$\dot{p}_m = F_d - F_1, \tag{13.22}$$

$$\dot{p}_m = F_1 - F_s - F_2, \tag{13.23}$$

$$\dot{q}_s = \frac{p_{m'}}{m'}, \tag{13.24}$$

where

$$F_1 = \Phi_{R_1}\left(\frac{p_m}{m} - \frac{p_{m'}}{m'}\right), \tag{13.25}$$

$$F_2 = \Phi_{R_2}\left(\frac{p_{m'}}{m'}\right), \tag{13.26}$$

$$F_s = \Phi_c(q_s). \tag{13.27}$$

These equations are in explicit form, and, after the constitutive behavior of the nonlinear elements is specified, they are ready for computer simulation.

Upon rethinking our original modeling assumptions in order to generate a causally complete bond graph and, thus, explicit state equations, two things of interest occurred. First, introduction of the originally unintended inertial

element, m', has increased the order of the system from two state equations to three state equations. This is not significant in this simple example, but if the model is large and many algebraic loops exist, such that many additional state variables are required, the modeler may decide to deal with the algebra iteratively as mentioned previously.

The second result of introducing m' into the model is that if m' is made very small, to simulate the originally intended system, a high-frequency component is introduced into the model. The inertial element, m', will interact with the spring and dampers to produce short-period and/or short-time-constant response components, and the computational time step must be appropriately shortened to capture these short time transients in the solution. In general, if the *parasitic elements* introduced to modify the causality of the original model are chosen so as to introduce dynamics about a factor of 10 faster than the original system, the simulation results will be imperceptibly different from the original system response. The increased time required for the simulation due to the shorter step size is about the same as the time needed to perform the algebraic iterations if the model is left in differential algebraic form.

Before leaving this section, the reader should realize that nonlinear systems do not officially have eigenvalues, time constants, or oscillation frequencies as defined in Chapter 6 for linear systems. So using these terms above while describing the effect of introducing parasitic elements to modify causality in nonlinear systems is not, strictly speaking, rigorous. However, I-elements interacting with C-elements will oscillate, I-elements interacting with R-elements will experience decaying motion over some period of time, and R-elements interacting with C-elements will dissipate energy over some time period. The engineer modeling a physical system will have no difficulty in interpreting the dynamics from introduced parasitic elements.

13.3 IMPLICIT EQUATIONS CAUSED BY DERIVATIVE CAUSALITY

In Chapter 5 we saw that after causality is assigned to a bond graph, some I- or C-elements may end up in derivative causality. As with incomplete causality from the previous section, the presence of derivative causality does not suggest anything about the quality of the model or one's modeling decisions. But it does indicate that an algebraic problem exists that must be dealt with prior to simulation of the system.

For linear systems, Chapter 5 demonstrated a procedure to handle the algebraic implications prior to deriving the state equations. For nonlinear systems, the procedure can still be used, but the algebraic manipulations may not be possible to perform.

In general, if derivative causality exists in a bond graph model, the state equations will have the general form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t, \dot{\mathbf{x}}), \tag{13.28}$$

where **x** is the vector of state variables and **u** is vector of inputs. The right-hand side of Eq. (13.28) represents the nonlinear functions of the states, inputs, and state derivatives. These equations are in implicit form and cannot be solved easily numerically because a knowledge of input and state variables does not immediately produce values of the derivatives of the state variables.

One can imagine a computer algorithm that, starting from the initial values of the state variables, guesses initial values of the state derivatives, $\dot{\mathbf{x}}$, and then uses some iterative procedure until the right-hand side of Eq. (13.28) reproduces the left-hand side to within some defined accuracy. Once the $\dot{\mathbf{x}}$ are known, an integration step can be taken, and the process begins again. This solution of implicit equations is quite computer intensive. More will be said about this in the section on automated simulation.

Another approach to dealing with derivative causality is the same as dealing with algebraic loops. We look at the modeling decisions that led to derivative causality and make new modeling decisions that do not introduce the problem. We again include physically motivated parasitic elements that permit explicit equations while retaining the intended performance of the model.

Consider the example system of Figure 13.6. It consists of three nonlinear springs and a nonlinear damper. The springs' constitutive laws are

$$F_1 = \Phi_{c_1}(q_1), \tag{13.29}$$

$$F_2 = \Phi_{c_2}(q_2), \tag{13.30}$$

$$F_3 = \Phi_{c_3}(q_3). \tag{13.31}$$

The damper is characterized by

$$v_d = \Phi_R^{-1}(F_d). \tag{13.32}$$

The bond graph for this system is shown in Figure 13.7, and causality assignment indicates that spring 3 is in derivative causality. Notice the causality on the damper dictates that the flow, v_d, is the output, and the effort, F_d, is the input.

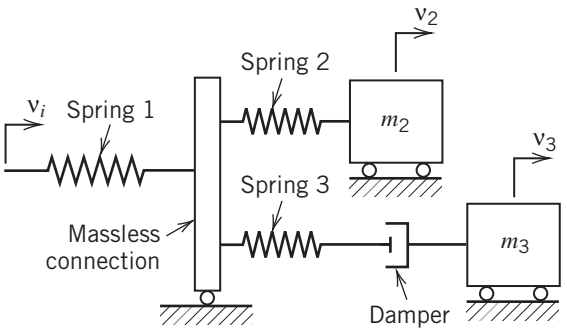


FIGURE 13.6. Nonlinear system exhibiting derivative causality.

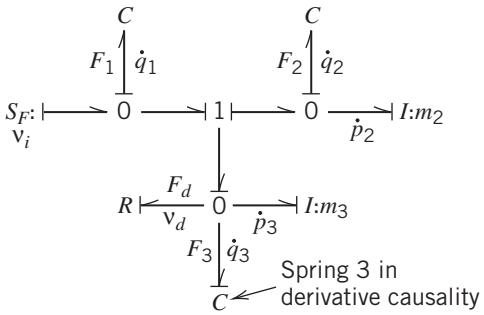


FIGURE 13.7. Bond graph for the system in Figure 13.6.

Procedures from Chapter 5, Section 5.5, show that state equation formulation will require the flow on that spring, $\dot{q}_3$, to be part of the equation derivation. From Eq. (13.31) we can write

$$q_3 = \Phi_{c_3}^{-1}(F_3), \tag{13.33}$$

which says we must be able to invert the constitutive relationships associated with derivative causal energy storage elements. This may not be possible, but assuming it is, we continue. Causality tells us that

$$F_3 = F_1 - F_2, \tag{13.34}$$

where F_1, F_2 are from Eqs. (13.29) and (13.30). Then,

$$\dot{q}_3 = \frac{d\Phi_{c_3}^{-1}(F_3)}{dF_3}\dot{F}_3, \tag{13.35}$$

where

$$\dot{F}_3 = \frac{d\Phi_{c_1}}{dq_1}\dot{q}_1 - \frac{d\Phi_{c_2}}{dq_3}\dot{q}_3. \tag{13.36}$$

The state equations, directly from the bond graph, are

$$\dot{q}_1 = v_1 - v_d - \frac{p_3}{m_3} - \dot{q}_3, \tag{13.37}$$

$$\dot{q}_2 = -\frac{p_2}{m_2} + v_d + \frac{p_3}{m_3} + \dot{q}_3, \tag{13.38}$$

$$\dot{p}_2 = F_2, \tag{13.39}$$

$$\dot{p}_3 = F_1 - F_2, \tag{13.40}$$

where F_1, F_2 come from Eqs. (13.29) and (13.30), and v_d comes from Eq. (13.32), with

$$F_d = F_1 - F_2. \tag{13.41}$$

Notice that q_3 is not an independent state variable even though spring 3 stores energy. Because of the derivative causality, q_3 is algebraically related to the state variables q_1 and q_2. Using Eq. (13.36) in Eq. (13.35) and substituting into Eqs. (13.37) and (13.38), we end up with equations of the general form of Eq. (13.28), where state derivatives appear on the right-hand side of the state equations.

Presuming that the operations indicated in Eqs. (13.35) and (13.36) can be carried out, the iterative solution of these implicit equations could be done. It may also be possible to perform the algebra to get all state derivatives to the left-hand side of Eqs. (13.37) and (13.38) and manipulate until explicit equations result. In general, this will not be possible for nonlinear systems because the constitutive behavior of the elements will not necessarily be differentiable analytical functions but may be bounded functions with discontinuities and/or tables of data from testing real devices.

An alternative approach is to look back at our modeling assumptions and revise them a bit so we do not end up with derivative causality. Figure 13.8

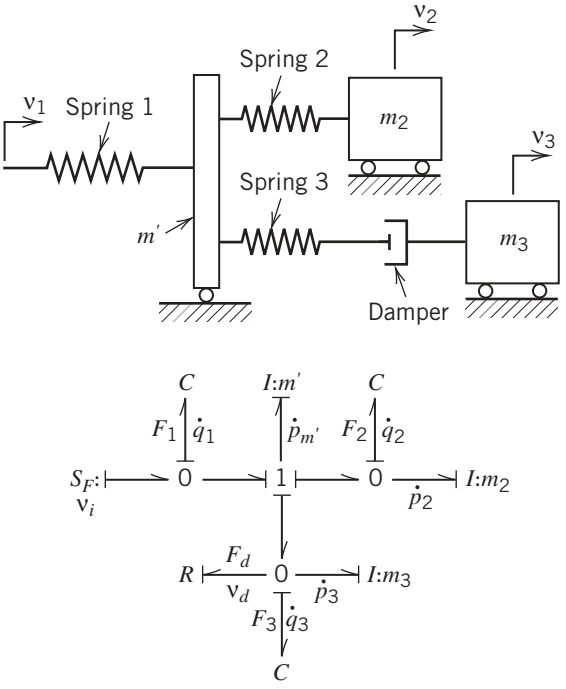


FIGURE 13.8. Revised system of Figure 13.6 without derivative causality.

shows the example system from Figure 13.6 without the assumption of a massless connection. This time a mass, m', has been included at the connection. The bond graph with m' included shows no derivative causality; thus, explicit equations will result from the bond graph.

This time the state equations are

$$\dot{q}_1 = v_i - \frac{p_{m'}}{m'}, \tag{13.42}$$

$$\dot{q}_2 = -\frac{p_2}{m_2} + \frac{p_{m'}}{m'}, \tag{13.43}$$

$$\dot{q}_3 = \frac{p_{m'}}{m'} - \frac{p_3}{m_3} - v_d, \tag{13.44}$$

$$\dot{p}_2 = F_2, \tag{13.45}$$

$$\dot{p}_3 = F_3, \tag{13.46}$$

$$\dot{p}_{m'} = F_1 - F_2 - F_3, \tag{13.47}$$

where F_1, F_2, and F_3 come from the constitutive relationships of Eqs. (13.29), (13.30), and (13.31) and v_d comes from Eq. (13.32), with $F_d = F_3$. Thus, we have explicit, first-order, nonlinear state equations ready for computer solution.

For the case of derivative causality, when we revisited our modeling assumptions and introduced a parasitic element, m', we ended up with two additional state equations. One is due to the introduced I-element, with mass m', and the other is due to switching the causality on spring 3 making q_3 an independent state variable. For low-order models requiring few parasitic element additions, the additional state equations will pose no computational problems. But for very large models, with many derivative causal elements, we may elect to deal with the problem by solving the implicit equations directly. The section on automated simulation will deal with this later.

Finally, when introducing parasitic elements to yield explicit equations, fast dynamic transients are introduced if the parasitic elements are sized to approximate the originally intended model. This was discussed in the previous section concerning algebraic loops. That discussion applies here as well.

13.4 AUTOMATED SIMULATION OF DYNAMIC SYSTEMS

Automated simulation here refers to the many commercial equation solvers that take equations of motion delivered in some prescribed format, march out solutions in time, and present results with very little computer programming required from the user. All of these packages accept equation formats as inputs, and some of them have graphical interfaces such that various graphical icons are interpreted as equations and delivered properly to the equation solver. The current software packages available throughout the world will not be specifically discussed in the text. Here we discuss some generalities of automated simulation.

There are several software packages available for very specific systems, such as automobile dynamics, train dynamics, magnetic circuits, acoustics, plus others. These have specific models already internally built, and the user must only provide data to describe the particular system. These are very useful programs for the intended use but do not represent the system dynamic modeling emphasized in this text. Here we speak of system models that have been constructed by the modeler to develop insight into system behavior or control philosophies for prototyping. Here we discuss software packages that allow simulation of these models in the most straightforward way.

13.4.1 Sorting of Equations

For an equation solver to simulate one time step into the future, it requires the current values of the state-variable derivatives. These derivatives are calculated one at a time from the equations or subroutines the user provides by automatically sorting the equations such that the updated information required to evaluate each derivative has already been computed. As a simple example, consider the equation

$$\frac{dx}{dt} = ax + b, \tag{13.48}$$

where

$$a = 6, \tag{13.49}$$

$$b = 8, \tag{13.50}$$

$$x = 2. \tag{13.51}$$

Without automatic sorting, dx/dt could not be evaluated unless Eqs. (13.49)–(13.51) have been previously evaluated so that a, b, and x have specific numerical values. With automatic sorting, the computer algorithm knows to evaluate a, b, and x prior to calculating dx/dt. This sorting feature may appear trivial or obvious, but it frees the user from providing equations in proper sequence and greatly facilitates putting several submodels together without paying attention to order.

An equally important feature of sorting comes from the following simple example. Consider

$$\frac{dx}{dt} = ax + b, \tag{13.52}$$

where

$$b = c - b + x, \tag{13.53}$$

$$c = 6, \tag{13.54}$$

$$a = 8, \tag{13.55}$$

$$x = 2. \tag{13.56}$$

Here we have a problem sorting the equations. In Eq. (13.53), b depends on itself, and there is no way an automatic sorting routine can find a sequential evaluation of Eqs. (13.53)–(13.56) that will allow an evaluation of dx/dt in Eq. (13.52). An automatic sorting algorithm cannot sort these equations. In principle, one could simply solve Eq. (13.53) for $b = \frac{1}{2}c + \frac{1}{2}x$ and be done with it. A computer algorithm capable of handling differential-algebraic equations* would take the values of c and x, then iterate on b until the right-hand side of Eq. (13.53) equaled the left-hand side. An explicit equation solver, given these equations, would respond with *can't sort* and quit. Thus, modern commercial software packages requiring explicit equations automatically inform the user when the model is *not* computable, but cannot simulate further.

As seen previously, bond graph techniques detect algebraic formulation problems prior to deriving a single equation. This allows modelers to address their modeling assumptions before attempting to simulate the model. This is a most useful virtue of bond graph modeling.

13.4.2 Implicit and Differential Algebraic Equation Solvers

When modeling assumptions lead to algebraic loops or derivative causality, explicit equation solvers cannot sort equations for simulation. If the algebra is not dealt with, then a solution can be marched out only by the iterative procedures described previously. The equations involved are *differential algebraic equations* (DAEs) or *implicit differential equations*.

As mentioned, there are commercial packages devoted to simulating very specific dynamic systems, such as automobiles and trains. These models are in the area of dynamics called "multibody system dynamics," where the inertial bodies are all assumed to be rigid and no relative motion is permitted at attachment points among the bodies. Bond graph models of such systems would be full of derivative causality. For these systems, special algorithms may be used to solve system equations. Because the equations have a particular form due to the particular algebraic problem, the software can march out solutions with no programming from the user.

The modeling procedure developed in this text often results in algebraic problems that must be addressed by the modeler prior to computer solution. Something as common as Coulomb friction, if involved in an algebraic loop, cannot be handled by a DAE solver without significant user attention. While DAE solvers are powerful tools for automated simulation, in this text we emphasize models that yield explicit equations.

13.4.3 Icon-Based Automated Simulation

There are commercial software packages that allow graphical system descriptions that are automatically translated into equations of required format and given

*There are software packages that will attempt to solve differential algebraic equations. These are not nearly as prevalent as explicit equation solvers.

to the equation solver. These icon-based packages are predominantly associated with special applications like electric circuits, hydraulic circuits, mechanical spring–mass–damper systems, magnetic circuits, and acoustic circuits. Some of these programs allow "functional block" inputs, like block diagrams, and are less specific to certain types of system than the aforementioned software.

There are several packages that translate bond graph descriptions into appropriate input files for equation solvers. Since bond graphs allow modeling of all types of energetic systems, linear or nonlinear, the graphical bond graph programs are the most appropriate for overall system dynamics coupling complex mechanical systems with actuators, sensors, and control systems in a variety of energy domains.

If linear modeling assumptions are used in the construction of a model, then only one parameter is needed to describe each basic —*R,* —*I,* —*C,* —*TF*—, or —*GY*— element. For multiport elements, the number of parameters increases but remains constrained. For nonlinear systems, there is an infinite variety of possible constitutive laws. For instance, if a bond graph had a C-element, a computer could interpret this symbol as having the linear constitutive law:

$$e = \frac{q}{C}, \tag{13.57}$$

and the user would only have to supply the value of the compliance parameter, C. If that same C-element were nonlinear, such that

$$e = \Phi_c(q), \tag{13.58}$$

the user would have to provide the computer the functional dependence through some user-provided function or tabular data. Some icon-based graphical bond graph processors interpret the bond graph first as though it were linear. They then allow the user to modify the input file to reflect the intended nonlinearities or simply assign values to appropriate parameters if the element were actually considered to be linear.

The title of the section is Automated Simulation of Dynamic Systems, but after reading all this description, the reader may feel there is nothing at all automated about simulation. It is true that low-order, linear systems can be simulated with virtually no effort from the user. But complex nonlinear systems do require significant user input. What is involved might be best called *computer-aided modeling and simulation* since a computer program can do much of the routine work of setting up a mathematical model for numerical simulation. When the program recognizes a problem, creative human intervention is often required.

Bond graphs allow organization of complex models into manageable parts, and they inform the user of computational problems through the assignment of causality. In the next section, a complex system simulation is demonstrated.

13.5 EXAMPLE NONLINEAR SIMULATION

In Chapter 12, thermodynamic modeling was developed, and a bond graph of a piston-ported, two-stroke engine was shown in Figure 12.20. The example model done here is a simplified version of a two-stroke engine called the "air motor."

The air motor is shown schematically in Figure 13.9. It consists of a slider–crank kinematic mechanism, which drives a piston in a cylinder. Just after top dead center (TDC), the inlet value is opened, exposing the piston to the supply pressure, P_s. The pressure, P, in the cylinder pushes the piston downward. As the piston passes the exhaust port, the piston modulates the flow area and effects the flow of air out of (or into) the cylinder. The idea of the simulation might be to determine the best location, x_e, of the exhaust port so as to extract the most useful work from the supply air. Our purpose here is just to demonstrate a nonlinear simulation. This particular example was chosen because all the components have been modeled previously in the text, but more importantly, there is no way that a linearized analysis will be sufficient here. The only practical way to gain insight into the behavior of this system is through simulation.

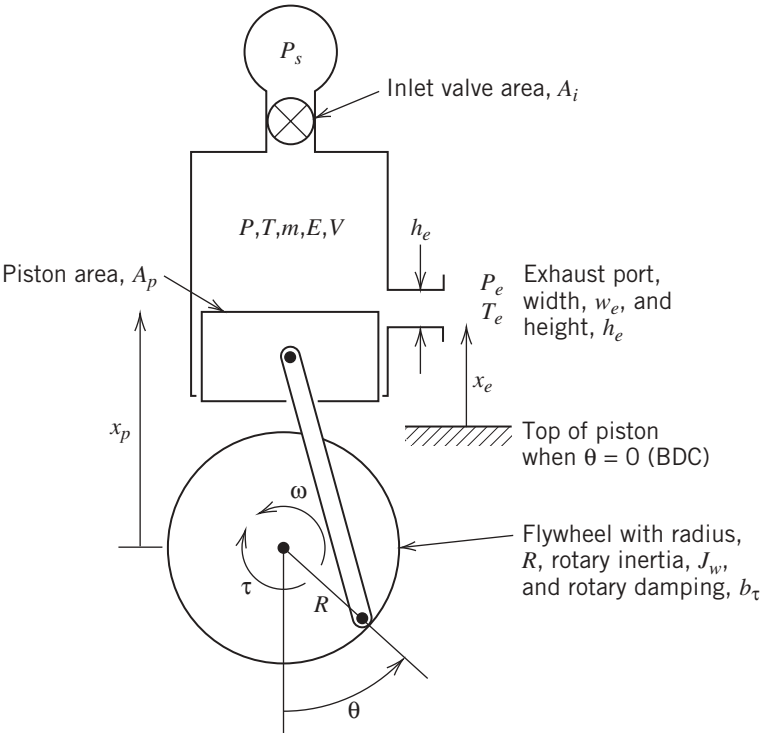


FIGURE 13.9. Schematic diagram of an air motor.

It is assumed that a single pressure, P, temperature, T, mass, m, energy, E, and volume, V, characterize the instantaneous properties of the cylinder head. It is further assumed that the piston is massless and there is no leakage past the piston. Finally, it is assumed that the inlet area, A_i, is exposed instantly upon command, the cylinder head can be modeled as a single thermodynamic accumulator (Section 12.5.1), and the inlet and exhaust ports are each characterized as isentropic nozzles (Section 12.5.2).

The bond graph for the air motor is shown in Figure 13.10. Causality has been assigned and indicates all integral causality with no algebraic loops. Thus, although complex, formulation will yield explicit equations ready for simulation. The thermodynamic accumulator will contribute three state variables, and the flywheel inertia will contribute one state variable. Only four state equations are needed for this simulation.

The 4-port R-element for the inlet requires specification of the supply pressure, P_s, and supply temperature, T_s. It also needs to know the flow area, A_i, which will be specified through some logic statements provided by the user. It will be open when θ exceeds 180° and closed when the piston position, x_p, is less than the exhaust port location, x_e.

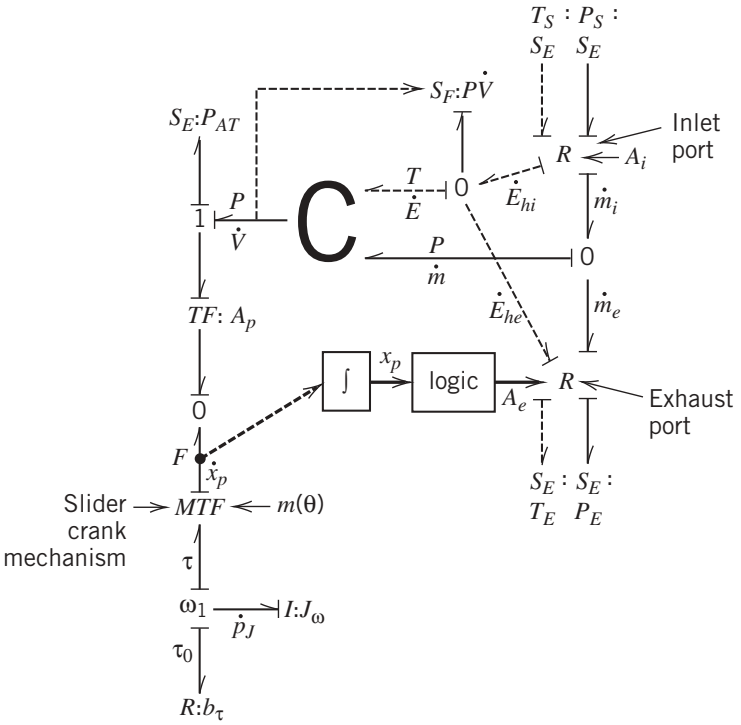


FIGURE 13.10. Bond graph of the air motor.

The 4-port R-element representing the exhaust port requires specification of the pressure, P_e, and temperature, T_e, just outside the port. The upstream conditions come from the causal outputs, P and T, from the 3-port C-element. The exhaust port area, A_e, is modulated by the piston position, x_p, through some user-supplied logic. This is indicated by the ideal measurement of $\dot{x}_p$ followed by signal processing blocks.

The slider–crank is represented as a modulated transformer with modulus, $m(\theta)$. Section 8.1 showed how to derive this modulus:

$$m(\theta) = \left[R \sin\theta - \frac{R^2 \sin\theta \cos\theta}{(L^2 - R^2 \sin^2\theta)^{1/2}} \right]. \tag{13.59}$$

This leads to the equation

$$\dot{x}_p = m(\theta)\omega. \tag{13.60}$$

(Note that in Eq. (13.59) the reference for $\theta = 0$ is bottom dead center while θ was measured from top dead center in Section 8.1, Figure 8.5.)

The state variables are E, m, and V, respectively the energy, mass, and volume of the accumulator, and p_J, the angular momentum of the flywheel.

Following the causal information, the equations can be derived as

$$\dot{E} = -P\dot{V} + \dot{E}_{hi} - \dot{E}_{he}, \tag{13.61}$$

$$\dot{m} = \dot{m}' - \dot{m}_e, \tag{13.62}$$

$$\dot{V} = -A_p \dot{x}_p \tag{13.63}$$

$$\dot{p}_J = -b_\tau \frac{p_J}{J_w} - \tau. \tag{13.64}$$

The enthalpy flow, $\dot{E}_{hi}$, and mass flow, $\dot{m}_i$, for the inlet port are outputs from the 4-port R-element described in Section 12.5.2. In that section a computational procedure is described by Eqs. (12.93)–(12.98). This will not be repeated here. We will just state here the causal implication of the 4-port; thus,

$$\dot{E}_{hi} = f_i(T_s, P_s, T, A_i), \tag{13.65}$$

$$\dot{m}_i = g_i(T_s, P_s, T, A_i). \tag{13.66}$$

Similarly, the energy and mass flows through the exhaust port can be expressed as

$$\dot{E}_{he} = f_e(T, P, T_e, P_e, A_e), \tag{13.67}$$

$$\dot{m}_e = g_e(T, P, T_e, P_e, A_e). \tag{13.68}$$

In these expressions, the outputs from the accumulator are related to the state variables by the compliance relationships,

$$T = \frac{1}{c_v}\frac{E}{m} \tag{13.69}$$

and

$$P = \frac{R}{c_v}\frac{E}{V}. \tag{13.70}$$

These constitutive laws are repeated from Section 12.5.1.

To complete the formulation, we need $\dot{x}_p$ from Eq. (13.63) and the torque, τ, from Eq. (13.64). From the bond graph,

$$\dot{x}_p = m(\theta)\frac{P_J}{J_w} \tag{13.71}$$

$$\tau = m(\theta)A_p(P - P_{\text{atm}}) \tag{13.72}$$

Atmospheric pressure, P_{atm}, is included in the bond graph to convert from absolute pressure, needed on the thermodynamic side of the accumulator, to gauge pressure on the mechanical side.

We must recognize the need for θ and x_p as modulating variables. Here we expand the state space by writing

$$\dot{\theta} = \frac{P_J}{J_w}, \tag{13.73}$$

and θ will be solved along with the other state equations. The displacement, x_p, is already established as a state variable by Eq. (13.71).

Finally, the area modulation of A_i for the inlet and that of A_e for the exhaust are needed to complete the formulation. These would both be specified by logic supplied by the user. For example, for A_i,

$$\begin{aligned} &\text{if } \theta =\geq 180^\circ \quad \text{and} \quad x_p > x_e, \quad A_i = A_{i\,\max} \\ &\text{otherwise,} \qquad\qquad\qquad\quad A_i = 0. \end{aligned} \tag{13.74}$$

This will keep the inlet open until the exhaust port begins to open. The user must be a little careful here to make sure that θ gets reset to zero after each full rotation since θ will increase continuously from integration of Eq. (13.73), and Eq. (13.74) would not work after one rotation.

The exhaust area logic would be

$$\begin{aligned} &A_e = 0; \\ &\text{if } x_p \leq (x_e + h_e), \quad A_e = W_e[x_e + h_e + x_p]; \\ &\text{if } x_p \leq x_e, \qquad A_e = w_e h_e. \end{aligned} \tag{13.75}$$

These statements will modulate the area as the piston uncovers the port, then keeps the area at its maximum value when the port is uncovered, and sets the area to 0 when the port is covered. There may be other strategies that the user will want to try. These are just examples.

Equations (13.61)–(13.75) can be delivered to an explicit equation solver in any order desired, if it has an automatic sorting routine that will evaluate each equation in its proper order to compute the state derivatives. An integrating algorithm will march the solution from one time step to the next. There is no need to eliminate non-state variables such as τ in Eq. (13.64) using Eq. (13.72) to put these relationships in the reduced functional explicit form,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \tag{13.76}$$

since we know that this combination of relationships, Eqs. (13.61)–(13.75), represents a causal, computable, model *equivalent* to Eq. (13.76) and the sorting of equations can be successfully carried out. In fact, some bond graph processors purposefully do not eliminate all the internal efforts and flows in a system to arrive at the form of Eq. (13.76) involving only state and input variables. Instead, they write all the element relationships in the proper causal form such that the result is equivalent to Eq. (13.76). Causal analysis will show that the set of equations can be successfully sorted. Then, after a simulation run, every effort and flow as well as all the state variables are available for plotting.

13.5.1 Some Simulation Results

The air engine parameters are

$D_p = 80$ mm, piston diameter;

$R = 40$ mm, crank radius;

$L = 40$ mm, connecting rod length;

$V_{sq} = 40\ \text{cm}^3$, volume in head at TDC;

$J_w = 1.64 \times 10^{-3}\ \text{k/g}^2$, flywheel rotary inertia;

$b_\tau =$ load resistance, to be varied to determine torque, speed, and power.

For the intake port,

$A_{i\,\max} = 2.9\ \text{cm}^2$, maximum inlet area;

$P_s = 200$ psig $= 14.3$ atm $= 14.3 \times 10^5\ \text{N/m}^2$ supply pressure;

$T_s = 20^\circ$ C $= 293$ K, supply temperature.

For the exhaust port,

$h_e = 17$ mm, height of port;

$w_e = 17$ mm, width of port;

x_e is variable, placement of port relative to crank center.

User might try $0, 20, 40$ mm, ...

For specific results, the initial conditions must be specified. The simulation is started at just past TDC with $\theta = 182^\circ$. The pressure in the cylinder, P, is P_{atm}, and the temperature, T, is 20°C. The cylinder volume, V, is $V_{sq} = 40\ \text{cm}^3$. From Eq. (13.70), we can calculate the initial energy, E, and from Eq. (13.69), we can calculate the initial mass, m. The initial angular momentum, p_J, is zero.

The simulation is initiated with these initial conditions, and simulation proceeds with the piston moving cyclically as energy and mass enter and exit the motor. The state equations were solved using a commercial software package and some representative responses are shown here.

Figure 13.11 shows the cylinder pressure, P, and exhaust and inlet areas as they depend on time. Figure 13.12 shows the cylinder temperature. The pressure and temperature both take several cycles to come to a more or less steady state. When the inlet area opens, the pressure spikes a bit due to the pressure difference between supply pressure and cylinder pressure and the fact that the piston is moving very slowly near TDC. The temperature shows the same behavior.

Figure 13.13 shows the mass flow rate through the inlet port and exhaust port. Interestingly, the inlet port shows a very brief negative spike. This is due to the pressure in the cylinder being higher than the supply pressure when the inlet opens. There is a rapid readjustment followed by flow into the cylinder as the piston is driven downward. The exhaust flow increases rapidly as the exhaust port opens and then decreases as the cylinder pressure drops. The increase in mass flow just before the exhaust port closes is due

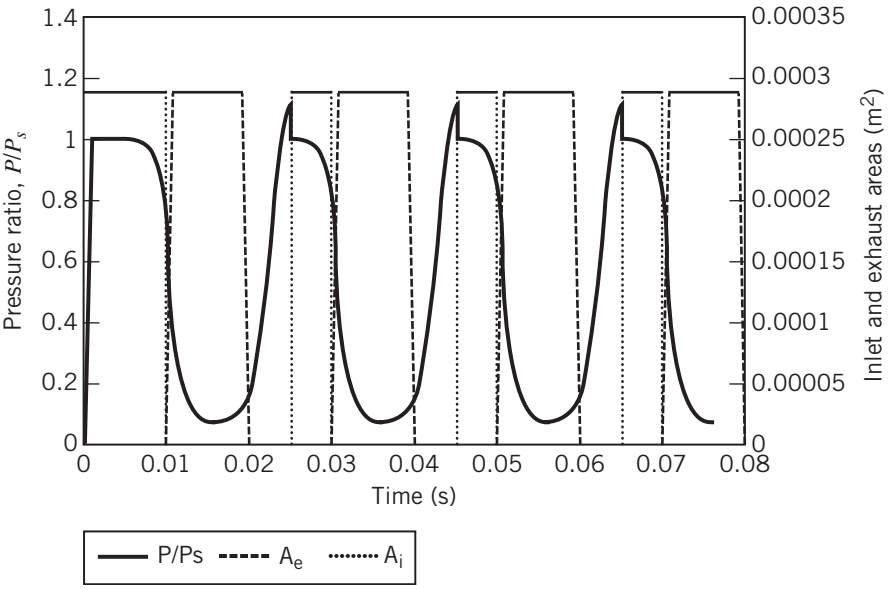


FIGURE 13.11. Pressure in the air engine cylinder.

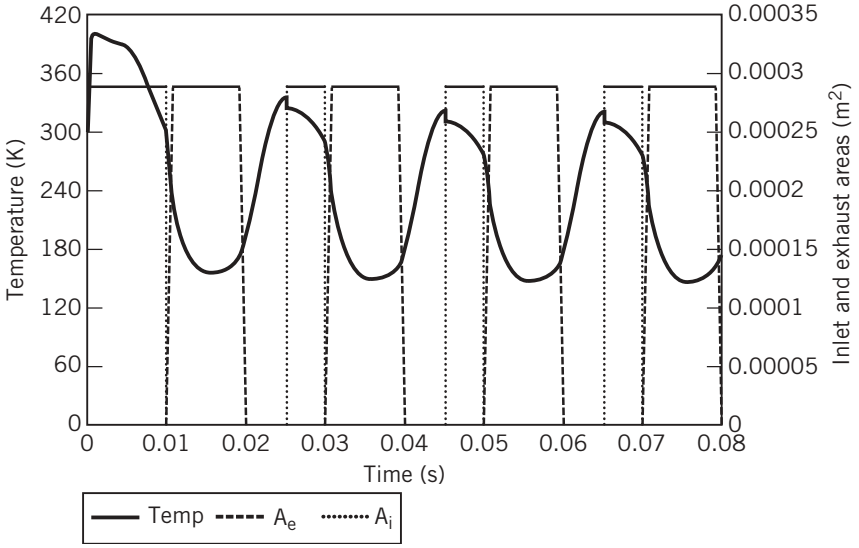


FIGURE 13.12. Temperature in the air engine cylinder.

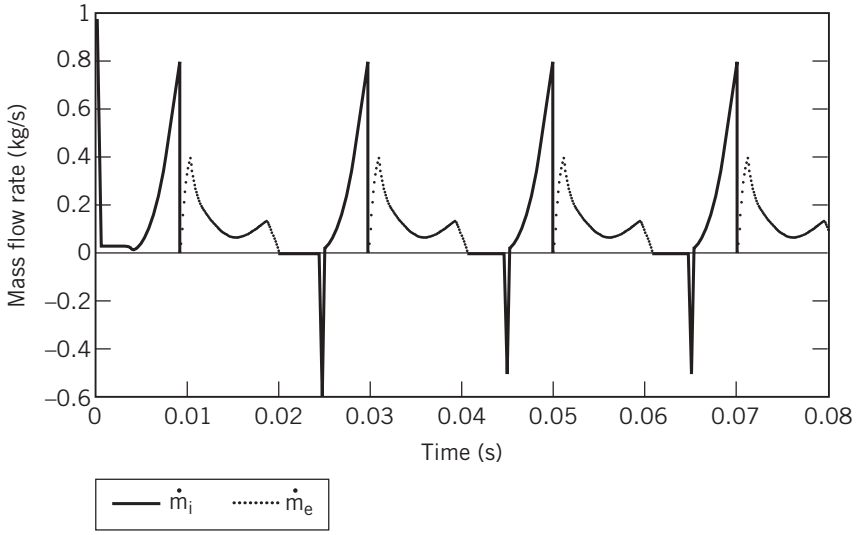


FIGURE 13.13. Mass flow rate through the inlet and exhaust ports.

to the piston having reached bottom dead center and moving upward. As the piston closes the exhaust area, it drives air out through the port.

Figure 13.14 shows the output torque of the engine. This is the torque measured on the torsional resistance, b_τ, from Figure 13.10. It is consistent with the pressure behavior. It is interesting that the torque peak does not coincide with the peak

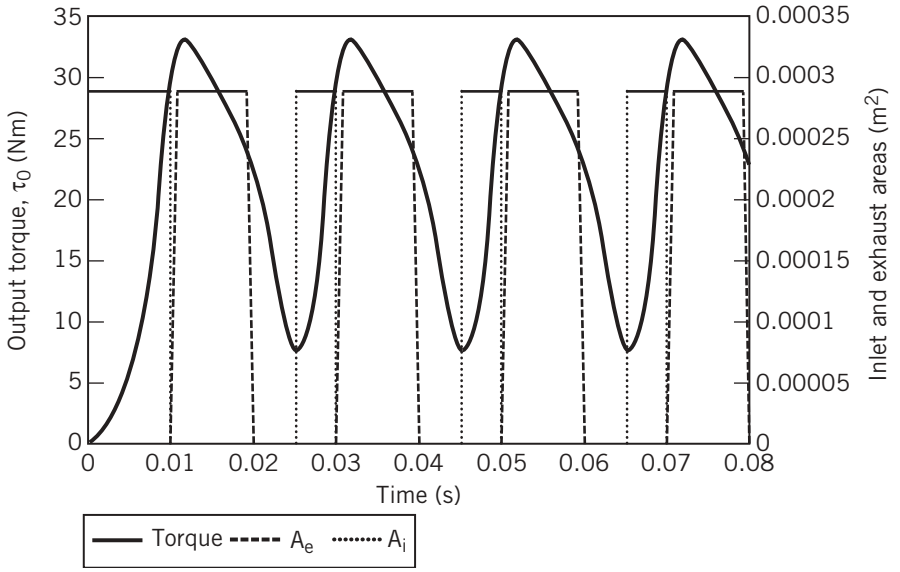


FIGURE 13.14. Output torque, τ, from the air engine.

pressure. This is due to the inherent behavior of a slider–crank mechanism. Near TDC, where the pressure is highest, there is no moment arm to transmit torque to the flywheel. As rotation proceeds, the moment arm grows as the pressure decreases.

13.6 SUMMARY

This chapter has attempted to expose the complexities of nonlinear simulation and the organization provided by using bond graphs to put complex systems together. Causality is a powerful tool in that formulation problems are exposed before equation formulation, providing the user with choices in how to proceed toward a computable model. Causality also dictates the input and output variables for the nonlinear constitutive laws for the elements. This allows a priori determination of whether a particular relationship must be inverted if required by the formulation demands.

It was shown here and in Chapter 9 that nonlinear geometry of mechanics requires modulation variables (displacements) that are not algebraically available from the state variables associated with stored energy. In this case, the state space must be expanded to include *free integrators,* which are solved along with the energy state variables to provide the needed displacements. Finally, an example was developed to demonstrate a complex simulation. It is hoped that the reader appreciates the usefulness of the bond graph in developing the simulation model.

REFERENCES

[1] B. Carnahan, H. A. Luther, and J. O. Wilkes, *Applied Numerical Methods*, New York: Wiley, 1969.

[2] A. Zeid and C.-H. Chung, "Bond graph modeling of multibody systems: a library of three-dimensional joints," *J. Franklin Inst.*, **329**, 605–636 (1992).

[3] D. Margolis and D. Karnopp "Analysis and simulation of planar mechanisms using bond graphs," *ASME J. Mech. Des.*, **101**, No. 2, 187–191 (1979).

PROBLEMS

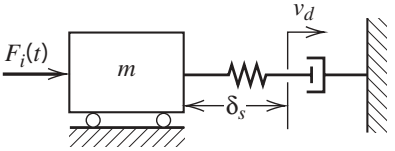
13-1. The system shown has a nonlinear spring characterized by

$$F_s = g_s \delta_s^3$$

and a nonlinear damper characterized by

$$F_d = g_d v_d^3.$$

Construct a bond graph model, derive equations, and put them into a form for computer simulation.



13-2. For the system of Problem 13-1, what changes must be made in the computer model if the spring and damper are characterized by square laws rather than cubic laws?

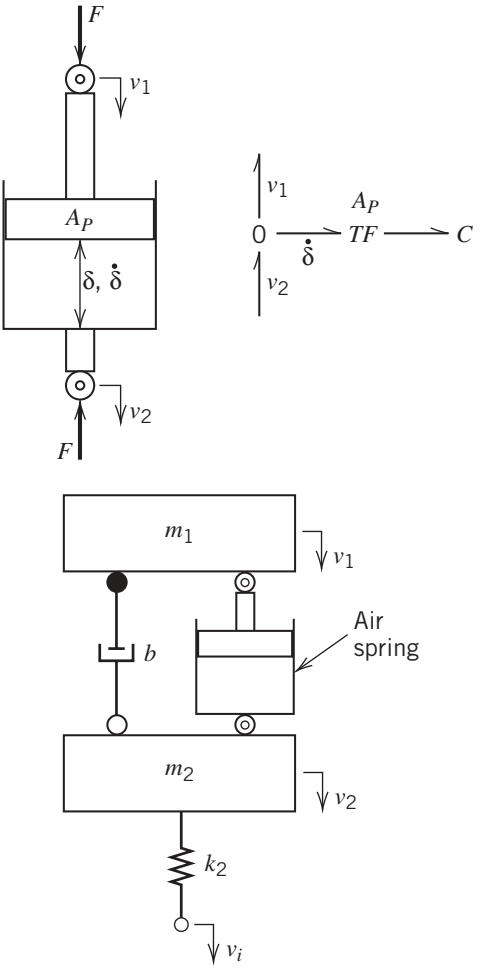
13-3. For the system of Problem 13-1, include friction between the mass and ground and reformulate the bond graph. Derive equations and put them into a form ready for computer simulation. Describe the program statements that would be necessary in order to include friction.

13-4. The device shown is an air spring that is assumed to behave isentropically. The constitutive relationship for this device has been determined to be

$$F = P_0 A_P \left[\frac{1 - \left(\frac{A_P \delta}{V_0} \right)^\gamma}{\left(\frac{A_P \delta}{V_0} \right)^\gamma} \right],$$

where P_0 is atmospheric pressure and V_0 is the chamber volume when $A_P\delta/V_0 = 1$. A bond graph for the device is shown with attachment points exposed.

Include this air spring in the quarter-car model, construct a bond graph model, and derive a complete set of state equations. Organize these equations into a form suitable for computer simulation. Explain any program statements needed to handle the special case when $A_P\delta/V_0 \to 0$.

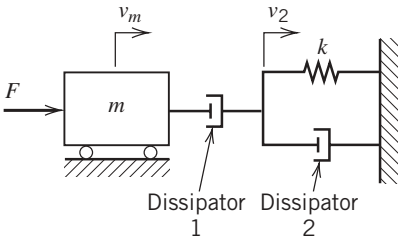


13-5. The system shown has two nonlinear dissipation elements with constitutive behavior,

$$F_1 = g_1 v_1^3, \quad v_1 = v_m - v_2,$$

$$F_2 = g_2 |v_2| v_2.$$

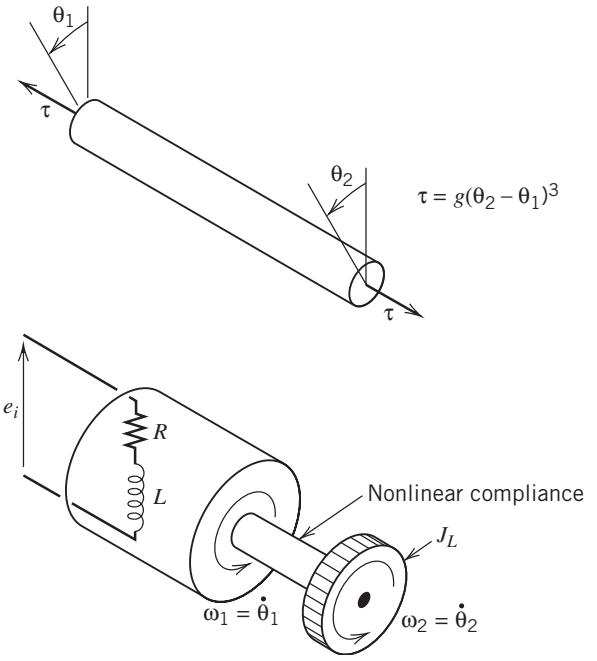
Construct a bond graph model and determine that an algebraic loop exists. Attempt to derive state equations using the procedure from Chapter 5, Section 5.4. Now add a parasitic element to eliminate the algebraic loop. Derive state equations and organize them into a form suitable for computer simulation.



13-6. The system shown is a dc motor with a flexible shaft and attached inertial load. The shaft is nonlinear and behaves according to

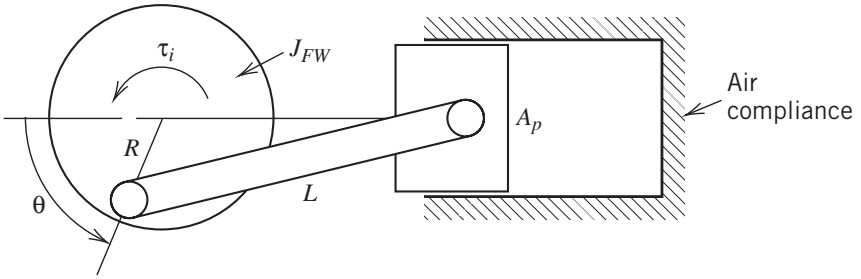
$$\tau = g(\theta_2 - \theta_1)^3.$$

Construct a bond graph model, assign causality, and convince yourself that derivative causality exists. Attempt to derive a computable model using the methods from Chapter 5, Section 5.5. Append a parasitic element to your model that eliminates derivative causality and assign a physical interpretation to this element. Derive the resulting explicit equations and put them into a form ready for computer simulation.

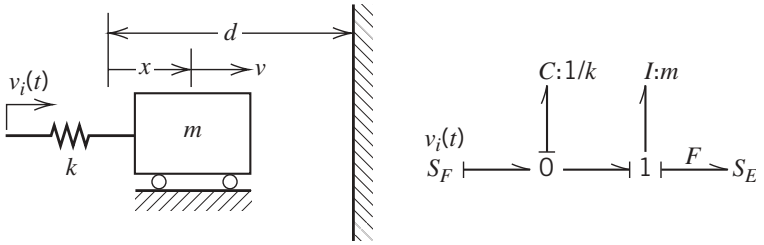


13-7. Figure 12.19 shows a thermodynamic system and its causal bond graph. Derive the state equations for this system and show they are appropriate for computer simulation.

13-8. The system shown couples the air spring from Problem 13-4 and the slider–crank device from Eq. (13.59). Construct a bond graph model of the system and derive state equations. Expand the state space as necessary and set up your equations for computer simulation.

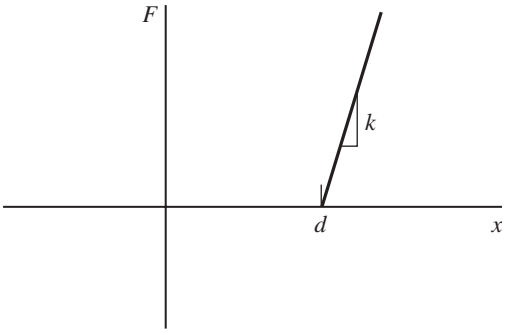


13-9. The mass, *m,* bounces off the rigid wall with no loss of energy. One way of simulating this phenomenon is to specify a boundary impulsive force, *F,* that acts over a specified but very short amount of time (perhaps one simulation time step) and accelerates the mass such that it comes off the wall with the same speed as the approach, but in the opposite direction.



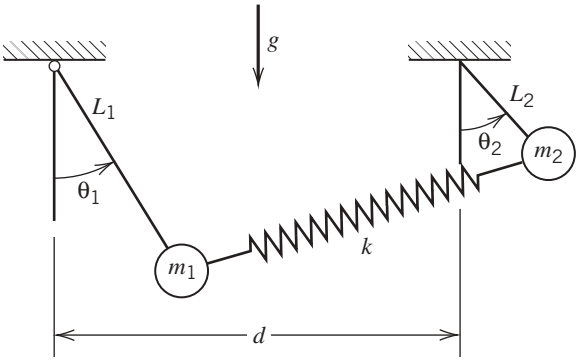
For an impact duration of Δt seconds, calculate the value of this force, assuming knowledge of the approach velocity; set up the state equations for simulation, including the logic statements needed to handle the boundary impulse.

13-10. For Problem 13-9, another way to handle the impact problem is to assume there is a spring between the mass and wall with the constitutive behavior shown.



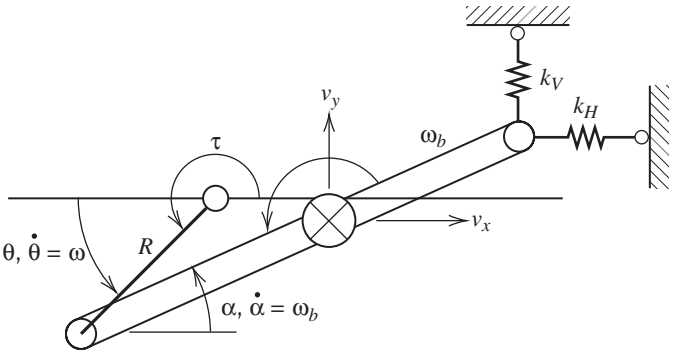
The spring constant, *k,* must be "very stiff" to simulate the originally intended system, and this may require shortening the simulation time step during contact with the wall. On the bond graph for Problem 13-9, put a C-element in place of the effort source and derive equations ready for simulation. Include the logic necessary to handle the "stiff" compliance.

13-11. Two unequal-length pendulums are attached by a spring as shown. Construct a bond graph model that would account for large angular deflections. Derive a *complete* state representation and set up for computer simulation. The spring is relaxed when both pendulums are vertical.



13-12. In the slider–crank mechanism, assume the only important inertial element is the connecting rod of mass *m,* centroidal moment of inertia *J,* and length *L.* A modified device is shown here where the horizontal sliding constraint has been replaced by two springs, k_H and k_V. If k_V was

very stiff, then the horizontal sliding constraint would be approached. Construct a bond graph model of this system and note the use of *MTF*s involving angles θ and α. Derive a complete state representation and organize them for simulation. (*Hint:* Transfer the center of mass motion to the end points of the rod; then enforce the velocity constraint at the crank, and derive the spring velocities at the other end.)



APPENDIX: TYPICAL MATERIAL PROPERTY VALUES USEFUL IN MODELING MECHANICAL, ACOUSTIC, AND HYDRAULIC ELEMENTS

Mass Density, ρ, [kg/m^3]

Solids	
Aluminum:	2,700 kg/m^3
Copper:	8,900 kg/m^3
Rubber, hard:	1,100 kg/m^3
Rubber, soft:	950 kg/m^3
Steel:	7,700 kg/m^3
Titanium:	4,500 kg/m^3
Liquids	
Hydraulic oil, well de-aerated:	900 kg/m^3
Water, fresh, at 20°C:	998 kg/m^3
Water, sea, at 134°C:	1,026 kg/m^3
Gases	
Air at 1 atm. and 20°C:	1.21 kg/m^3
Air at 1 atm. and 0°C:	1.29 kg/m^3
Hydrogen at 1 atm. and 0°C:	0.09 kg/m^3

Modulus of Elasticity, E, [Pa = N/m^2]

Aluminum:	71,000 N/mm^2 = 71 × 10^9 Pa
Copper:	122,000 N/mm^2 = 122 × 10^9 Pa
Hard rubber:	2,300 N/mm^2 = 2.3 × 10^9 Pa
Soft rubber:	5 N/mm^2 = 0.005 × 10^9 Pa
Steel:	206,000 N/mm^2 = 206 × 10^9 Pa
Titanium:	110,000 N/mm^2 = 110 × 10^9 Pa

Bulk Modulus, B, [Pa = N/m^2]

Water at 20°C:	2.18 × 10^9 Pa
Hydraulic oil (well de-aerated):	1.52 × 10^9 Pa

Copyright © 2012 John Wiley & Sons, Inc.

Speed of Sound, c, [m/s]

Air at 1 atm. and 20°C:	343 m/s
Air at 1 atm. and 0°C:	332 m/s
Hydrogen at 1 atm. and 0°C:	1,269.5 m/s
Water, fresh, at 20°C:	1481 m/s
Water, sea, at 0°C:	1500 m/s

Coefficient of Shear Viscoscity, μ, [Pa · s = Ns/m^2]

Air at 1 atm. and 20°C:	1.8×10^{-5} Pa · s
Castor oil:	0.96 Pa · s
Water, fresh, at 20°C:	1.0×10^{-3} Pa · s

Ratio of Specific Heats, $\gamma = c_p/c_v$

Air	1.40
Carbon dioxide	
(low frequency)	1.30
(high frequency)	1.40
Nitrogen	1.40

INDEX

0-junction, 58, 59
  causality, 63
1-junction, 60, 63
  causality, 63
1-port, 19
  compliance, 41
  inertia, 43
  resistor, 38
2-port, 18
  gyrator, 53
  transformer, 51

Accumulator, thermodynamic, 579
Acoustic approximation, 122, 576
Active bond, 29, 33, 385, 387
Actuators, 392
Air motor, bond graph, 617
Air motor simulation, 620
Air spring, 598
Algebraic loop, 183, 604
  nonlinear, 605
A-matrix, 167
Amplifiers, 385, 388, 392
Angular momentum, 413
Augmenting the bond graph, 168
Automated simulation, 12, 198, 614
Automatic control, 258
  proportional, 260
    proportional plus derivative, 265
    proportional plus derivative plus integral, 269
    lead-lag, 284
    motion, 289
    state variable feedback, 296
    vibration control, 301
Automotive cornering forces, 349, 468
Automotive cornering stiffness, 468
Automotive shock absorber, 156

Bar, vibrating, 471, 482, 513, 515
Beam:
  Bernoulli-Euler, 480
  Timoshenko, 476
  transverse motion, 476
Bernoulli resistor, 571, 595
Bernoulli's equation, 569
Bernoulli-Euler beam, 480
Bicycle car model, 468, 469
Block diagrams, 30, 67, 240, 255, 392, 395
B-matrix, 168
Bode plot, 247, 316
Body fixed coordinates, 116, 412, 416
  modulated gyrator, 118, 414, 463
Bond, 27
  active, 29
Bond graph, 5, 12.
  air motor, 616
  algebraic loop, 183
  augmentation, 169
  co-energy variables, 176
  derivative causality, 65, 170, 172
  equation formulation problems, 183
  output variables, 10, 30, 220
  reducible loop, 97
  state variable, 9
Bulk modulus, 129, 159, 373, 568

Capacitance:
  elastic pipe, 126, 120
  rigid pipe, 126, 120
  water storage tank, 126
Capacitor, 40
  moveable plate, 380
Cardan angles, 416
Casimir form, 352, 358, 364
Causality, 31, 63
Causality assignment procedure, 171
Causal stroke, 3-
C-element, 42
Center of mass, 106
C-field, 326
Characteristic equation, 224, 228, 485
Characteristic value, 228
Closed loop system, 259
Closed loop transfer function, 287
Co-energy variables, 176
Common effort junction, 58
Common flow junction, 60
Complex frequency response, 247
Complex plane, 234
Compliance, 40
Compliance matrix, 330

Copyright © 2012 John Wiley & Sons, Inc.

Compressible gas dynamics, 549, 578
Compression, isentropic, 42, 574, 583
Conductance causality, 39, 350
Conduction heat transfer, 548
Constant coefficient linear systems, 219
Constitutive relationship, 38, 40, 43
Control, see Automatic control
Control actions, 259
  proportional, 260
  proportional plus derivative, 265
  proportional plus derivative plus integral, 269
Control volume, 552, 565, 571, 578
Coordinate transformations, 362, 416
Coordinates, body-fixed, 116, 412, 412
Cramer's rule, 240, 241
Critical damping, 232

Damped natural frequency, 232
Damped oscillator, 232
Damper, semi-active, 392
Damping ratio, 232
D-C motor, 18, 24, 140, 374
Delta function, 475
Derivative causality, 65
  implicit equations, 191, 608
Descriptor form of equations, 434, 437
Differential algebraic equations, 437, 604
Differential equations, partial, 475, 481, 566
Displacement variable, 20
Distributed parameter systems, 470
Drive train model, 29, 69, 435

Effort junction, 60
Effort source, 48
Eigenvalue, 11, 222, 224
Elastic modulus, 128, 328
Electrical alternators, 377, 400
Electrical circuits, 78
Electrical circuit construction procedure, 79
Electrical circuit power convention, 79
Electrical circuit reference voltage, 79
Electrical networks, 84
Electrical systems, 78
  D-C motor, 18, 24, 140, 374
  Pi network, 89
  Tee network, 90
  voice coil, 139
Electrical transformer, 52, 84
Electrically linear, 382
Electrohydraulic valve, 7
Electromechanical transducers, 139
Energy, 17
  Gibbs, 554
  Helmholtz, 554
  internal, 550, 554
Energy variables, 20
Enthalpy, 554, 575
Entropy, 549, 552, 558
  Start here
Equation derivation, 175
  algebraic loop, 183, 604
  derivative causality, 65, 188, 191, 608
Equations, sorting, 203, 613
Euler angles, 416
Euler's equations, 414
Euler's formula, 222
Eulerian description, 565, 571, 577
Explicit equations, 201, 601

Faraday' law, 139, 356, 520
Ferromagnetic material, 520
Fields, 326
  explicit, 329
  IC, 445
  implicit, 2329
Field strength, 521, 523
Finite lump, 470
First order system, 223, 226
Flow junction, 58
Flow source, 48
Flow variable, 19
Fluid capacitance, 125
Fluid circuit construction procedure, 132
Fluid dynamic systems, 565, 578
Fluid inertia, 44, 130
Fluid line, 134
Fluid resistance, 39
Fluid system, acoustic muffler, 122, 135
Flux, 47, 351, 377
Flux linkage variable, 23, 178, 196, 377, 383, 519
Forced response, 222, 229, 487
Force-free boundaries, 492
Force-free modes, 492
Forcing frequency, 245
Forward loop transfer function, 287
Free response, 218, 222
Frequency, damped natural, 232, 234
  undamped natural, 230, 264
Frequency equation, 485
Frequency response, 11, 244
  complex, 247
    normalization, 252
  examples, 249
Friction, dry, 171, 372, 562

Gear set, 51, 105
Generator, a-c, 377, 379, 385
Generator, d-c, 141, 375
Gibbs equation, 553
Gibbs free energy, 554
Gyrator, 53
  causality, 63
  modulated, 55, 118, 354, 376
Gyrator transducers, 139
Gyroscope, 53

Half car model, 108
Hamiltonian form, 440
Hamiltonian *IC* field, 440
Heat conduction, 548
Heat transfer, 548
Heave-pitch vehicle model, 108
Helmholtz free energy, 554
Hydraulic and acoustic circuits, 121
Hydraulic models, acoustic approximation, 121
Hydraulic motor, 133
Hydraulic pump, 133
Hydraulic ram, 52, 137, 372
Hydraulic shock absorber, 156, 393
Hydraulic valve, 7, 391
Hydrostatic system, 121, 122, 126, 575

I-element, 45
  linear, 40
  nonlinear, 43
I-field causality, 285, 287
Imaginary part of, 226
Implicit equations, 534
  derivative causality, 534
  nonlinear example, 535
Inductance matrix, 286, 307
Inertia, 45
  causality, 65

Inertia matrix, 445
Initial condition problem, 78, 222, 226
Initial conditions, 78, 222, 226
Input/output concept, 240, 248
Instruments, 33, 371
Internal energy, 550, 554
Irreversible thermodynamics, 351
Isentropic nozzle, 585
Isentropic process, 129, 574
Isothermal models, 551

Junctions, 58
  0-junction, 58
  1-junction, 60
Junction structure, 89, 170, 326

Karnopp–Margolis method, 605
Kinematic constraints, 111, 116
Kinematic linkages, 56
Kinetic energy of rigid body, 106
Kirchhoff's laws, 164, 171

Ladder network, 89
Lagrange multipliers, 435
Lagrange's equations, 440
Lagrangian description, 571, 577
Lagrangian *IC* field, 445
Laplace transforms, 219
Lead-lag control, 284
Legendre transformation, 554, 555
Lever, 51
Linear system, 11
Lorentz force law, 140, 356
Loudspeaker, 18, 34, 139
Lumped models, 470
Lumped parameter systems, 470

Magnetic bond graph variables, 519
Magnetic circuits, 519
Magnetic energy, 524
Magnetic flux, 519
  density, 519
Magnetic induction, 519
Magnetic saturation, 521
Magnetizing force, 521
Magnetomechanical elements, 532
Magnetomotive force, 523, 539
Mass controlled region, 500
Mathematical model, 1, 4
Maxwell reciprocity, 352, 382
Mechanical systems, 61, 91
  in rotation, 100
  rotation construction procedure, 100
  in translation, 91
  translation construction procedure, 93
Mechatronic system, 1
MGY, 55, 118, 354, 376
Microphone, 380
Modal analysis, 482
Modal bond graph systems, 443
Modal damping, 503
Modal inertia, 504
Modal mass, 504
Modal stiffness, 488, 500
Model, 4
Mode shapes, 485
Modulated gyrator, 55, 118, 354, 376
Modulated transformer, 35, 55, 108, 121, 354, 357
Modulus of elasticity, 128, 328
Momentum variable, 20
Motion control, 289
Motor, a-c, 377, 385, 536
Motor, d-c, 18, 24, 140, 374
Motors, synchronous, 536
Moveable plate capacitor, 380
MTF, 35, 55, 108, 121, 354, 357
Multibody systems, 416, 433
Multi-energy-domain models, 136
Multiport, 17

Nonlinear simulation, example, 198, 602
Normal modes, 482
Nozzle, 571
  isentropic, 585

Open loop system, 258
Oscillator
  Damped, 232
  Undamped, 230
One-junction, 60
Onsager reciprocity, 351, 352
Open loop transfer function, 287
Orthogonality, 488, 496
Output variables, 10, 30, 220, 226, 259
Output/input, concept, 30
Overdamped, 232

p, momentum variable, 20
Parasitic elements, 372, 608
Perfect gas, 556, 561
Permanent magnet, 56, 141, 372, 520, 525
Permeability, 522, 528
Permeance, 522, 524, 529
Phase angle,
Pi network, 90
Plane motion, 106
Pneumatic actuator, 588, 596
Pneumatic systems, 122, 566, 575
Pneumatic tires, 446
Poles and zeros, 268
  open loop
  closed loop
Power, 19
Power convention, 25
Power variables, 19
Pressure:
  dynamic, 122, 571
  static, 122, 571
Pressure momentum, 22
Proportional control, 260
Proportional plus derivative control, 265
Proportional plus derivative plus integral control, 269
P, PD, PID control, 260–269
Pseudo-bond graphs, 326, 548, 578
Pump, 133, 138, 379

q, displacement variable, 20
Quarter car model, 94, 96, 108
Quarter car, simulation, 198

Real part of, 231, 233, 237
R-element 38
Reluctance, 522
Residual compliance, 501
Resistance:
  controlled, 387
  eletrothermal, 563
  modulated, 386
Resistance causality, 350
Resistor, 38
  Bernoulli, 571,595
  R-field, 326
Rigid-body motion, 412
Root locus, 273
  examples, 281
Rotating frame, 118

Second-order system, 225
Semi-active damper, 392
Sensor, 259
Separation of variables, 482
Sequential causality assignment, 171
Servo-valve, hydraulic, 8, 389
Shock absorber, 156
SI units, 22

Side branch accumulator, 252
Signals, input, 67
Signals, output, 67, 394
Simulation of state equations, 162
Simulation, 12
Simulation, air motor, 616
Simulation, automated, 12, 162, 612
Simulation, nonlinear, 198
Sinusoidal forcing, 244
Slider-crank mechanism, 34, 616
Slip angle 448, 468
Solenoid, 349, 383
Sorting of equations, 203
Source element, effort, 48
Source element, flow, 48
Source, controlled, 88, 386
Specific heats, 42, 550, 556, 581
Speed of sound, 127
Stable systems, 235
State determined systems, 9
State equation derivation, 175
State equations, 175, 202, 219
State equations, algebraic loop, 183
State equations, explicit, 201, 601
State equations, implicit, 188, 608
State equations, linear, 167
State equations, linear, matrix form, 167
State equations, sorting, 203
State equations, standard form, 220
State variables, 9, 170, 202
State variables, choosing, 170
State variable feedback, 296
State variables, vector, 160
Steady state response, 245
Stiffness controlled region, 500
Stiffness matrix, 330
String, vibrating, 514
Superposition, 222
Suspension, automotive, 94, 108, 198, 283
Suspension, controlled, 393
Symmetry of matrix, 331
Synchronous motors, 536

Tee network, 90
Tetrahedron of state, 21, 47
Thermal compliance, 549, 592
Thermal energy, 558
Thermal resistance, 558, 561
Thermal systems, 548, 565
Thermodynamic accumulator, 579–584
Time constant, 224
Timoshenko beam, 479, 480, 481
Transducers, 133, 136
Transducers, energy-storing, 380–385
Transducers, gyrator type, 54, 139
Transducers, magnetic, 377, 378, 532–543
Transducers, power, 372, 375, 377
Transducers, transformer, 137
Transfer functions, 239–244
  closed loop, 287
  examples, 249
  open loop, 287
  forward loop, 287
Transformation matrix, 360, 443
Transformation, Legendre, 554, 555
Transformation, power-conserving, 418
Transformations, coordinate, 362, 416, 419
Transformations, velocity, 428
Transformer transducers, 137
Transformer, 51
Transformer, causality, 66
Transformer, modulated, 55, 56, 108, 121, 354, 356, 363, 418, 456
Transformer, multiport, 359
Trivial solution, 223, 226, 236

Undamped natural frequency, 230
Undamped oscillator, 230
Underdamped system, 232, 234
Unstable systems, 235

Valve, four-way, 389
Valve, hydraulic, 389–391, 406
Variable, displacement, 20
Variable, effort, 19
Variable, energy, 20
Variable, flow, 19
Variable, momentum, 20
Variables, acoustic systems, 19
Variables, electrical systems, 24
Variables, hydraulic systems, 23
Variables, mechanical rotation, 23
Variables, mechanical translation, 20
Vector matrix form, 220
Vehicle dynamics, 411, 445
Vibration control, 301
Vibration of a bar, 472–475, 483–491
Vibration of a beam, 471–482
Viscosity, 123
Voice coil actuator, 53, 56, 139, 306, 539
Voice coil, 54

Wankel compressor or engine, 590–591
Wave equation, 475, 576
Wheatstone bridge, 83
Word bond graph, 27, 29, 35

Zero-junction, 53, 55
Zeros and poles, 268