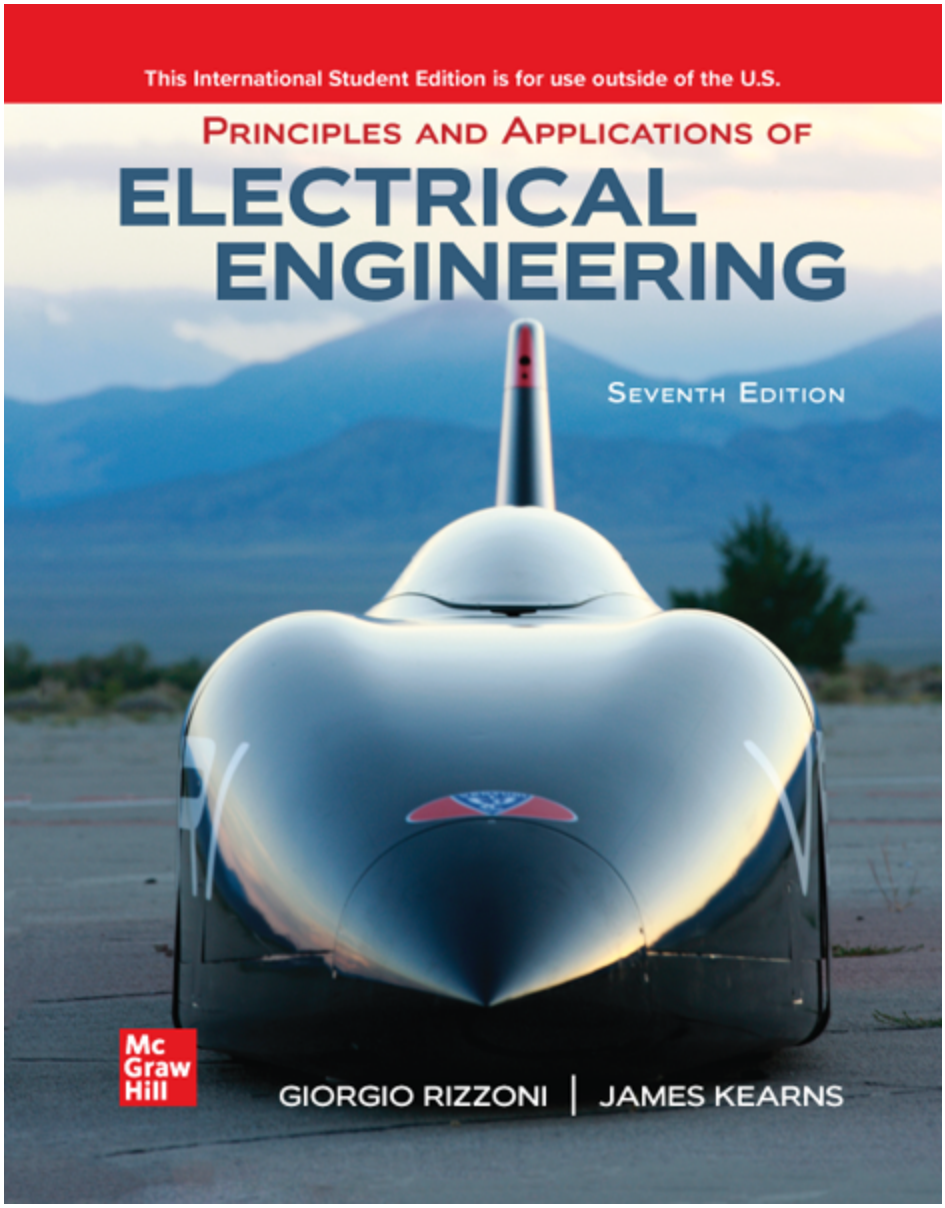

This International Student Edition is for use outside of the U.S.
PRINCIPLES AND APPLICATIONS OF
ELECTRICAL ENGINEERING
SEVENTH EDITION
Mc Graw Hill
GIORGIO RIZZONI | JAMES KEARNS

PRINCIPLES AND APPLICATIONS OF ELECTRICAL ENGINEERING

Seventh Edition

Giorgio Rizzoni
The Ohio State University

James Kearns
York College of Pennsylvania

PRINCIPLES AND APPLICATIONS OF ELECTRICAL ENGINEERING

Published by McGraw Hill LLC, 1325 Avenue of the Americas, New York, NY 10121.

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 LWI 24 23 22 21

ISBN 978-1-260-59809-4
MHID 1-260-59809-8

Cover Image: *Denis Boussard*

mheducation.com/highered

About the Authors

Giorgio Rizzoni, the *Ford Motor Company Chair in ElectroMechanical Systems*, is a Professor of Mechanical and Aerospace Engineering and of Electrical and Computer Engineering at The Ohio State University (OSU). He received his B.S. in 1980, his M.S. in 1982, and his Ph.D. in 1986, in Electrical and Computer Engineering, all from the University of Michigan. Since 1999 he has been the director of the Ohio State University Center for Automotive Research (CAR), an interdisciplinary research center in the OSU College of Engineering.

Dr. Rizzoni's research interests are in the dynamics and control of future ground vehicle propulsion systems, including advanced engines, alternative fuels, electric and hybrid-electric drivetrains, energy storage systems, and fuel cell systems. He has contributed to the development of a graduate curriculum in these areas and has served as the director of three U.S. Department of Energy Graduate Automotive Technology Education Centers of Excellence: *Hybrid Drivetrains and Control Systems* (1998–2004), *Advanced Propulsion Systems* (2005–2011), and *Energy Efficient Vehicles for Sustainable Mobility* (2011–2016).

In 1999 Dr. Rizzoni established an automotive industry research consortium that today sees the participation of over 20 automotive OEMs and suppliers; in 2008 he created the SMART@CAR consortium, focusing on plug-in hybrid and electric vehicles and vehicle-grid interaction, with funding from electric utilities, automotive OEMS, and electronics suppliers. Through the Ohio Third Frontier Wright Project Program he created a *Center of Excellence for Commercial Hybrid Vehicles* in 2009, and a *Center of Excellence for Energy Storage Technology* in 2010.

Dr. Rizzoni is a Fellow of IEEE (2004), a Fellow of SAE (2005), a recipient of the 1991 National Science Foundation Presidential Young Investigator Award, and of several other technical and teaching awards.

The OSU Center for Automotive Research

The OSU Center for Automotive Research, CAR, is an interdisciplinary research center in the OSU College of Engineering founded in 1991 and located in a 50,000 ft^2 building complex on the west campus of OSU. CAR conducts interdisciplinary research in collaboration with the OSU colleges of Engineering, Medicine, Business, and Arts and Sciences, and with industry and government partners. CAR research aims to: develop efficient vehicle propulsion and energy storage systems; develop new sustainable mobility concepts; reduce the impact of vehicles on the environment; improve vehicle safety and reduce occupant and pedestrian injuries; increase vehicle autonomy and intelligence; and create quieter and more comfortable automobiles. A team of 50 administrative and research staff supports some 40 faculty, 120 graduate and 300 undergraduate students and maintains and makes use of advanced experimental facilities. Dr. Rizzoni has led CAR for over a decade, growing its research expenditures from $1M per year to over $10M today, and engaging CAR in a broad range of technology commercialization activities, start-up company incubation and spin-out, as well as providing a broad range of engineering services to the automotive industry.

CAR is also the home of the OSU Motorsports program, which supports the activities of five student vehicle competition programs of several student vehicle competition programs including: the Buckeye Bullet (holder of all current U.S. and FIA electric vehicle land speed records), the EcoCAR hybrid-electric vehicle team, the Formula Buckeyes and Baja Buckeyes SAE teams, and the Buckeye Current electric motorcycle racing team.

Jim Kearns is an Associate Professor of Electrical & Computer Engineering at York College of Pennsylvania. He received a B.S. in Mechanical Engineering (SEAS) and a B.S. in Economics (Wharton) from the University of Pennsylvania in 1982. Subsequently, he received his M.E. from Carnegie-Mellon University in 1984, and his Ph.D. from the Georgia Institute of Technology in 1990, both in Mechanical Engineering. While at Georgia Tech he was the recipient of a Presidential

Fellowship. Subsequently, he worked as a Postdoctoral Fellow at the Applied Research Laboratory of the University of Texas—Austin.

In 1992, Dr. Kearns took his first teaching position at the Universidad del Turabo in Gurabo, Puerto Rico, where he worked with a small group of faculty and staff to build and develop a new school of engineering. In addition to other duties, he was tasked with developing a curriculum on electromechanics. During this time Dr. Kearns spent his summers at Sandia National Laboratories as a University Fellow.

In 1996, Dr. Kearns was the second full-time engineering faculty member hired by York College of Pennsylvania to (once again) develop a new engineering program with an emphasis on Mechatronics. As a result of that work, Jim was asked in 2003 to develop new electrical and computer engineering programs at YCP. Jim served as program coordinator until July 2010.

Throughout Dr. Kearns professional career he has been involved in teaching and research related to physical acoustics and electromechanical systems. His interest in electrical engineering began during his Ph.D. studies, when he built spark generators, DC power supplies, and signal amplifiers for his experiments. His steady pursuit of electromechanical engineering education has been the hallmark of his professional career. Dr. Kearns has been involved in a variety of pedagogical activities, including the development and refinement of techniques in electrical engineering education.

Dr. Kearns is a member of IEEE and ASEE. He is active in faculty governance at York College, where he is a past chair of its Tenure and Promotion committee and its Student Welfare committee. Dr. Kearns recently completed a four-year term as Vice-President and then President of the York College Academic Senate.

About the Cover

On the cover, an image of the Venturi Buckeye Bullet 3 at the Bonneville Salt Flats, UT, in 2016. On Monday, September 19, 2016, The Ohio State University's Venturi Buckeye Bullet 3 student team and driver Roger

Schroer took this electric streamliner vehicle to a world record two-way average top speed of 341.4 miles per hour (549.4 kilometers per hour). The vehicle was designed and built by a team of engineering students at The Ohio State University, advised by Dr. Giorgio Rizzoni.

Brief Contents

Mc Graw Hill Education
5011111

Contents

Preface

The pervasive presence of electronic devices and instrumentation in all aspects of engineering design and analysis is one of the manifestations of the electronic revolution that has characterized the last 60 years. Every aspect of engineering practice, and of everyday life, has been affected in some way or another by electrical and electronic devices and instruments. Laptop and tablet computers along with so-called "smart" phones and touchscreen interfaces are perhaps the most obvious manifestations. These devices, and their underlying technology, have brought about a revolution in computing, communication, and entertainment. They allow us to store, process, and share professional and personal data and to access audio (most notably, music) and video of every variety. These advances in electrical engineering technology have had enormous impacts on all other fields of engineering, including mechanical, industrial, computer, civil, aeronautical, aerospace, chemical, nuclear, materials, and biological engineering. This rapidly expanding electrical and electronic technology has been adopted, leveraged, and incorporated in engineering designs across all fields. As a result, engineers work on projects requiring effective communication across multiple disciplines, one of which is nearly always electrical engineering.

0.1 OBJECTIVES

Engineering education and professional practice continue to undergo profound changes in an attempt to best utilize relevant advances in electronic technology. The need for textbooks and other learning resources that relate these advances to engineering disciplines beyond electrical and computer engineering continues to grow. This fact is evident in the ever-expanding application and integration of electronics and computer technologies in commercial products and processes. This textbook and its associated learning resources represent one effort to make the principles of

electrical and computer engineering accessible to students in various engineering disciplines.

The principal objective of the book is to present the *principles* of electrical, electronic, and electromechanical engineering to an audience of engineering majors enrolled in introductory and more advanced or specialized electrical engineering courses.

A second objective is to present these principles with a focus on important results and common yet effective *analytical and computational* tools to solve practical problems.

Finally, a third objective of the book is to illustrate, by way of concrete, fully worked examples, a number of relevant *applications* of electrical engineering. These examples are drawn from the authors' industrial research experience and from ideas contributed by practicing engineers and industrial partners.

These three objectives are met through the use of various pedagogical features and methods.

0.2 ORGANIZATION

The basic organizational structure of a generic chapter remains essentially unchanged from the previous edition. Example problems and associated methods and procedures of problem solving remain organized so that students are able to easily and efficiently locate them when doing homework and preparing for exams. Page xAdditional unguided exercises are provided to test student understanding. Relevant and stimulating applications to practical measurement challenges are included in nearly every chapter.

A continued and enhanced emphasis on problem solving can be found in this edition. All the highlighted *Focus on Problem Solving* boxes have been reviewed and revised to clarify and add additional detail to the steps needed by students to successfully complete end-of-chapter homework problems.

An effort was also made to reduce the aesthetic complexity of the book, without sacrificing technical content or overall aesthetic appeal. Effective reading is promoted by less clutter and visual "noise." A thorough,

exhaustive, page-by-page search was made to locate errors in the text, equations, figures, references to equations and figures, examples, and homework problems.

The book is now divided into five major parts:

I. **Circuit Analysis**
II. **Systems and Instrumentation**
III. **Analog Electronics**
IV. **Digital Electronics**
V. **Electric Power and Machines**

The pedagogical enhancements made within each part are discussed below.

0.3 PEDAGOGY AND CONTENT

Part I: Circuit Analysis

Once again, the first part of the book has undergone a significant revision from the previous edition.

Chapter 1 begins with an emphasis on developing a student's ability to recognize structure within a circuit diagram. It is the authors' experience that this ability is key to student success. Yet, many books contain little content on developing this ability. The result is that many students wander into more difficult topics still viewing a circuit as simply an unruly collection of wires and elements.

The approach taken in this book is to encourage students to initially *focus on nodes*, rather than elements, in a circuit. For example, some of the earliest exercises in this book ask students to count the number of nodes in a circuit diagram. One immediate advantage of this patient approach is that students learn to disregard the particular aesthetic structure of a circuit diagram and instead focus on the technical structure and content. Chapter 1 also immediately engages students in the terminology, laws, and methods needed to solve basic DC problems and introduces the first of many electromechanical analogies.

Chapter 2 introduces students to more sophisticated analytic methods with a focus on appreciating the implications and utility of equivalent networks. The students' skill at recognizing circuit structure is further developed by the introduction of elements in series and parallel, applied to the more general concept of equivalent resistance between two nodes. The principle of superposition and the *source-load perspective* followed by Thévenin and Norton equivalent networks complete Chapter 2. The section on the source-load perspective revisits the concepts of voltage and current division to develop their graphical solution as the intersection of a Page xisource's load line with the load's *v-i* relation. This section is not essential but it can be very helpful to students when introduced prior to the usually difficult topic of Thévenin equivalent networks.

Methods of Problem Solving were enhanced and clarified. Throughout these chapters students are encouraged to think of problem solving in two steps: first **simplify**; then **solve**. In addition to being an effective problem-solving method, this method provides context for the power and importance

of equivalent networks in general, and Thévenin's theorem, in particular. Chapter 3 continues the emphasis on equivalent networks applied to AC circuit analysis. In the following chapters on transient analysis and frequency response, foundational first- and second-order circuit *archetypes* are identified. Students are encouraged to continue to use Thévenin and Norton equivalent networks to simplify, when possible, transient circuit problems to these archetypes, which, in effect, become targets for students.

Finally, emphasis continues to be placed on visualizing phasors in the complex plane and understanding the key role of the unit phasor and Euler's theorem. Throughout the chapter on AC circuits students are encouraged to focus on the concepts of impedance and power triangles, and their similarity. Single-phase AC power concepts are now addressed in the chapter on AC circuits, whereas material related to transformers and three-phase power were moved to Part V.

Part II: Systems and Instrumentation

This part of the textbook brings together all of the material related to measurement and instrumentation found in the sixth edition and represents a significant change. The chapter on operational amplifiers continues to emphasize three amplifier archetypes (the unity-gain buffer, the inverting amplifier, and the noninverting amplifier) before introducing variations and applications, which are now more readily related to issues and challenges commonly encountered when conducting measurements using electronic instrumentation. The discussion of instrumentation amplifiers, in particular, was expanded and clarified. It is hoped that the reorganization of this material will bring greater relevance and practicality to students at an early stage of their study and allow instructors to complete this material and that in Part I in a one-semester course.

Part III: Analog Electronics

While much of the content on electronics is unchanged from the sixth edition, the problem-solving strategies and techniques for transistor circuits were further enhanced and clarified. The focus on simple but useful circuit examples was not changed.

The emphasis on large-signal models of BJTs and FETs and their applications was retained; however, an appropriate, but limited, presentation of small-signal models was included to support the discussion of AC amplifiers. These chapters present an uncomplicated and practical treatment of the analysis and design of simple amplifiers and switching circuits.

The chapter on power electronics is no longer included in the textbook but can be found in the online resources that support the book.

Part IV: Digital Electronics

The chapters on digital electronics remain largely unchanged except for a needed update of the material on encoders, gate arrays, and programmable logic devices. A greater number of end-of-chapter problems are now included in the chapter on digital systems.

It should be noted that the chapters on communication systems have been removed from the textbook but can be found in the online resources that support it.

Part V: Electric Power and Machines

Part V reflects a change in the organization of the book that brings together those aspects of electrical engineering that are related to electric power systems. Every instructor understands that there is no unique way of presenting introductory electrical engineering material, and the positioning of Part V in the book is somewhat arbitrary, as the section could really be placed anywhere after Section I. Chapter 13 covers the fundamentals of electric power systems, largely unchanged from previous editions, introducing AC power, complex power, and elements of three-phase power systems. Chapters 14 and 15 offer an introductory treatment of electrical machines, with focus on DC, and AC synchronous and induction machines. Two ancillary chapters are available online for instructors who wish to have a more in-depth treatment of electromechanical systems: one on power electronics, which introduces devices and systems for electric power conversion; the other on special-purpose electric machines, which presents a survey of electric machines commonly used in industrial systems and consumer products, such as step motors, brushless DC machines, switched reluctance machines, and single-phase AC machines. The content of Chapters 14 and 15 and of the ancillary chapters was developed by the first author for use in a required junior-year system dynamics course for mechanical engineers, and in a technical elective on mechatronics systems.

0.4 NOTATION

The notation used in this book for various symbols (variables, parameters, and units) has been updated but still follows generally accepted conventions. Distinctions in notation can be subtle. Luckily, very often the context in which a symbol appears makes its meaning clear. When the meaning of a symbol is not clear from its context a correct reading of the notation is important. A reasonably complete listing of the symbols used in this book and their notation is presented below.

For example, an uppercase roman font is used for units such as volts (V) and amperes (A). An uppercase italics math font is used for real parameters and variables such as resistance (R) and DC voltage (V). Notice the difference between the variable V and the unit V. Further, an uppercase bold

math font is used for complex quantities such as voltage and current phasors (**V** and **I**) as well as impedance (**Z**), conductance (**Y**), and frequency response functions (**H** and **G**). Lowercase italic symbols are, in general, time dependent variables, such as voltage (v or $v(t)$) and current (i or $i(t)$), where (t) is an explicit indication of time Page xiiidependence. Lowercase italic variables may represent constants in specific cases. Uppercase italic variables are reserved for constant (time-invariant) values exclusively.

Various subscripts are also used to denote particular instances or multiple occurrences of parameters and variables. Exponents are italicized superscripts.

Finally, in electrical engineering the imaginary unit $\sqrt{-1}$ is always represented by j rather than i, which is used by mathematicians. The reason for the use of j instead of i should be obvious!

Quantity	Symbol	Description
Voltage	v or $v(t)$	Time Dependent and Real
	V	Time Invariant and Real
	V	Complex Phasor
Effective (rms) voltage	$\tilde{V}$	Time Invariant and Real
Current	i or $i(t)$	Time Dependent and Real
	I	Time Invariant and Real
	I	Complex Phasor
Effective (rms) current	$\tilde{I}$	Time Invariant and Real
Volts	V	Unit of voltage
Amperes	A	Unit of current
Resistance	R	Real
Inductance	L	Real
Capacitance	C	Real
Reactance	X	Frequency Dependent and Real
Impedance	**Z**	Frequency Dependent and Complex
Conductance	**Y**	Frequency Dependent and Complex
Transfer Function	**G** or **H**	Frequency Dependent and Complex
Cyclical Frequency	f	Time Invariant and Real
Angular Frequency	ω	Time Invariant and Real
Angle	θ	Time Invariant and Real
Amplitude	A	Time Invariant and Real

0.5 SYSTEM OF UNITS

This book employs the International System of Units (also called SI, from the French Système International des Unitès). SI units are adhered to by virtually all professional engineering societies and are based upon the seven fundamental quantities listed in Table 0.1. All other units are derived from these base units. An example of a derived unit is the radian, which is a measure of plane angles. In this book, angles are in units of radians unless explicitly given otherwise as degrees.

Since quantities often need to be described in large multiples or small fractions of a unit, the standard prefixes listed in Table 0.2 are used to denote SI units in powers of 10. In general, engineering units are expressed in powers of 10 that are multiples of 3. For example, 10^{-4} s would be expressed as 100×10^{-6} s, or 100 μs.

Page xiv

Tables 0.1 and 0.2 are useful references when reading this book.

Table 0.1 SI units

Quantity	Unit	Symbol
Length	Meter	m
Mass	Kilogram	kg
Time	Second	s
Electric current	Ampere	A
Temperature	Kelvin	K
Substance	Mole	mol
Luminous intensity	Candela	cd

Table 0.2 Standard prefixes

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

0.6 ADDITIONAL FEATURES OF THE SEVENTH EDITION

Pedagogy

The seventh edition continues to offer all the time-tested pedagogical features available in the earlier editions.

- **Learning Objectives** offer an overview of key chapter ideas. Each chapter opens with a list of major objectives, and throughout the chapter the learning objective icon indicates targeted references to each objective.
- **Focus on Problem Solving** sections summarize important methods and procedures for the solution of common problems and assist the student in developing a methodical approach to problem solving.
- **Clearly Illustrated Examples** illustrate relevant applications of electrical engineering principles. The examples are fully integrated with the Focus on Problem Solving material, and each one is organized according to a prescribed set of logical steps.
- **Check Your Understanding** exercises follow each set of examples and allow students to confirm their mastery of concepts.
- **Make the Connection** sidebars present analogies that illuminate electrical engineering concepts using other concepts from engineering disciplines.

- **Focus on Measurements** boxes emphasize the great relevance of electrical engineering to the science and practice of measurement.

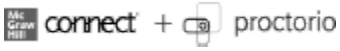
Instructor Resources on Connect:

Instructors have access to these files, which are housed in Connect.

- **PowerPoint presentation slides** of important figures from the text
- **Instructor's Solutions Manual** with complete solutions

Page xv

Remote Proctoring & Browser-Locking Capabilities



New remote proctoring and browser-locking capabilities, hosted by Proctorio within Connect, provide control of the assessment environment by enabling security options and verifying the identity of the student.

Seamlessly integrated within Connect, these services allow instructors to control students' assessment experience by restricting browser activity, recording students' activity, and verifying students are doing their own work.

Instant and detailed reporting gives instructors an at-a-glance view of potential academic integrity concerns, thereby avoiding personal bias and supporting evidence-based claims.

Writing Assignment

Available within McGraw Hill Connect®, the Writing Assignment tool delivers a learning experience to help students improve their written communication skills and conceptual understanding. As an instructor you can assign, monitor, grade, and provide feedback on writing more efficiently and effectively.

0.7 ACKNOWLEDGMENTS

The authors would like to recognize the help and assistance of reviewers, students, and colleagues who have provided invaluable support. In particular, Dr. Ralph Tanner of Western Michigan University has painstakingly reviewed the book for accuracy and has provided rigorous feedback, and Ms. Jiyu Zhang, PhD student at Ohio State, has been generous in her assistance with the electromechanical systems portion of the chapter. The authors are especially grateful to Dr. Domenico Bianchi and Dr. Gian Luca Storti for creating many new homework problems and solutions and for their willingness to pitch in whenever needed. This seventh edition is much improved due to their efforts.

The authors also wish to acknowledge Dr. Jason Forsyth (James Madison University) and Dr. James Moscola (York College of Pennsylvania) for updating the material on combinational logic modules in Chapter 11.

Throughout the preparation of this edition, Kathryn Rizzoni has provided editorial support and has served as an interface to the editorial staff at MHHE. We are grateful for her patience, her time invested in the project, her unwavering encouragement, her kind words, and her willingness to discuss gardening and honeybees.

The book has been critically reviewed by:

- Riadh Habash—The University of Ottawa
- Ahmad Nafisi—California Polytechnic State University
- Raveendra Rao—The University of Western Ontario
- Belinda Wang—The University of Toronto
- Brian Peterson—United States Air Force Academy
- John Durkin—University of Akron
- Chris Klein—Ohio State University
- Ting-Chung Poon—Virginia Tech
- James R. Rowland—University of Kansas
- N. Jill Schoof—Maine Maritime Academy

- Shiva Kumar—McMaster University
- Tom Sullivan—Carnegie Mellon University
- Dr. James Moscola—York College of Pennsylvania

- Dr. Jason Forsyth—James Madison University

In addition, we would like to thank the many colleagues who have pointed out errors and inconsistencies and who have made other valuable suggestions.

Comments by Giorgio Rizzoni

As always, a new edition represents a new era. I am truly grateful to my friend and co-author, Jim Kearns, for taking on a new challenge and for bringing his perspective and experience to the book. Jim and I share a passion for teaching, and throughout this project we have invariably agreed on which course to take. It is not easy to find a suitable co-author in the life of a project of this magnitude, and I have been fortunate to find a friend willing to undertake a new journey with me.

When the first edition of the book was nearing, so was the birth of our first child, Alessandro (Alex). The second and third editions were marked by the births of Maria Caterina (Cat), and Michael. Time passes, Kathryn continues to be my best friend and partner, and now Cat is a successful industrial designer, Alex is studying electrical engineering (imagine that!), and Michael computer science and engineering. The years go by, but my family continues to be an endless source of joy, pleasant surprises and, always, smiles. Many thanks to Kathryn, Alex, Cat, and Michael for always being there to support and encourage me.

Comments by James Kearns

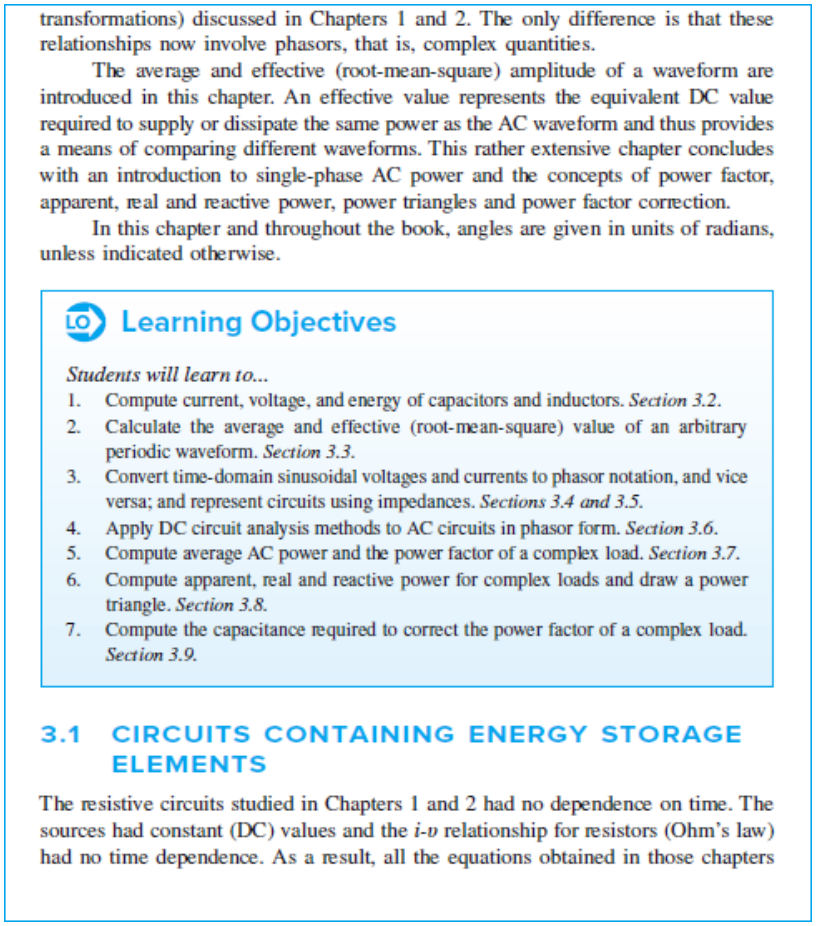
My association with this remarkable book continues to be a great privilege and honor. Its contents continue to reflect the enormous effort and expertise of the principal author, and my dear friend, Dr. Giorgio Rizzoni. His leadership and vision were essential to the creation of this new edition. I remain awestruck by his seemingly unbounded energy and enthusiasm and humbled by his kind, considerate, and generous ways.

As with all things, the love and support of my family and friends sustained me throughout this work. My children, Kevin, Claire, and Caroline, continue to bless, inspire, and inform my daily life.

Finally, I wish to once again thank my parents for their many years of unconditional love and support. Despite having lost both of them in recent years they remain very much alive within me and present in my work. Can anyone ever begin to measure or repay the gift of loving parents?

Guided Tour

Learning Objectives offer an overview of key chapter ideas. Each chapter opens with a list of major objectives, and throughout the chapter the learning objective icon indicates targeted references to each objective.



transformations) discussed in Chapters 1 and 2. The only difference is that these relationships now involve phasors, that is, complex quantities.

The average and effective (root-mean-square) amplitude of a waveform are introduced in this chapter. An effective value represents the equivalent DC value required to supply or dissipate the same power as the AC waveform and thus provides a means of comparing different waveforms. This rather extensive chapter concludes with an introduction to single-phase AC power and the concepts of power factor, apparent, real and reactive power, power triangles and power factor correction.

In this chapter and throughout the book, angles are given in units of radians, unless indicated otherwise.

Learning Objectives

Students will learn to...

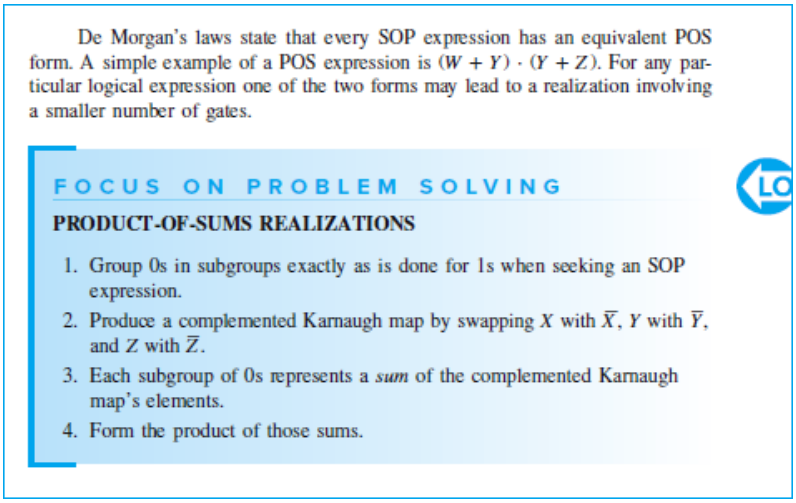
1. Compute current, voltage, and energy of capacitors and inductors. *Section 3.2.*
2. Calculate the average and effective (root-mean-square) value of an arbitrary periodic waveform. *Section 3.3.*
3. Convert time-domain sinusoidal voltages and currents to phasor notation, and vice versa; and represent circuits using impedances. *Sections 3.4 and 3.5.*
4. Apply DC circuit analysis methods to AC circuits in phasor form. *Section 3.6.*
5. Compute average AC power and the power factor of a complex load. *Section 3.7.*
6. Compute apparent, real and reactive power for complex loads and draw a power triangle. *Section 3.8.*
7. Compute the capacitance required to correct the power factor of a complex load. *Section 3.9.*

3.1 CIRCUITS CONTAINING ENERGY STORAGE ELEMENTS

The resistive circuits studied in Chapters 1 and 2 had no dependence on time. The sources had constant (DC) values and the *i-v* relationship for resistors (Ohm's law) had no time dependence. As a result, all the equations obtained in those chapters

Focus on Problem Solving sections summarize important methods and procedures for the solution of common problems and assist the student in developing a methodical approach to

problem solving.



De Morgan's laws state that every SOP expression has an equivalent POS form. A simple example of a POS expression is $(W + Y) \cdot (Y + Z)$. For any particular logical expression one of the two forms may lead to a realization involving a smaller number of gates.

FOCUS ON PROBLEM SOLVING

PRODUCT-OF-SUMS REALIZATIONS

1. Group 0s in subgroups exactly as is done for 1s when seeking an SOP expression.
2. Produce a complemented Karnaugh map by swapping X with $\overline{X}$, Y with $\overline{Y}$, and Z with $\overline{Z}$.
3. Each subgroup of 0s represents a *sum* of the complemented Karnaugh map's elements.
4. Form the product of those sums.

Page xviii

Clearly illustrated examples present relevant applications of electrical engineering principles. The examples are fully integrated with the Focus on "Problem" Solving material, and each one is organized according to a prescribed set of logical steps.

EXAMPLE 3.5 Calculating Inductor Current From Voltage

Problem

Use a time plot of the voltage across an inductor and its initial current to calculate the current through it as a function of time.

Solution

Known Quantities: Inductor voltage; initial condition (current at $t = 0$); inductance value.

Find: Inductor current.

Schematics, Diagrams, Circuits, and Given Data:

$$v(t) = \begin{cases} 0\,\text{V} & t < 0\,\text{s} \\ -10\,\text{mV} & 0 < t < 1\,\text{s} \\ 0\,\text{V} & t > 1\,\text{s} \end{cases}$$

$$L = 10\,\text{mH}; \quad i_L(t = 0) = I_0 = 0\,\text{A}$$

The voltage across the inductor is plotted in Figure 3.16(a).

Analysis: Use the integral i-v relationship for an inductor to obtain the current through it:

$$i_L(t) = i_L(t_0) + \frac{1}{L}\int_{t_0}^{t} v(\tau)\,d\tau \qquad t \geq t_0$$

$$= \begin{cases} I_0 + \frac{1}{L}\int_0^t (-10 \times 10^{-3})\,d\tau = 0 + \frac{-10^{-2}}{10^{-2}}t = -t\,\text{A} & 0 \leq t \leq 1\,\text{s} \\ -1\,\text{A} & t \geq 1\,\text{s} \end{cases}$$

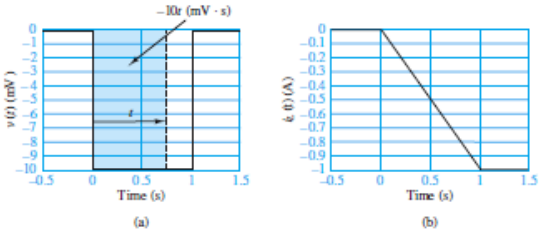


Figure 3.16

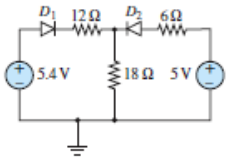
The inductor current is plotted in Figure 3.16b.

Comments: The inductor voltage can change instantaneously and thus be discontinuous!

Check Your Understanding exercises follow each set of examples and allow students to confirm their mastery of concepts.

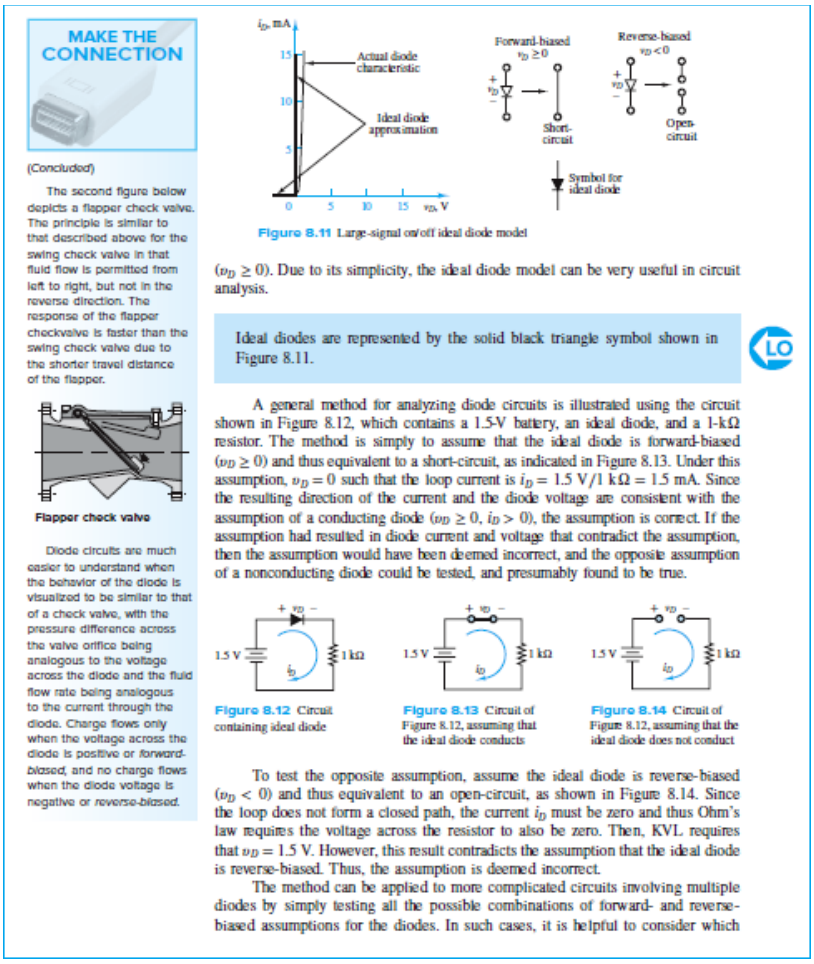
CHECK YOUR UNDERSTANDING

Determine which of the diodes conduct in the circuit shown below. Each diode has an offset voltage of 0.6 V.



Answer: Both diodes conduct.

Make the Connection sidebars present analogies that illuminate electrical engineering concepts using concepts from other engineering disciplines.



MAKE THE CONNECTION

(Concluded)

The second figure below depicts a flapper check valve. The principle is similar to that described above for the swing check valve in that fluid flow is permitted from left to right, but not in the reverse direction. The response of the flapper checkvalve is faster than the swing check valve due to the shorter travel distance of the flapper.

Flapper check valve

Diode circuits are much easier to understand when the behavior of the diode is visualized to be similar to that of a check valve, with the pressure difference across the valve orifice being analogous to the voltage across the diode and the fluid flow rate being analogous to the current through the diode. Charge flows only when the voltage across the diode is positive or *forward-biased*, and no charge flows when the diode voltage is negative or *reverse-biased*.

Figure 8.11 Large-signal on/off ideal diode model

($v_D \geq 0$). Due to its simplicity, the ideal diode model can be very useful in circuit analysis.

Ideal diodes are represented by the solid black triangle symbol shown in Figure 8.11.

LO

A general method for analyzing diode circuits is illustrated using the circuit shown in Figure 8.12, which contains a 1.5-V battery, an ideal diode, and a 1-kΩ resistor. The method is simply to assume that the ideal diode is forward-biased ($v_D \geq 0$) and thus equivalent to a short-circuit, as indicated in Figure 8.13. Under this assumption, $v_D = 0$ such that the loop current is $i_D = 1.5\ \text{V}/1\ \text{k}\Omega = 1.5\ \text{mA}$. Since the resulting direction of the current and the diode voltage are consistent with the assumption of a conducting diode ($v_D \geq 0$, $i_D > 0$), the assumption is correct. If the assumption had resulted in diode current and voltage that contradict the assumption, then the assumption would have been deemed incorrect, and the opposite assumption of a nonconducting diode could be tested, and presumably found to be true.

Figure 8.12 Circuit containing ideal diode

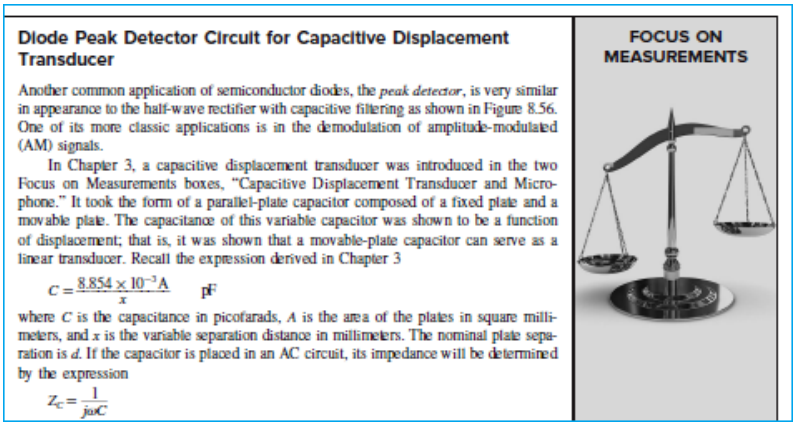
Figure 8.13 Circuit of Figure 8.12, assuming that the ideal diode conducts

Figure 8.14 Circuit of Figure 8.12, assuming that the ideal diode does not conduct

To test the opposite assumption, assume the ideal diode is reverse-biased ($v_D < 0$) and thus equivalent to an open-circuit, as shown in Figure 8.14. Since the loop does not form a closed path, the current i_D must be zero and thus Ohm's law requires the voltage across the resistor to also be zero. Then, KVL requires that $v_D = 1.5$ V. However, this result contradicts the assumption that the ideal diode is reverse-biased. Thus, the assumption is deemed incorrect.

The method can be applied to more complicated circuits involving multiple diodes by simply testing all the possible combinations of forward- and reverse-biased assumptions for the diodes. In such cases, it is helpful to consider which

Focus on Measurements boxes emphasize the great relevance of electrical engineering to the science and practice of measurement.



Diode Peak Detector Circuit for Capacitive Displacement Transducer

FOCUS ON MEASUREMENTS

Another common application of semiconductor diodes, the *peak detector*, is very similar in appearance to the half-wave rectifier with capacitive filtering as shown in Figure 8.56. One of its more classic applications is in the demodulation of amplitude-modulated (AM) signals.

In Chapter 3, a capacitive displacement transducer was introduced in the two Focus on Measurements boxes, "Capacitive Displacement Transducer and Microphone." It took the form of a parallel-plate capacitor composed of a fixed plate and a movable plate. The capacitance of this variable capacitor was shown to be a function of displacement; that is, it was shown that a movable-plate capacitor can serve as a linear transducer. Recall the expression derived in Chapter 3

$$C = \frac{8.854 \times 10^{-3} A}{x} \quad \text{pF}$$

where C is the capacitance in picofarads, A is the area of the plates in square millimeters, and x is the variable separation distance in millimeters. The nominal plate separation is d. If the capacitor is placed in an AC circuit, its impedance will be determined by the expression

$$Z_C = \frac{1}{j\omega C}$$

Affordability & Outcomes = Academic Freedom!

You deserve choice, flexibility and control. You know what's best for your students and selecting the course materials that will help them succeed should be in your hands.

Thats why providing you with a wide range of options that lower costs and drive better outcomes is our highest priority.

Mc Graw Hill connect®

Students—study more efficiently, retain more and achieve better outcomes. Instructors—focus on what you love—teaching.

They'll thank you for it.

Study resources in Connect help your students be better prepared in less time. You can transform your class time from dull definitions to dynamic discussion. Hear from your peers about the benefits of Connect at **www.mheducation.com/highered/connect**

Study anytime, anywhere.

Download the free ReadAnywhere app and access your online eBook when it's convenient, even if you're offline. And since the app automatically syncs with your eBook in Connect, all of your notes are available every time you open it. Find out more at **www.mheducation.com/readanywhere**

Learning for everyone.

McGraw Hill works directly with Accessibility Services Departments and faculty to meet the learning needs of all students. Please contact your Accessibility Services office and ask them to email accessibility@mheducation.com, or visit **www.mheducation.com/about/accessibility.html** for more information.

Learn more at: www.mheducation.com/realvalue

Rent It

Affordable print and digital rental options through our partnerships with leading textbook distributors including Amazon, Barnes & Noble, Chegg, Follett, and more.

Go Digital

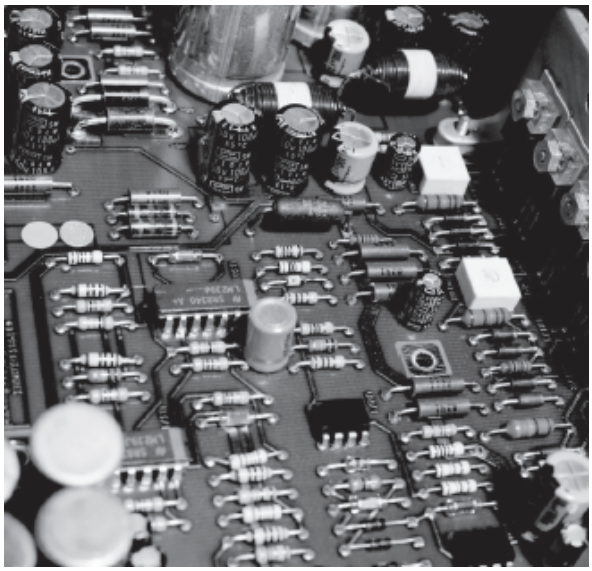
A full and flexible range of affordable digital solutions ranging from Connect, ALEKS, inclusive access, mobile apps, OER and more.

Get Print

Students who purchase digital materials can get a loose-leaf print version at a significantly reduced rate to meet their individual preferences and budget.

PRINCIPLES AND APPLICATIONS OF ELECTRICAL ENGINEERING

PART I
CIRCUITS



Arthur S. Aubry/Stockbyte/Getty Images

Page 3

C H A P T E R
1

FUNDAMENTALS OF ELECTRIC CIRCUITS

Chapter 1 is the foundation for the entire book and presents the fundamental laws that govern the behavior of electric circuits. Basic features and terminology of electric circuits, such as nodes, branches, meshes, and loops, are defined, and the three fundamental laws of circuit analysis, Kirchhoff's current and voltage laws and Ohm's law, are introduced. The concept of electric power and the passive sign convention are introduced along with basic circuit elements—sources and resistors. Basic analytic techniques of node voltage and mesh current analyses are introduced along with some engineering applications.

Page 4

Learning Objectives

Students will learn to...

1. Identify the principal *features of electric circuits or networks*: nodes, loops, meshes, and branches. *Section 1.1*.

2. Apply definitions of charge, current and voltage. *Section 1.2*.
3. Identify *sources* and their *i-v characteristics*. *Section 1.3*.
4. Apply the *passive sign convention* to compute the power consumed or supplied by circuit elements. *Section 1.4*.
5. Apply *Kirchhoff's laws* to simple electric circuits. *Section 1.5*.
6. Apply *Ohm's law* to calculate unknown voltages and currents in simple circuits *Section 1.6*.
7. Apply the *Node Voltage method* to solve for unknown voltages and currents in resistive networks. *Section 1.7*.
8. Apply the *Mesh Current method* to solve for unknown voltages and currents in resistive networks. *Section 1.8*.
9. Apply the Node Voltage and Mesh Current methods to solve for unknown voltages and currents in resistive networks with *dependent sources*. *Section 1.9*.

1.1 FEATURES OF NETWORKS AND CIRCUITS

"A *network* can be defined as a collection of interconnected objects. In an electric network, *elements*, such as resistors, are connected by wires. An electric *circuit* can be defined as an electric network within which at least one closed path exists and around which electric charge may flow. All electric circuits are networks but not all electric networks contain a circuit. In this book, a circuit is any network that contains at least one complete and closed path.

There are two principal quantities within a circuit: current and voltage. *The primary objective of circuit analysis is to determine one or more unknown currents and voltages.* Once these currents and voltages are determined, any other aspect of the circuit, such as its power requirements, efficiency, and speed of response, can be computed.

Two useful concepts for circuit analysis are those of a source and of a load. In general, the load is the circuit element or segment of interest to the designer or user of the circuit. By default, the source is everything else not included in the load. Typically, the source provides energy and the load consumes it for some purpose, such as the lifting of a weight. For example, consider the simple physical circuit of a headlight attached to a car battery as shown in Figure 1.1(a). For the driver of the car, the headlight may be the circuit element of interest since it enables the driver to see the road at night. From this perspective, the headlight is the load and the battery is the source as shown in Figure 1.1(b), which is intuitively appealing because power flows from the source (the battery) to the load (the headlight). However, in general, it is not required nor necessarily true that power flows in this manner. Electric power is discussed later in this chapter.

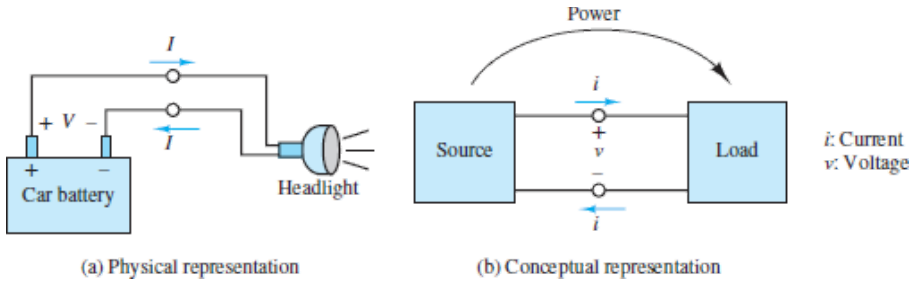


Figure 1.1 (a) Physical model and (b) generalized conceptual representation of an electrical system. See the notation rules for V, I, v, and i listed in the preface.

The use of the term *source* can be confusing at times because, as will be discussed later in this chapter, there are circuit elements known as *ideal voltage and current sources*, which have well-defined attributes and circuit symbols. These Page 5ideal sources, along with other circuit elements, are often the constituents of the source portion of a circuit, as well as the load portion. In this book, ideal sources are referred to as either voltage or current sources, explicitly, to avoid confusion.

Other key conceptual features of electric circuits are the *ideal wire*, *node*, *branch*, *loop*, and *mesh*. The concept of a node (see below) is particularly useful for correctly interpreting circuit diagrams and constructing circuit prototypes on breadboards.

Many students struggle with circuit analysis simply because they lack an organizing perspective with which to interpret circuit diagrams. One particularly helpful perspective is to see electric circuits as comprised of elements situated between nodes. This perspective enables students to see beyond the particular aesthetic presentation of a circuit diagram, to see its substance and not be fooled by its appearance. Once the concept of a node is well understood circuits that previously appeared complicated often become meaningful and clear.

Ideal Wire

Electric circuit and network *diagrams* are used to model actual electric circuits and networks. These diagrams contain *elements* connected by *ideal wires*. An ideal wire is able to conduct electric charge without any loss of electric potential. In other words, no work is required to move an electric charge along an ideal wire. Luckily, in many applications, actual wires are well approximated by ideal wires. However, there are applications where wiring accounts for significant losses of potential (e.g., long-distance transmission lines and microscopic integrated circuits). In these applications, the ideal wire approximation must be used with care. In this book, all wires in circuit and network diagrams are ideal, unless indicated otherwise.

Node

A **node** consists of one or more ideal wires connected together such that an electric charge can travel between any two points on the node without traversing a circuit element, such as a resistor. Thus, every point on a node has the same electric potential, which is known as the *node voltage* and its value is relative to a reference potential.

A *junction* is a point where two or more wires are joined together. A node may contain one or more junctions or none at all, such as when a single wire directly connects two elements. A junction is part of a node but is not a node itself.

Page 6

It is crucial to correctly identify and count nodes in the analysis of electric circuits. Figure 1.2 illustrates a helpful way to mark nodes. There are three nodes in Figure 1.2(a) and two nodes in Figure 1.2(b). It is sometimes convenient to use the concept of a **supernode**, which is simply a closed boundary enclosing two or more nodes, as shown in Figure 1.2(c).

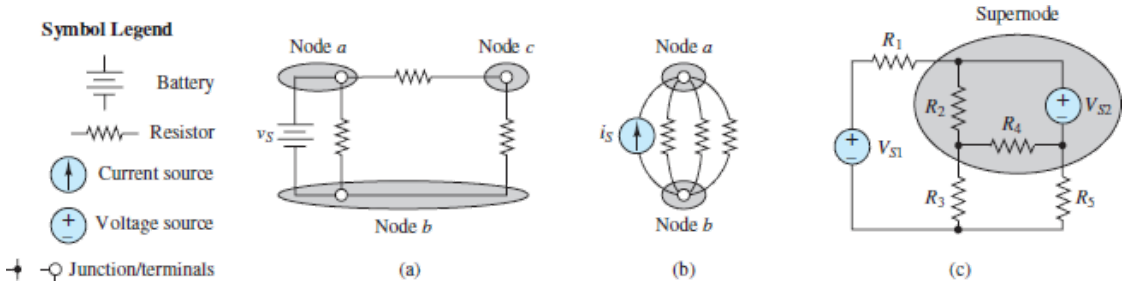
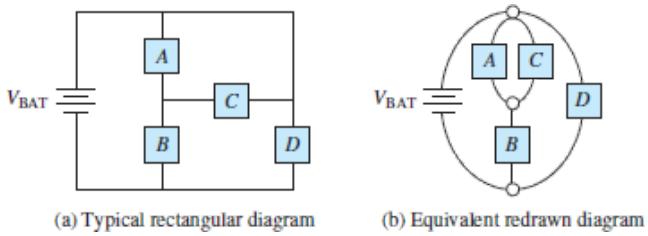


Figure 1.2 Illustrating nodes and supernodes in circuit diagrams

It is also important to realize that since no work is required to move an electric charge along an ideal wire, the length and shape of an ideal wire has no impact on the behavior of a circuit. Likewise, since nodes are comprised of ideal wires, the extent and shape of a node has no impact on the behavior of a circuit. As a result, a node may be redrawn in any manner as long as the newly drawn node is attached to the same elements as the original node. Circuit diagrams are typically drawn, by convention, in a rectangular manner, with all wires drawn either side to side or up and down. However, many students find it helpful to redraw circuits so as to clarify the number and location of nodes in a circuit. Figure 1.3 shows two identical circuits drawn in two different ways. Can you tell that these circuits have the same number of nodes?



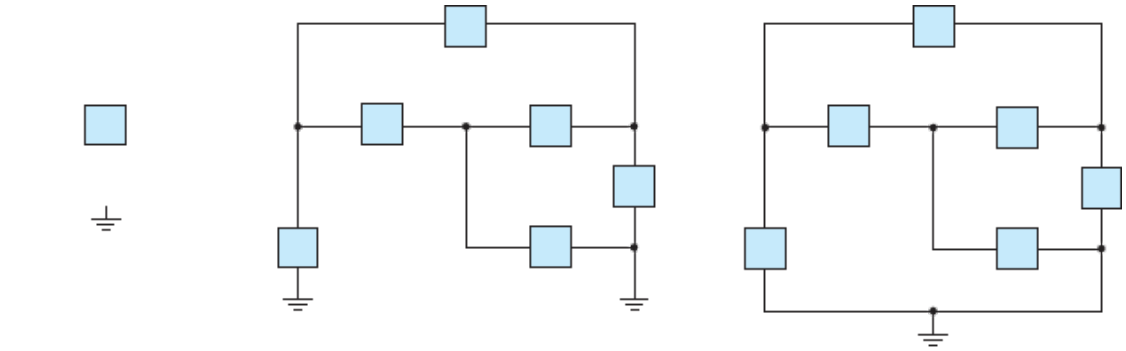
(a) Typical rectangular diagram (b) Equivalent redrawn diagram

Figure 1.3 (a) A typical rectangular circuit diagram and (b) an equivalent redrawn diagram. A circuit can be redrawn to have almost any appearance; however, the behavior of the circuit is unchanged as long as the number of nodes and the elements between those nodes remain unchanged.

Keep in mind that all forms of potential, including voltage, are relative quantities. For this reason, it is important to refer to the voltage *across* an element. In circuit diagrams, the voltage across an element is indicated by the paired symbols + and −. Taken together as a single symbol they indicate the *assumed polarity* of the voltage.

Sometimes it is convenient to establish a *reference node.* Any one node in a network can be designated as the reference node. Then, all other node voltages are Page 7determined relative to that reference node. The value of the reference node can be chosen freely, although a value of zero is usually chosen, for simplicity. It is often true that a smart choice of reference node will simplify the analysis that follows. A good rule of thumb is to select a node that is connected to a large number of elements.

A reference node is designated by the symbol shown in Figure 1.4(a). This symbol is also used to designate *earth ground* in applications. To reduce the apparent complexity of some circuits, multiple reference symbols are used to minimize the amount of displayed reference node wiring. It is simply understood that all nodes to which these symbols are attached are, in fact, connected by ideal wires and therefore part of one large reference node. Figure 1.4(b) and (c) illustrate this practice.



(a) Generic element and reference node symbol (b) Circuit with two reference node symbol (c) Equivalent circuit

Figure 1.4 There can be one and only one reference node in a network although the reference node symbol may appear more than once in order to reduce the amount of displayed reference node wiring. The reference node symbol is also used to designate a connection to earth ground in practical circuits.

Elements that sit between the same two nodes are said to be in *parallel.*

Branch

A **branch** is defined in this book as a single electrical pathway, consisting of wires and elements. A branch may contain one or more circuit elements as shown in Figure 1.5. By definition, the current *through* any one element in a branch is the Page 8same as the current through every other element in that branch; that is, there is one current in a branch, the *branch current.*

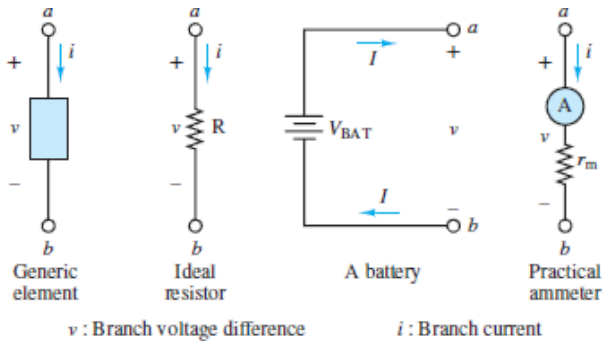


Figure 1.5 Examples of circuit branches

Elements that sit along the same branch are said to be in *series.*

Loop

A **loop** is any closed pathway. Figure 1.6(a) shows that different loops in the same circuit may share common elements and branches. It is interesting, and perhaps initially confusing, to note that a loop does not necessarily have to correspond to a closed electrical pathway, consisting of wires and elements. Figure 1.6(b) shows one example in which a loop passes directly from node *a* to node *c*.

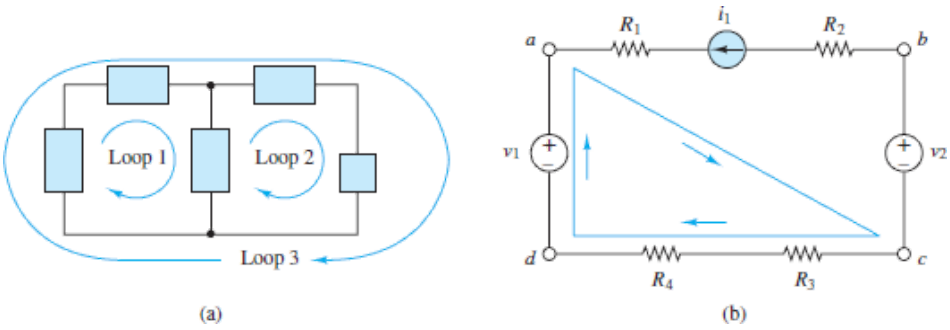


Figure 1.6 A skydiver understands all too well that her fate is unchanged by the choice of reference potential.

LO

Mesh

A **mesh** is a closed electrical pathway that does not contain other closed electrical pathways. In Figure 1.6(a), loops 1 and 2 are meshes, but loop 3 is not a mesh because it contains loops 1 and 2. The circuit in Figure 1.6(b) has one mesh. Figure 1.7 illustrates how simple it is to visualize meshes.

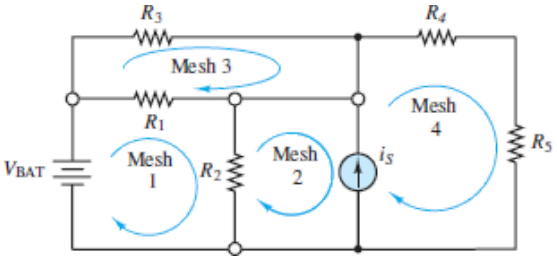


Figure 1.7 Circuit with four meshes. How many different closed electrical pathways are in this circuit? [Answer: 14]

EXAMPLE 1.1

Problem

Identify the branch and node voltages and the loop and mesh currents in the circuit of Figure 1.8.

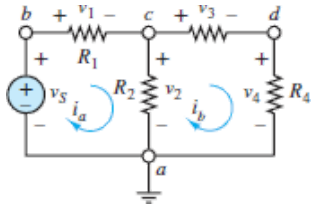


Figure 1.8

Solution

The following node and branch voltages may be identified:

Node voltages	Branch voltages	Relationship
$v_a = 0$ (reference)		
v_b	v_S	$v_S = v_b - v_a$
	v_1	$v_1 = v_b - v_c$
v_c	v_2	$v_2 = v_c - v_a$
	v_3	$v_3 = v_c - v_d$
v_d	v_4	$v_4 = v_d - v_a$

Comments: Currents i_a and i_b are mesh currents.

EXAMPLE 1.2 Counting Nodes in a Network

Problem

Count the total number of nodes in each of the four networks.

Solution

Known Quantities: Wires and elements.

Find: The number of nodes in each network diagram.

Schematics, Diagrams, Circuits, and Given Data: Figure 1.9 contains four elements: two resistors and two ideal voltage sources, one independent and one dependent. Figure 1.10 contains five elements: four resistors and one independent ideal voltage source. Figure 1.11 contains five elements: four resistors and one operational amplifier. Figure 1.12 contains three elements: two headlamps and one 12-V battery.

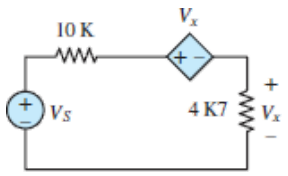


Figure 1.9

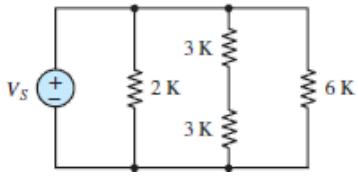


Figure 1.10

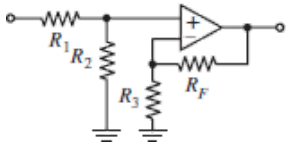


Figure 1.11

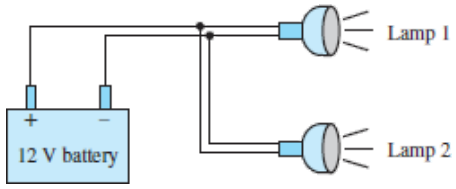


Figure 1.12

Assumptions: All wires are ideal.

Analysis:

In Figure 1.9, all four elements are in a single electrical loop. There is one node between each pair of elements. Thus, there are *four nodes* in this network.

In Figure 1.10, the voltage source and the 2K and 6K resistors sit between two large nodes, one along the top of the network and the other along the bottom of the network. Each of these nodes contains two wire junctions. Another node is between the two 3K resistors. Thus, there are *three nodes* in this network.

In Figure 1.11, there is one node to the left of R_1 and one node to the right of the operational amplifier and R_F. A third node is located between R_1, R_2, and the +

(noninverting) terminal of the operational amplifier. A fourth node is located between R_3, R_F, and the − (inverting) terminal of the operational amplifier. A fifth, and final, node is the reference node, which is between R_2 and R_3. Thus, there are *five nodes* in this network.

In Figure 1.12, there is one node between the positive + battery terminal and one terminal on each headlamp. There is another node between the negative − battery terminal and the second terminal on each headlamp. Thus, there are *two nodes* in this network.

Comments: The notation K is short for kΩ. The placement of K indicates the location of a decimal point. For example, 4K7 is 4.7 kΩ.

1.2 CHARGE, CURRENT AND VOLTAGE

The earliest accounts of electricity date from about 2,500 years ago, when it was discovered that a piece of amber was capable of attracting very light objects, such as feathers. The word *electricity* originated about 600 B.C.; it comes from *elektron*, which was the ancient Greek word for amber. Following the work of Alessandro Volta and his invention of the copper-zinc battery, it was determined that static electric effects and the current in metal wires connected to a battery were both due to the same fundamental nature of matter, namely, the atomic structure of matter, consisting of a nucleus—neutrons and protons—surrounded by electrons.

Charles Coulomb (1736–1806).
(INTERFOTO/Personalities/Alamy Stock Photo)

The unit of charge, the **coulomb (C)**, is named after Charles Coulomb. The unit of current, the ampere (A), is named after the French scientist André-Marie Ampère.

The fundamental electric quantity is **charge**. The unit of charge is the **coulomb (C)**. The electron and proton each carry one unit of charge but of opposite sign. By convention, the electron is deemed to be negatively charged.

$$q_e = -1.602 \times 10^{-19}\text{C} \qquad q_p = +1.602 \times 10^{-19}\text{C} \tag{1.1}$$

Electrons and protons are often referred to as **elementary charges**. The amount of charge associated with an electron may seem rather small. However, typical currents involve the flow of large numbers of charged particles.

Current

Electric current is defined as the rate at which charge passes through an area, such as the cross-sectional area of a wire. Figure 1.13 depicts a macroscopic view of Page 11current i in a wire, where Δq units of charge flow through the cross-sectional area A in a period Δt. The resulting current i is

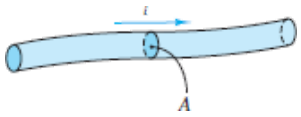


Figure 1.13 Current in an electric conductor is defined as the net flow rate of charge through the cross-sectional area A.

$$i \equiv \frac{\Delta q}{\Delta t} \quad \frac{\text{C}}{\text{s}} \tag{1.2}$$

The arrow symbol associated with the current i is its *assumed* direction through the wire segment. A negative value for i would indicate a direction opposite to the assumed direction. When large numbers of discrete charges cross A in a very small period, the current i can be written in differential form.

$$i \equiv \frac{dq}{dt} \quad \frac{\text{C}}{\text{s}} \tag{1.3}$$

The unit of current is the **ampere**, where 1 ampere (A) = 1 coulomb/second (C/s). By convention, in electrical engineering positive current is the direction of positive charge flow. This convention can be confusing since the mobile charge carriers in metal wires and many other conductors are electrons from the *conduction band* of the material. However, when an electron travels in one direction the effect on the distribution of *net charge* is the same as if a proton had travelled in the opposite direction. In other words, positive current represents the *relative* flow of positive charges.

Voltage

Typically, work is required to move charge between two nodes in a circuit. The total *work per unit charge* is called **voltage**, and the unit of voltage is the **volt** in honor of

Alessandro Volta.

$$1 \text{ volt (V)} = 1\frac{\text{joule (J)}}{\text{coulomb (C)}}$$

The voltage, or **potential difference**, across two nodes in a circuit is the energy (in joules) per unit charge (1 coulomb) needed to move charge from one node to the other. The direction, or *polarity*, of the voltage is related to whether energy is being gained or lost by the charge in the process.

Note that the word *potential* is quite appropriate as a synonym of voltage, in that voltage is the potential energy per unit charge across two nodes in a circuit. If the lightbulb is disconnected from the circuit, a voltage v_{ab} still exists across the battery terminals, as illustrated in Figure 1.14. This voltage represents work done on the battery to separate positive ions from negative ions. The potential energy associated with the separated ions is available to do work on an external element attached to the battery terminals. That work is expressed as positive charge flowing (current) through the element from high to low potential. Thus, the battery is able to *supply* energy to the attached element and, likewise, the attached element is able to *consume* or *dissipate* energy from the battery.

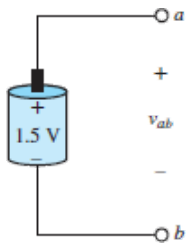


Figure 1.14 The voltage v_{ab} across the open terminals of the battery represents the potential energy available to move charge from *a* to *b* once a closed circuit is established.

The Reference Node and Ground

Earth ground represents a specific, and usually clearly marked, node in many circuits. Residential electric circuits are connected to earth ground through a large conductor, such as a copper spike or water pipe, that is buried in the earth. When present in a circuit diagram, earth ground is always chosen as the reference node Page 12because the earth's potential is relatively stable and uniform due to its ability to store and distribute large quantities of charge. In circuits where earth ground is not present, some other relatively large conductor can serve as a stable ground node, such as a metal enclosure or chassis of an instrument.

In practice, the voltage value assigned to a reference node, such as earth ground, while typically zero, is not consequential. A simple analogy with fluid flow illustrates this rule. Consider a tank of water, as shown in Figure 1.15, located at a certain height above the ground. The potential energy difference per unit mass due to gravity $u_{12} = g(h_1 - h_2)$ is completely analogous to the potential energy difference per unit charge $v_a - v_b$. Now assume that the height h_3 at ground level is chosen to be the zero potential energy reference. Is the flow of water in the pipe changed due to this choice? Of course not. Is the flow of water in the pipe dependent upon the height $h_2 - h_3$ of the support structure? Again, the answer is no. The truth of these statements is demonstrated by rewriting the *head* of the water tank $h_1 - h_2$ as $(h_1 - h_3) - (h_2 - h_3)$ such that

$$u_{12} = g(h_1 - h_2) = [g(h_1 - h_3)] - [g(h_2 - h_3)] = u_{13} - u_{23}$$

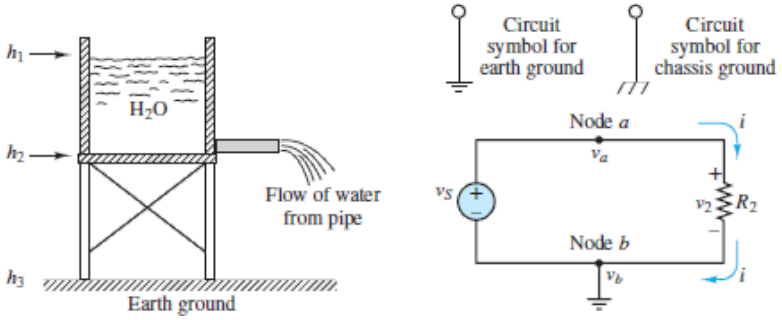


Figure 1.15 An analogy between water flow and electric current illustrates the relation between potential differences and a ground reference potential.

Even though the values of u_{13} and u_{23} depend upon h_3, the difference between them u_{12} does *not* depend upon h_3. It is the change in potential energy that matters in the water tank problem. So it is with electric circuits. The voltage *across* an element does not depend upon the selection of a reference node nor upon the arbitrary voltage value assigned to the reference node.

Another familiar scenario is that of a skydiver leaping from an airplane and parachuting to the surface below (see Figure 1.16). To quantify the potential energy U of the skydiver it is first necessary to choose a reference height h_0 such that $U = mg\Delta h = mg(h - h_0)$, where h represents the height of the skydiver. One possible choice for a reference height is the height of the airplane such that the potential energy of the skydiver is negative ($U < 0$)? However, such a choice would be strange and perhaps misleading. The surface of the earth is a more meaningful reference to the skydiver, who knows that a soft landing depends upon dissipating most of the initial potential energy through collisions with air molecules rather than through a collision with the surface. The skydiver knows that her fate is unchanged by her

choice of reference; however, some choices are more meaningful than others. So it often is with electric circuits.

Figure 1.16 A skydiver understands all too well that her fate is unchanged by the choice of reference potential

EXAMPLE 1.3 Charge and Current in a Conductor

Problem

Find the total charge in a cylindrical conductor (solid wire) and compute the current through the wire.

Solution

Known Quantities: Conductor geometry, charge density, charge carrier velocity.

Find: Total charge of carriers Q; current in the wire I.

Schematics, Diagrams, Circuits, and Given Data:

Conductor length: $L = 1$ m.

Conductor diameter: $2r = 2 \times 10^{-3}$ m.

Charge density: $n = 10^{29}$ carriers/m^3.

Charge of one electron: $q_e = -1.602 \times 10^{-19}$.

Charge carrier average net speed: $u = 19.9 \times 10^{-6}$ m/s.

Assumptions: None.

Analysis: To compute the total charge in the conductor, first determine the volume of the conductor:

$$\text{Volume} = \text{length} \times \text{cross-sectional area}$$

$$\text{Vol} = L \times \pi r^2 = (1\text{ m})\left[\pi\left(\frac{2 \times 10^{-3}}{2}\right)^2 \text{m}^2\right] = \pi \times 10^{-6}\text{ m}^3$$

Next, compute the number of carriers (electrons) in the conductor and the total charge:

$$\text{Number of carriers} = \text{volume} \times \text{carrier density}$$

$$N = \text{Vol} \times n = (\pi \times 10^{-6}\ \text{m}^3)\left(10^{29}\ \frac{\text{carriers}}{\text{m}^3}\right) = \pi \times 10^{23}\ \text{carriers}$$

$$\text{Charge} = \text{number of carriers} \times \text{charge/carrier}$$

$$Q = N \times q_e = (\pi \times 10^{23}\ \text{carriers}) \times \left(-1.602 \times 10^{-19}\ \frac{\text{C}}{\text{carrier}}\right) = -50.33 \times 10^3\ \text{C}$$

To compute the current, consider the average net speed of the charge carriers and the charge density per unit length of the conductor:

$$\text{Current} = \text{carrier charge density per unit length} \times \text{carrier average net speed}$$

$$I = \left(\frac{Q}{L}\ \frac{\text{C}}{\text{m}}\right) \times \left(u\ \frac{\text{m}}{\text{s}}\right) = \left(-50.33 \times 10^3\ \frac{\text{C}}{\text{m}}\right)\left(19.9 \times 10^{-6}\ \frac{\text{m}}{\text{s}}\right) = -1\ \text{A}$$

Comments: Charge carrier density is a function of material properties. Carrier average net speed is a function of the applied electric field.

Page 14

1.3 *i-v* CHARACTERISTICS AND SOURCES

It is possible to create an *i-v* plot for any circuit element. The functional relationship between i and v may be quite complex and not easily expressed in a closed mathematical form, such as $i = f(v)$ However, the plot of the ***i-v* characteristic** (or **volt-ampere characteristic**) for most circuit elements is either known or can be determined experimentally.

For example, consider the incandescent (tungsten filament) lightbulb shown in Figure 1.17(a). The *i-v* characteristic of the lightbulb can be determined by varying the voltage over some predetermined range and recording the resulting current for each particular voltage in that range. The plot of the *i-v* data will be similar to that shown in Figure 1.17(b). A positive voltage across the bulb results in a positive current through it, and conversely, a negative voltage across the bulb results in a negative current through it. In both cases charge flows from high to low potential, releasing energy that is dissipated by the bulb as light and heat.

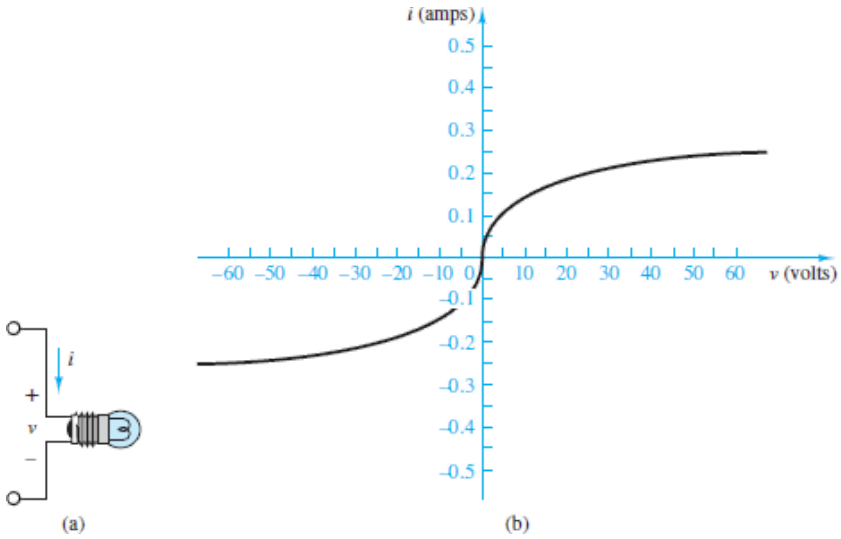


Figure 1.17 (a) Depiction of how to measure the *i*-*v* characteristic of an incandescent (tungsten filament) lightbulb; (b) typical *i*-*v* plot of such a lightbulb

The *i*-*v* characteristics of ideal voltage and current sources are simple yet helpful visual aids. An ideal source is one that can provide any amount of energy without affecting the behavior of the source itself. **Ideal sources** are divided into two types: voltage sources and current sources.

Ideal Voltage Sources

An **ideal voltage source** generates a prescribed voltage across its terminals independent of the current through its terminals. The circuit symbol for an ideal voltage source is shown in Figure 1.18(a). Notice that the current is defined as being directed from low to high potential. In other words, the voltage source is *supplying* energy to the flowing charge.

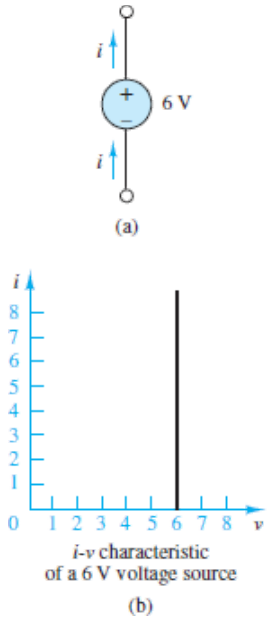


Figure 1.18 (a) An ideal voltage source; and (b) a typical *i-v* characteristic, which indicates energy is supplied by the source to the flowing charge

Page 15

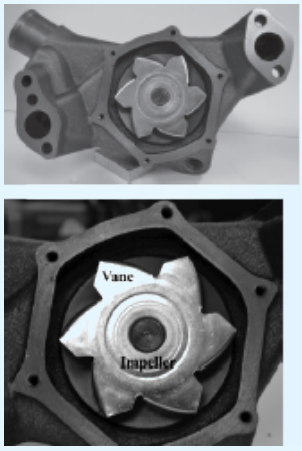
A typical *i-v* characteristic is shown in Figure 1.18(b). The current supplied by the source is determined by the circuit connected to it. It is important to recognize that an ideal voltage source guarantees a particular *change* in voltage from the node attached to its − terminal to the node attached to its + terminal. (Note: The + and − polarity markers do *not* indicate positive and negative voltage values relative to some zero reference. Do not make this mistake when solving problems!)

An ideal voltage source provides a prescribed voltage across its terminals independent of the current through those terminals. The amount of current through the source is determined by the circuit connected to it.

MAKE THE CONNECTION

Hydraulic Analog of a Voltage Source

The role played by a voltage source in an electric circuit is very similar to that played by a velocity pump in a hydraulic circuit. In a velocity or roto-dynamic pump, such as a centrifugal pump, impeller vanes add kinetic energy (velocity) to the fluid flow. This increase in kinetic energy is translated to an increase in pressure *across* the pump. The pressure difference *across* the pump is analogous to the voltage, or potential difference, *across* the voltage source.



A centrifugal pump (*Giorgio Rizzoni*).

Various types of batteries, electronic power supplies, and function generators approximate ideal voltage sources when used in proper circumstances. However, all such real devices have limits on the amount of current that can be supplied without impacting the voltage across the source. This behavior can be seen in a typical 12-V car battery. A digital voltmeter can be used to observe the voltage across a car battery as various electrical devices in the car are turned on and off. Very little change in the battery voltage will be observed, even when power windows are engaged. However, when the car is started, the battery voltage will drop significantly during the short period needed for the engine to start.

Figure 1.19 depicts various symbols for voltage sources. The output voltage of an ideal source can be a function of time. In this book, unless otherwise noted, a generic voltage source is denoted by a lowercase v. If it is necessary to emphasize that the source produces a time-varying voltage, then the notation $v(t)$ is employed. Finally, a constant, or dc voltage source is denoted by the uppercase character V.

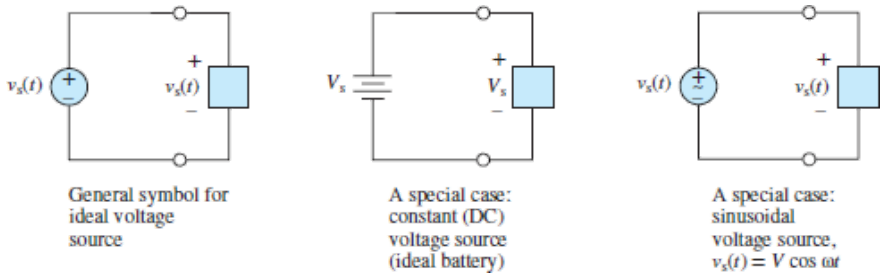


Figure 1.19 Three common ideal voltage sources

Ideal Current Sources

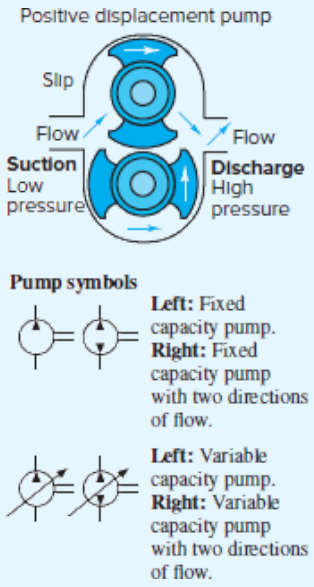
An **ideal current source** generates a prescribed current through its terminals independent of the voltage across its terminals. The circuit symbol for an ideal current source is shown in Figure 1.20(a). Notice that the current is defined as being Page 16directed from low to high potential. In other words, the current source is *supplying* energy to the flowing charge.

MAKE THE CONNECTION

Hydraulic Analog of a Current Source

The role played by a current source in an electric circuit is very similar to that of a positive displacement pump in a hydraulic circuit. In a positive displacement pump, such as a peristaltic or reciprocating pump, an internal mechanism, such as a roller,

piston, or diaphragm, forces a particular volume of fluid to be pumped *through* a hydraulic line. The volume flow rate *through* the pump is analogous to the charge flow rate *through* the current source.



A typical *i-v* characteristic is shown in Figure 1.20(b). The voltage *across* the current source is determined by the circuit connected to it. It is important to recognize that an ideal current source guarantees a particular current *through* its terminals, such that the current entering the − terminal is the same as the current exiting the + terminal. (Again, the + and − polarity markers do not indicate positive and negative voltage values relative to some zero reference. Do not make this mistake when solving problems!)

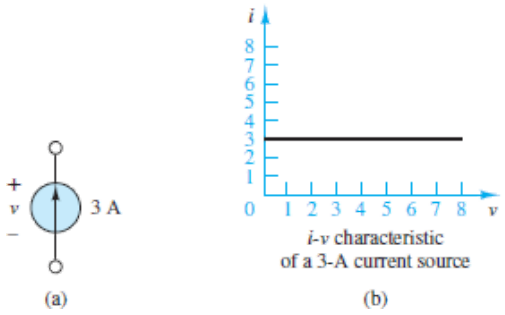


Figure 1.20 (a) An ideal current source; and (b) a typical *i-v* characteristic, which indicates energy is supplied by the source to the flowing charge

An ideal current source provides a prescribed current through its terminals independent of the voltage across those terminals. The amount of voltage

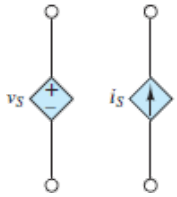
across the source is determined by the circuit connected to it.

Practical approximations to ideal current sources are not as common nor numerous as those for ideal voltage sources. However, in general, an ideal voltage source in series with an output resistance that is large in comparison to the input resistance of the circuit attached to its terminals provides a nearly constant current and thus approximates an ideal current source. A battery charger is a common and approximate example of an ideal current source.

Dependent (Controlled) Sources

The ideal *independent* sources described above are able to generate a prescribed voltage or current independent of the circuit attached to its terminals. Another category of sources, whose output (current or voltage) depends on some other voltage or current in a circuit, is known as **dependent** (or **controlled**) **sources**. As shown in Figure 1.21, the circuit symbols for these sources are diamonds to distinguish them from independent sources. The table illustrates the relationship between the source voltage v_S or source current i_S and the circuit voltage v_x or circuit current i_x, which they depend upon and which can be any voltage or current elsewhere in the overall circuit.

Page 17



Source type	Relationship
Voltage controlled voltage source (VCVS)	$v_S = \mu v_x$
Current controlled voltage source (CCVS)	$v_S = r i_x$
Voltage controlled current source (VCCS)	$i_S = g v_x$
Current controlled current source (CCCS)	$i_S = \beta i_x$

Figure 1.21 Symbols for dependent sources

Dependent sources are very useful in describing the behavior of transistors and other electronic devices.

1.4 POWER AND THE PASSIVE SIGN CONVENTION

The power supplied or dissipated by a circuit element can be represented by the following relationship:

$$\text{Power} = \frac{\text{work}}{\text{time}} = \frac{\text{work}}{\text{charge}} \frac{\text{charge}}{\text{time}} = \text{voltage} \times \text{current} \tag{1.4}$$

Thus,

Electric power, P, is the product of voltage *across* an element and current *through* it.

$$\boxed{p = vi} \tag{1.5}$$

The unit of voltage (joules per coulomb) multiplied by that of current (coulombs per second) equals the unit of power (joules per second, or watts).

The power associated with a circuit element can be positive or negative. Positive power is, by convention, the rate at which energy is transferred from the flowing charge to an element. Negative power implies energy is transferred by an element to the flowing charge. Consider Figure 1.22(a), in which electric charge flows from low to high potential. Clearly, work has been done *by* element *A on* the flowing charge as its potential is raised. The rate at which this work is done *by* element *A* is its power. In this case, power is considered negative when energy is *supplied* or *released* by the element. In Figure 1.22(b), charge flows from high to low potential. Here, work has been done *on* element *B by* the flowing charge as its potential is lowered. The rate at which this work is done *on* element *B* is its power. In this case, power is considered positive when energy is *dissipated* or *stored* by the element.

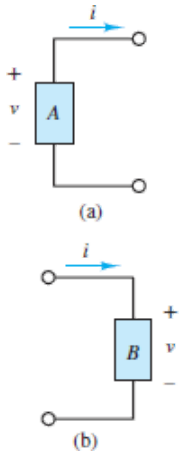


Figure 1.22 Assuming positive values for i and υ, the active sign convention shown in (a) implies energy is supplied or released by element A while the passive sign convention shown in (b) implies energy is consumed or stored by element B.

In the *passive sign convention*, current is directed from high to low potential. In this convention, energy is released by the flowing charge and consumed or stored by the element. The rate at which energy is transferred from the flowing charge to the element is considered positive power.

Page 18

Figure 1.23 shows the four quadrants of a generic i-υ plot, where i and υ are assumed to observe the passive sign convention. Power is positive in the first and third quadrants; negative in the second and fourth quadrants. The i-υ plot of a typical incandescent lightbulb shown in Figure 1.17(b) reveals that its power is always positive. In other words, the lightbulb always dissipates energy.

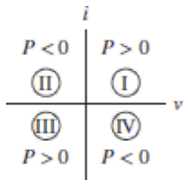


Figure 1.23 The four quadrants of a generic i-υ plot, where i and υ are assumed to observe the passive sign convention

Passive elements are defined as those that do not require an external source of energy to *enable* them. Common passive elements are resistors, capacitors, inductors, diodes, and electric motors. Passive elements can dissipate energy (e.g., resistors) and/or store and release energy (e.g., capacitors and inductors).

Active elements, on the other hand, are defined as those that do require an external source of energy to be enabled. Common active elements are transistors, amplifiers, and voltage and current sources. There are electronic devices that can operate either as passive or active elements. For example, a photodiode can act either as a light sensor (passive element) or as a solar cell (active element).

The electrical engineering community has uniformly adopted the passive sign convention. All the constitutive laws (e.g., Ohm's law) introduced in this book assume that convention. It is often necessary to assume directions for unknown currents and/or assume polarities for unknown voltages when solving circuit problems. It is important that these assumptions be made in accord with the passive sign convention. As long as the passive sign convention is observed it is not necessary to foresee actual current directions nor actual voltage polarities. Instead, when a current direction or voltage polarity is assumed incorrectly, the solution will yield a negative result, indicating that the assumed direction or polarity is opposite the actual.

FOCUS ON PROBLEM SOLVING

THE PASSIVE SIGN CONVENTION

1. Assign a current through each passive element. The direction of each current be assumed arbitrarily.
2. For each *passive* element, assign a voltage across the element such that assigned current through the element is directed from high to low poter Other valid descriptions are that current enters the + terminal or exits th terminal of the element.
3. The power associated with each passive element is equal to vi. Positive po indicates that the element is either dissipating or storing energy.

EXAMPLE 1.4 Use of the Passive Sign Convention

Problem

Apply the passive sign convention to solve for the voltages and *mesh current* in the circuit of Figure 1.24.

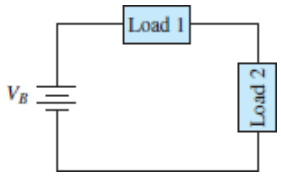


Figure 1.24

Solution

Known Quantities: Voltage of the battery and the power dissipated by elements 1 and 2.

Find: Mesh current and the voltage across each load.

Schematics, Diagrams, Circuits, and Given Data: Figure 1.25(a) and (b). The voltage of the battery is $V_B = 12\text{V}$. The power dissipated by element 1 is $P_1 = 0.8\text{ W}$ and by element 2 is $P_2 = 0.4\text{ W}$..

Assumptions: None.

Analysis: This problem can be solved using the passive sign convention in two different approaches. The first approach assumes a clockwise mesh current, while the second approach assumes a counterclockwise current. For either approach, the passive sign convention is used to label the change in voltage across each load. Figure 1.25(a) and (b) show the result of these two approaches. Notice that the change in voltage across each element was chosen so that the assumed current is directed from high to low potential.

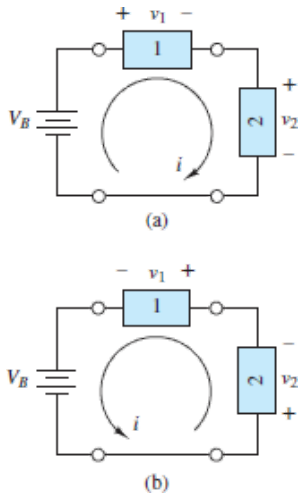


Figure 1.25

The polarity of the battery is indicated by the alternating sequence of long and short bars. The positive and negative terminals of the battery are connected to a long and short bar, respectively.

A four-step solution using the first approach, as depicted in Figure 1.25(a), is given below.

1. Assume a clockwise direction for the current.
2. Label the change in voltage across each (passive) element so that the current is directed from high to low potential.
3. Express the power dissipated by each element using the relation $P = vi$, which is valid when the passive sign convention is observed.

$$P_1 = v_1 i = 0.8\,\text{W}$$
$$P_2 = v_2 i = 0.4\,\text{W}$$

   The power associated with the battery is expressed as $P_B = -V_B i$, which requires a negative sign $-vi$ because the current through the battery is directed from low to high potential, opposite of the passive sign convention.
4. Conservation of energy requires that the total power associated with the circuit be zero. Thus,

$$P_1 + P_2 + P_B = 0$$
$$P_B = -P_1 - P_2 = -0.8\,\text{W} - 0.4\,\text{W} = -1.2\,\text{W} = -V_B i$$

   It is now possible to use the three vi equations to solve for the three unknown variables i, v_1, and v_2. Since $V_B = 12$ V, the current i is:

$$i = \frac{-1.2\,\text{W}}{-12\,\text{V}} = 0.1\,\text{A}$$

As a result, the change in voltage across each element is:

$$v_1 = \frac{0.8\ \text{W}}{0.1\ \text{A}} = 8\,\text{V}$$
$$v_2 = \frac{0.4\ \text{W}}{0.1\ \text{A}} = 4\,\text{V}$$

A four-step solution using the second approach, as depicted in Figure 1.25(b), is given below.

1. Assume a counterclockwise direction for the current.
2. Label the change in voltage across each (passive) element so that the current is directed from high to low potential.

3. Express the power dissipated by each element using the relation $P = vi$, which is valid when the passive sign convention is observed.

$$P_1 = v_1 i = 0.8\,\text{W}$$
$$P_2 = v_2 i = 0.4\,\text{W}$$

The power associated with the battery is expressed here as $P_B = +V_B i$, which now requires a positive sign $+vi$ because the current through the battery is directed from high to low potential.

4. Conservation of energy requires that the total power associated with the circuit be zero. Thus,

$$P_1 + P_2 + P_B = 0$$
$$P_B = -P_1 - P_2 = -0.8\,\text{W} - 0.4\,\text{W} = -1.2\,\text{W} = V_B i$$

It is now possible to use the three vi equations to solve for the three unknown variables i, v_1, and v_2. Since $V_B = 12$ V, the current i is:

$$i = \frac{-1.2\,\text{W}}{12\,\text{V}} = -0.1\,\text{A}$$

As a result, the change in voltage across each element is:

$$v_1 = \frac{0.8\ \text{W}}{-0.1\ \text{A}} = -8\,\text{V}$$
$$v_2 = \frac{0.4\ \text{W}}{-0.1\ \text{A}} = -4\,\text{V}$$

Comments: Notice that the *actual* current present in the circuit and the *actual* change in voltage across each element is the same for each solution approach. For instance, using the first approach the current was found to be 0.1 A clockwise, while using the second approach the current was found to be −0.1 A counterclockwise. The negative sign found for the current in the second approach indicates that the actual current is directed clockwise, not counterclockwise. This example provides a good demonstration of the fact that it is not necessary to foresee the actual direction of unknown currents and voltages when solving a circuit problem. The important point is to observe the passive sign convention.

Also note that conservation of energy is required for electric circuits, just as it is for any other physical system. For electric circuits: *Power supplied always equals power consumed.*

EXAMPLE 1.5 Power Calculations

Problem

For the circuit shown in Figure 1.26, determine which components are consuming power and which are supplying power. Is conservation of power satisfied? Explain your answer.

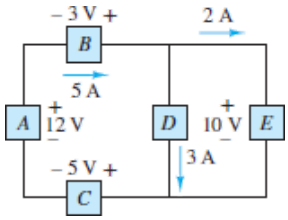


Figure 1.26

Solution

Known Quantities: All currents and voltages.

Find: Which components are consuming power, and which are supplying power? Verify conservation of power.

Page 21

Analysis: The power associated with each element can be computed using $P = vi$ when the passive sign convention is observed or $P = -vi$ when it is not observed.

$$P_A = -(12\,\text{V})(5\,\text{A}) = -60\,\text{W}$$
$$P_B = -(3\,\text{V})(5\,\text{A}) = -15\,\text{W}$$
$$P_C = (5\,\text{V})(5\,\text{A}) = 25\,\text{W}$$
$$P_D = (10\,\text{V})(3\,\text{A}) = 30\,\text{W}$$
$$P_E = (10\,\text{V})(2\,\text{A}) = 20\,\text{W}$$

Notice that the total power sums to zero. The same results can be expressed more literally as:

- *A* supplies 60 W
- *B* supplies 15 W
- *C* dissipates (consumes) 25 W
- *D* dissipates (consumes) 30 W

- *E* dissipates (consumes) 20 W
- Total power supplied equals 75 W
- Total power dissipated (consumed) equals 75 W
- Total power supplied = total power dissipated

Comments: Notice that whether power is calculated using $P = vi$ or $P = -vi$ depends entirely upon whether the passive sign convention is observed for any particular element.

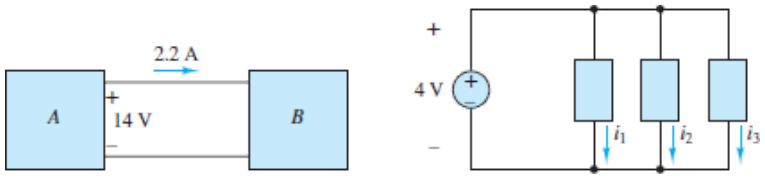
CHECK YOUR UNDERSTANDING

Compute the current through each of the headlamps shown in Figure 1.12 assuming each headlamp consumes 50 W. How much power is the battery providing?

Answer: $I_1 = I_2 = 4.17$ A; 100 W.

CHECK YOUR UNDERSTANDING

Determine which circuit element, *A* or *B*, in the figure on the left is supplying power and which is dissipating power. Also determine how much power is dissipated and supplied.



If the voltage source in the figure on the right supplies a total of 10 mW and $i_1 = 2$ mA and $i_2 = 1.5$ mA, what is the current i_3? If $i_1 = 1$ mA and $i_3 = 1.5$ mA, what is i_2?

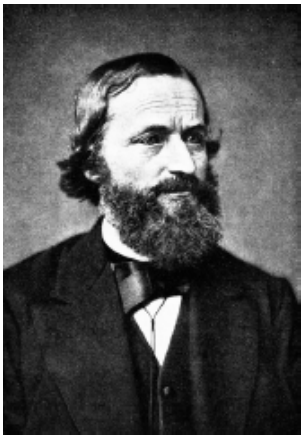
Answer: *A* supplies 30.8 W; *B* dissipates 30.8 W. $i_3 = -1$ mA; $i_2 = 0$ mA.

1.5 KIRCHHOFF'S LAWS

Earlier in this chapter, a circuit was defined as an electric network within which at least one closed path exists and around which electric charge may flow. In fact, conservation of electric charge requires a closed path for any non-zero current.

LO

To have a non-zero current, there must be a closed electrical path (i.e., a circuit).



Gustav Robert Kirchhoff (1824–1887) (*bilwissedition Ltd. & Co. KG/Alamy Stock Photo*)

For example, Figure 1.27 depicts a simple circuit, composed of a battery (e.g., a 1.5-V lithium battery) and a lightbulb. Conservation of charge requires that the current i from the battery to the lightbulb is equal to the current from the lightbulb to the battery. No current (nor charge) is "lost" around the closed circuit. This principle was observed by the German scientist G. R. Kirchhoff[1] and is known as **Kirchhoff's current law (KCL)**. This law states that *the net sum of the currents crossing any closed boundary must equal zero*.

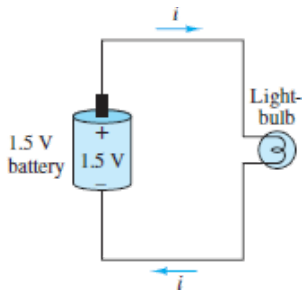


Figure 1.27 A simple electric circuit composed of a battery, a lightbulb, and two nodes

LO

$$\sum_{n=1}^{N} i_n = 0 \qquad \text{Kirchhoff's current law (KCL)} \tag{1.6}$$

where the sign of currents entering the region surrounded by the closed boundary must be opposite to the sign of currents exiting the same region. In other words, the sum of currents "in" must equal the sum of currents "out."

LO

$$\sum_{\text{in}} i = \sum_{\text{out}} i \qquad \text{Alternate KCL} \tag{1.7}$$

An application of Kirchhoff's current law is illustrated in Figure 1.28, where the simple circuit of Figure 1.27 has been augmented by the addition of two lightbulbs. The relationship between the currents is found by applying either version of KCL. To express the net sum of currents it is necessary to select a sign convention for currents entering and exiting a node. One possibility is to consider all currents entering a node as positive and all currents exiting a node as negative. (This particular sign convention is completely arbitrary.) The result of using this sign convention and applying the first version of KCL to node 1 is

$$i - i_1 - i_2 - i_3 = 0 \qquad \text{which is equivalent to} \qquad i = i_1 + i_2 + i_3$$

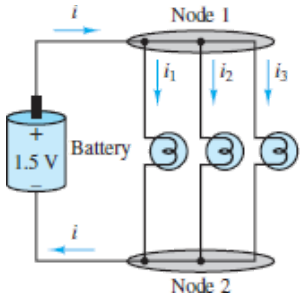


Figure 1.28 KCL applied at node 1 results in $i - i_1 - i_2 - i_3 = 0$, or equivalently $i = i_1 + i_2 + i_3$.

Note that the latter expression is exactly what would have been found if the alternate version of KCL had been applied. Also note that the result is the same if the opposite sign convention (i.e., currents entering and exiting the node are negative and positive, respectively) is used.

Page 23

Consider again the simple circuit of a battery and a lightbulb shown in Figure 1.29. **Kirchhoff's voltage law** (**KVL**) states that *the net change in electric potential around a closed loop is zero.* In mathematical terms:

$$\sum_{n=1}^{N} v_n = 0 \qquad \text{Kirchhoff's voltage law} \tag{1.8}$$

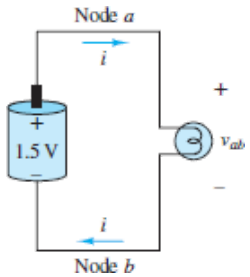


Figure 1.29 KVL applied clockwise from node *b* around the single loop circuit results in $1.5V - v_{ab} = 0$, or equivalently $v_{ab} = 1.5V$.

Here, v_n are the *changes* in voltage from *one node to another* around a closed loop.

When summing these changes in voltage, it is necessary to account for the polarity of the change. Changes in voltage from the minus sign − to the plus sign +

are considered positive (i.e., a rise in voltage), while those from plus to minus are considered negative (i.e., a drop in voltage). These two symbols act together to indicate the *assumed* direction of the change in voltage from one node to another.

An alternate but equivalent expression for KVL is that the sum of all voltage rises around a loop must equal the sum of all voltage drops around the same loop.

$$\sum_{\text{rises}} v = \sum_{\text{drops}} v \qquad \text{Alternate KVL} \tag{1.9}$$

In Figure 1.29, the *voltage across the lightbulb* is the change in electric potential from node *a* to node *b*. This change can also be expressed as the difference between two node voltages, v_a and v_b. The values of node voltages are relative to a reference node. Any single node may be chosen as the reference with its value set to zero, for simplicity. For the circuit in Figure 1.29 select node *b* as the reference and set its value as $v_b = 0$. Observe that the battery's positive terminal is 1.5 V *above the reference*, so that $v_a = 1.5$ V. In general, the battery guarantees that node *a* will always be 1.5 V above node *b*.

$$v_a = v_b + 1.5 \text{ V}$$
$$v_a = 1.5 \text{ V} \qquad \text{when} \qquad v_b = 0 \text{ acts as the reference.}$$

The notation used to express the *change* in voltage across the lightbulb, *from* node *b* *to* node *a*, is v_{ab}, where

$$v_{ab} \equiv v_a - v_b = 1.5 \text{ V}$$

EXAMPLE 1.6 Kirchhoff's Current Law Applied to an Automotive Electrical Harness

Problem

Figure 1.30 shows an automotive battery connected to a variety of elements in an automobile. The elements include headlights, taillights, starter motor, fan, power locks, and dashboard panel. The battery must supply enough current to satisfy each of the elements. Apply KCL to a model of the electrical system to find a relationship between the currents in the circuit.

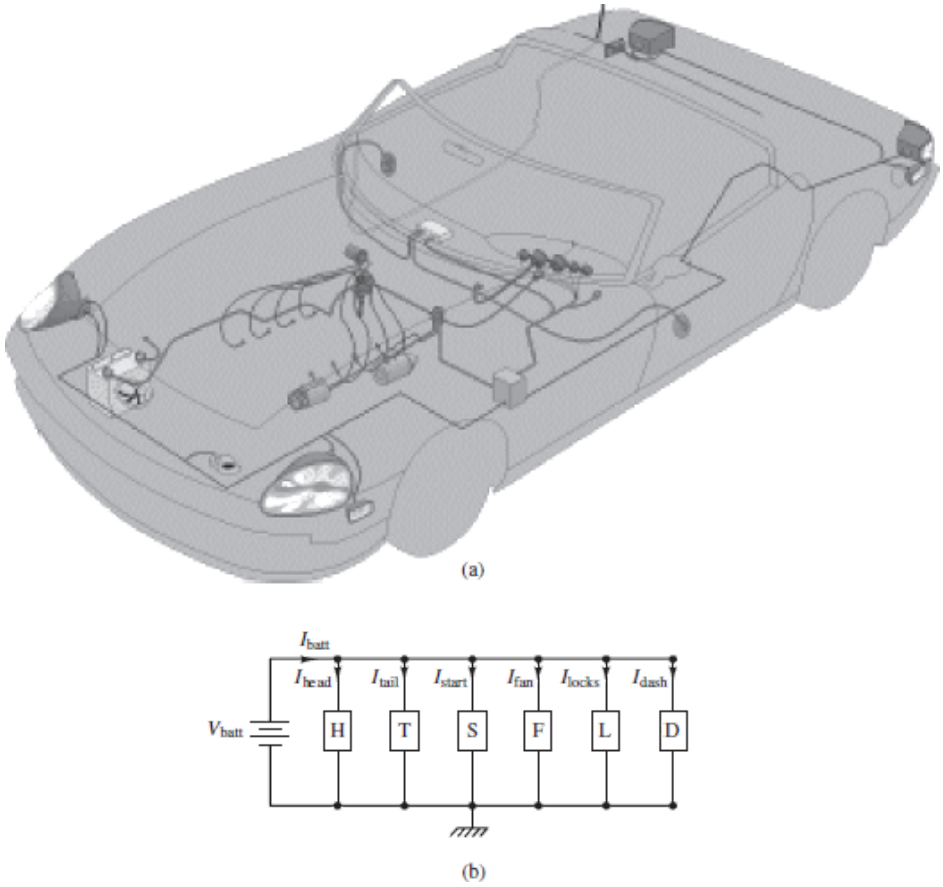


Figure 1.30 (a) Automotive electrical harness; (b) model electric circuit diagram

Solution

Known Quantities: Components of electrical harness: headlights, taillights, starter motor, fan, power locks, and dashboard panel.

Find: Expression relating battery current to harness currents.

Schematics, Diagrams, Circuits, and Given Data: Figure 1.30.

Assumptions: None.

Analysis: Figure 1.30(b) depicts the model electric circuit, illustrating that the current supplied by the battery is divided among the various elements. The application of KCL to the upper node yields

$$I_{batt} - I_{head} - I_{tail} - I_{start} - I_{fan} - I_{locks} - I_{dash} = 0$$

or

$$I_{batt} = I_{head} + I_{tail} + I_{start} + I_{fan} + I_{locks} + I_{dash}$$

Page 25

EXAMPLE 1.7 Application of KCL

Problem

Determine the unknown currents in the circuit of Figure 1.31.

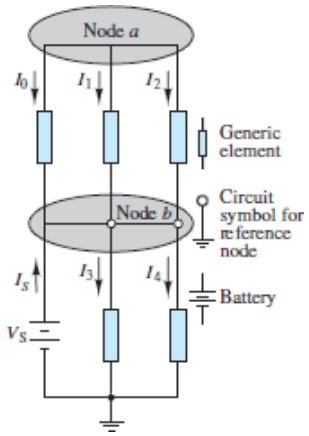


Figure 1.31 KCL yields $I_0 + I_1 + I_2 = 0$ at node *a* and $I_0 + I_1 + I_2 + I_S = I_3 + I_4$ at node *b*.

Solution

Known Quantities:

$$I_S = 5\text{ A} \qquad I_1 = 2\text{ A} \qquad I_2 = -3\text{ A} \qquad I_3 = 1.5\text{ A}$$

Find: I_0 and I_{24}.

Analysis: Two nodes are clearly shown in Figure 1.31 as node *a* and node *b*; the third node in the circuit is the reference node. Apply KCL at each of the three nodes.

At node *a*:

$$I_0 + I_1 + I_2 = 0 \qquad \text{from} \qquad \sum i_{out} = \sum i_{in}$$
$$I_0 + 2 - 3 = 0$$
$$\therefore \qquad I_0 = 1\text{ A}$$

Note that the assumed direction of all three currents is away from the node. However, I_2 has a negative value, which means that its actual direction is toward the node. The magnitude of I_2 is 3A.

At node *b*:

$$I_0 + I_1 + I_2 + I_S = I_3 + I_4 \qquad \text{from} \qquad \sum i_{in} = \sum i_{out}$$
$$1 + 2 - 3 + 5 = 1.5 + I_4$$
$$\therefore \quad I_4 = 3.5 \text{ A}$$

At the reference node: Assume that currents entering a node are positive and currents exiting a node are negative.

$$-I_S + I_3 + I_4 = 0$$
$$-5 + 1.5 + I_4 = 0$$
$$\therefore \quad I_4 = 3.5 \text{ A}$$

Comments: The result obtained at the reference node is exactly the same as that calculated at node *b*. Applying KCL to every node in a circuit will result in a redundant equation.

EXAMPLE 1.8 Application of KCL

Problem

Apply KCL to the circuit of Figure 1.32, using the concept of a supernode to determine the source current i_{s1}.

Page 26

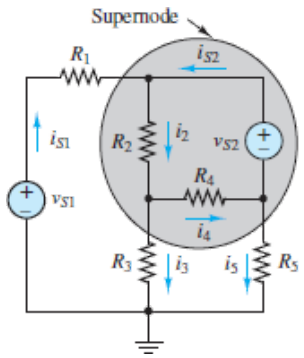


Figure 1.32 KCL applied at the boundary of the supernode yields $i_{S1} = i_3 + i_5$.

Solution

Known Quantities:

$$i_3 = 2\ \text{A} \qquad i_5 = 0\ \text{A}$$

Find: i_{S1}.

Analysis: Apply KCL at the closed boundary of the so-called supernode to obtain

$$i_{S1} = i_3 + i_5 \qquad \text{from} \qquad \sum i_{in} = \sum i_{out}$$
$$i_{S1} = 2 + 0 = 2\ \text{A}$$

Comments: Notice that the same result for i_{S1} is obtained by applying KCL at the bottom node. This result is another example of a redundant equation that is sometimes obtained by applying KCL at two different closed boundaries or nodes.

EXAMPLE 1.9 Kirchhoff's Voltage Law—Electric Vehicle Battery Pack

Problem

Figure 1.33(a) depicts the battery pack in the Smokin' Buckeye electric race car, which consists of thirty-one 12-V batteries in series.

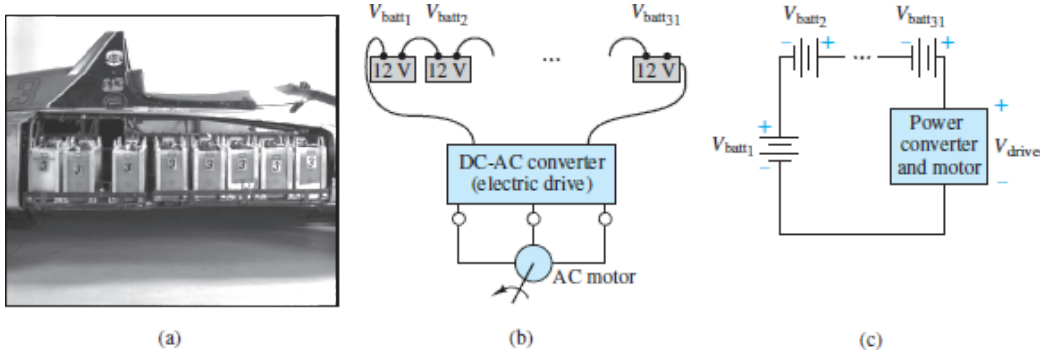


Figure 1.33 Electric vehicle battery pack illustrates KVL. (*Courtesy: David H. Koether Photography)*

Solution

Known Quantities: Nominal characteristics of **Optima™ lead-acid batteries.**

Find: Expression relating battery and electric motor drive voltages.

Schematics, Diagrams, Circuits, and Given Data: $V_{batt} = 12$ V; Figure 1.33(a), (b), and (c).

Assumptions: None.

Analysis: Figure 1.33(b) models the electric circuit, illustrating the batteries in series with the electric drive that powers the vehicle's 150-kW three-phase induction motor. Apply KVL around the circuit of Figure 1.33(c):

$$\sum_{n=1}^{31} V_{batt_n} - V_{drive} = 0$$

Thus, the electric drive is nominally supplied by a $31 \times 12 = 372$-V battery pack. In practice, the voltage across a lead-acid battery depends upon the state of charge of the battery. When fully charged, the battery pack of Figure 1.33(a) supplies closer to 400 V (i.e., roughly 13 V per battery).

EXAMPLE 1.10 Application of KVL

Problem

Determine the unknown voltage v_2 by applying KVL to the circuit of Figure 1.34.

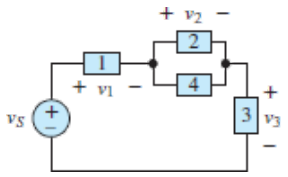


Figure 1.34 A circuit with four generic elements and one ideal voltage source

Solution

Known Quantities:

$$v_S = 12 \text{ V} \qquad v_1 = 6 \text{ V} \qquad v_3 = 1 \text{ V}$$

Find: v_2.

Analysis: Apply KVL starting at the reference node and proceeding clockwise around the large outer loop (the outer perimeter) of the circuit:

$$\begin{aligned} v_S - v_1 - v_2 - v_3 &= 0 \\ v_S - v_1 - v_3 &= v_2 \\ 12 - 6 - 1 &= v_2 = 5 \text{ V} \end{aligned}$$

Comments: Note that v_2 is the voltage across elements 2 and 4. These two elements are in *parallel* because they are located between the same two nodes. One can also say that the two branches that contain these elements are in parallel.

EXAMPLE 1.11 Application of KVL

Problem

Use KVL to determine the unknown voltages v_1 and v_4 in the circuit of Figure 1.35.

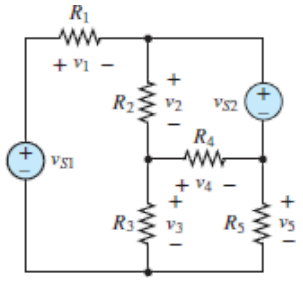


Figure 1.35 Circuit for Example 1.11

Solution

Known Quantities:

$$v_{S1} = 12 \text{ V} \qquad v_{S2} = -4 \text{ V} \qquad v_2 = 2 \text{ V} \qquad v_3 = 6 \text{ V} \qquad v_5 = 12 \text{ V}$$

Find: v_1, v_4.

Analysis: To determine the unknown voltages, apply KVL clockwise around the left and upper-right meshes:

$$\begin{aligned} v_{S1} - v_1 - v_2 - v_3 &= 0 \\ v_2 - v_{S2} + v_4 &= 0 \end{aligned}$$

After substituting numerical values, the equations become:

$$\begin{aligned} 12 - v_1 - 2 - 6 &= 0 \\ v_1 &= 4 \text{ V} \\ 2 - (-4) + v_4 &= 0 \\ v_4 &= -6 \text{ V} \end{aligned}$$

It is possible to solve for v_1 and v_4 using other loops in the circuit. For instance, apply KVL clockwise around the lower-right mesh to find v_4:

$$\begin{aligned} v_3 - v_4 - v_5 &= 0 \\ 6 - v_4 - 12 &= 0 \\ v_4 &= -6 \text{ V} \end{aligned}$$

Or apply KVL clockwise around the outer most loop to find v_1:

$$\begin{aligned} v_{S1} - v_1 - v_{S2} - v_5 &= 0 \\ 12 - v_1 - (-4) - 12 &= 0 \\ v_1 &= 4 \text{ V} \end{aligned}$$

Comments: Notice that there are seven closed wire loops in the circuit. KVL could be applied around any of these loops to find an equation. The key is to find two linearly independent equations that involve the two unknowns.

CHECK YOUR UNDERSTANDING

Apply KVL to each of the other three closed wire loops in Figure 1.35 that were not explored in Example 1.11. Compare the results to those found in the example. Are the results consistent?

CHECK YOUR UNDERSTANDING

Repeat the exercise of Example 1.7 when *I* 0 = 0.5 A, *I* 2 = 2 A, *I* 3 = 7 A, and *I* 4 = − 1 A. Find *I* 1 and *I S*.

Answer: $I_1 = -2.5$ A and $I_S = 6$ A

CHECK YOUR UNDERSTANDING

Use the result of Example 1.8 and the following data to compute the current *iS*2 in the circuit of Figure 1.32.

$$i_2 = 3\,\text{A} \qquad i_4 = 1\,\text{A}$$

Answer: $i_{\omega} = 1$ A

Page 29

1.6 RESISTANCE AND OHM'S LAW

When charge flows through a wire or circuit element, it encounters **resistance**, the magnitude of which depends on the *resistivity* of the material and the geometry of the wire or element. In practice, all circuit elements exhibit some resistance, which leads to energy dissipation in the form of heat. Whether this loss of electrical energy as heat is detrimental depends upon the purpose of the circuit element. For example, a typical electric toaster relies on the conversion of electrical energy to heat within its resistive coils to accomplish its purpose, the making of toast. All electric heaters rely upon this process, in one form or another. On the other hand, heat loss due to resistance in residential wiring is costly, and potentially dangerous. Resistance in microcircuitry generates heat that effectively limits the speed of microprocessors and the number and scale of transistors that can be packed into a given volume.

The resistance of a cylindrical wire segment, as shown in Figure 1.36(a), is given by

$$R = \rho\frac{l}{A} = \frac{l}{\sigma A} \tag{1.10}$$

where ρ and σ are the material properties *resistivity* and *conductivity*, respectively, and l and A are the segment length and cross-sectional area, respectively. As evident in the above equation, conductivity is simply the inverse of resistivity. The unit of resistance R is **ohms (Ω)**, where

$$1\ \Omega = 1\ \text{V/A} \tag{1.11}$$

The resistance of an actual wire or circuit element is usually accounted for in a circuit diagram by an **ideal resistor**, which lumps the entire distributed resistance R of the wire or element into one single element. Ideal resistors exhibit a linear i-v relationship known as **Ohm's law**, which is

$$v = iR \qquad \text{Ohm's law} \tag{1.12}$$

In other words, the voltage across an ideal resistor is directly proportional to the current through it. The constant of proportionality is the resistance R. The circuit symbol and i-v characteristic for an ideal resistor are shown in Figure 1.36(b) and (c), respectively. Notice the passive sign convention used in the circuit symbol diagram, as appropriate, since a resistor is a passive element.

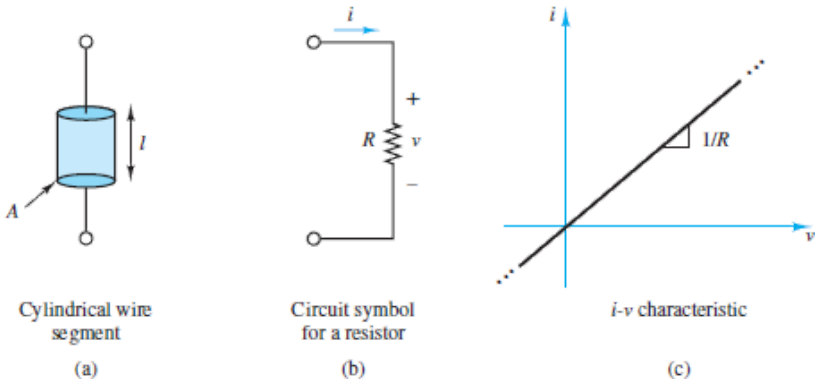


Figure 1.36 (a) Resistive wire segment; (b) ideal resistor circuit symbol; (c) the i-v relationship (Ohm's law) for an ideal resistor

Page 30

It is often convenient to define the *conductance*, G (unit is siemens, S), of a circuit element as the inverse of its resistance.

$$G = \frac{1}{R} \qquad \text{siemens (S)} \qquad \text{where} \qquad 1\ \text{S} = \frac{1\text{A}}{\text{V}} \tag{1.13}$$

In terms of conductance, Ohm's law is

$$i = Gv \tag{1.14}$$

Ohm's law is an *empirical* relationship that finds widespread application in electrical engineering. It is a simple yet powerful approximation of the physics of

electrical conductors. However, the linear *i*-*v* relationship usually does not apply over very large ranges of voltage or current. For some conductors, Ohm's law does not approximate the *i*-*v* relationship even over modest ranges of voltage or current. Nonetheless, most conductors exhibit piecewise linear *i*-*v* characteristics for one or more ranges of voltage and current, as shown in Figure 1.37 for an incandescent lightbulb and a semiconductor diode.

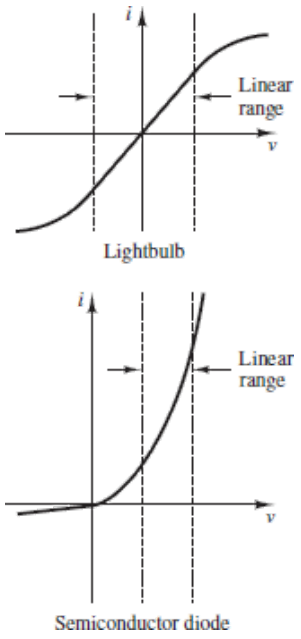


Figure 1.37 Piecewise linear segments within non-linear *i*-*v* characteristics

Short- and Open-Circuits

Two convenient idealizations, the **short-circuit** and the **open-circuit**, are limiting cases of Ohm's law as the resistance approaches zero or infinity, respectively. Formally, a short-circuit is an element *across* which the voltage is zero, regardless of the current *through* it. Figure 1.38 depicts the circuit symbol for an ideal short-circuit.

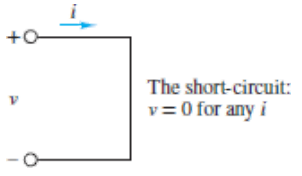


Figure 1.38 The short-circuit

In practice, any conductor will exhibit some resistance. For practical purposes, however, many elements approximate a short-circuit under certain conditions. For

example, a large-diameter copper pipe is effectively a short-circuit in the context of a residential electric power supply, while in a low-power microelectronic circuit (e.g., an iPhone®) a typical ground plane is 35×10^{-6} m thick, which is adequate for a short-circuit in that context. A typical solderless breadboard is designed to accept 22-gauge solid jumper wires, which act effectively as short-circuits between elements on the breadboard. Table 1.1 lists the resistance per 1,000 ft of some commonly used wire, as specified by the *American Wire Gauge Standards*.

Table 1.1 Resistance of copper wire

AWG size	Number of strands	Diameter per strand (in)	Resistance per 1,000 ft (Ω)
24	Solid	0.0201	28.4
24	7	0.0080	28.4
22	Solid	0.0254	18.0
22	7	0.0100	19.0
20	Solid	0.0320	11.3
20	7	0.0126	11.9
18	Solid	0.0403	7.2
18	7	0.0159	7.5
16	Solid	0.0508	4.5
16	19	0.0113	4.7
14	Solid	0.0641	2.52
12	Solid	0.0808	1.62
10	Solid	0.1019	1.02
8	Solid	0.1285	0.64
6	Solid	0.1620	0.4
4	Solid	0.2043	0.25
2	Solid	0.2576	0.16

Page 31

The limiting case for Ohm's law when $R \rightarrow \infty$ is called an **open-circuit**. Formally, an open-circuit is an element *through* which the current is zero, regardless of the voltage *across* it. Figure 1.39 depicts the circuit symbol for an ideal open-circuit.

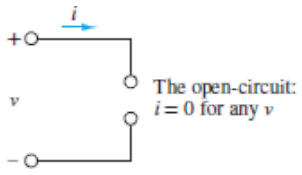


Figure 1.39 The open-circuit

In practice, it is easy to approximate an open-circuit. For moderate voltage levels, any gap or break in a conducting path amounts to an open-circuit. However, at sufficiently high voltages such a gap will become ionized and its resistance will

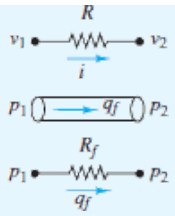
decrease suddenly and dramatically, effectively producing a short-circuit across the gap. If sufficient charge is available, the result will be arcing in which a pulse of charge jumps the gap. Subsequently, the pulse discharge results in a decrease in the voltage *across* the gap and the ionized path collapses. The result is that the gap has returned to its open-circuit approximation. This phenomenon is employed in spark plugs to ignite the air-fuel mixture in a spark-ignition internal combustion engine. Any insulating material will break down when a sufficiently high voltage is applied across it.

The *dielectric strength* is a measure of the maximum electric field (voltage per unit distance) that an insulating material can sustain without breaking down and allowing charge to flow. This measure is somewhat dependent upon temperature, pressure, and the material thickness; however, typical values are 3 kV/mm for air at sea level and room temperature, 10 kV/mm for window glass, 20 kV/mm for neoprene rubber, 30 kV/mm for pure water, and 60 kV/mm for PTFE, commonly known as Teflon.

MAKE THE CONNECTION

Hydraulic Analog of Electrical Resistance

A useful analogy can be made between the electric current through electric components and the flow of incompressible fluids (e.g., water, oil) through hydraulic components. The fluid flow rate *through* a pipe is analogous to current *through* a conductor. Similarly, pressure drop *across* a pipe is analogous to voltage *across* a resistor. The resistance of the pipe to fluid flow is analogous to electrical resistance: The pressure difference across the pipe causes fluid flow, much as a potential difference across a resistor causes charge to flow. The figure below depicts how pipe flow is often modeled as current through a resistance.



Analogy between electrical and fluid resistance

Discrete Resistors

Various types of *discrete resistors* are used in laboratory experiments, tinkering projects, and commercial hardware, and are available in a wide range of nominal values, tolerances, and power ratings. Each type has a particular temperature range within which it is designed to operate. In fact, some discrete resistors (known as thermistors) are designed to be highly sensitive to temperature and to be used as temperature transducers.

The majority of discrete resistors have a cylindrical shape and are color coded for their nominal value and tolerance. Several common types of resistors are: *carbon composites*, in which the resistance is set by a mixture of carbon and ceramic powder (Figure 1.40); *carbon film*, in which the resistance is set by the length and width of a thin strip of carbon wrapped around an insulating core; and thin metal film, in which the resistance is set by the characteristics of a thin metal film also wrapped around an insulating core (Figure 1.41).

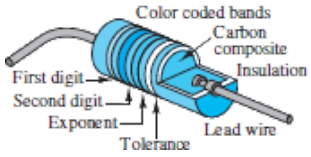


Figure 1.40 Carbon composite resistor

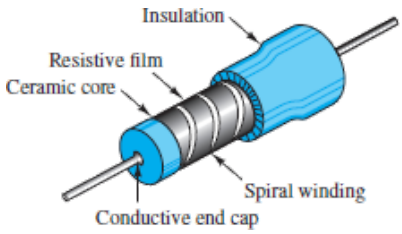


Figure 1.41 Thin-film resistor

Discrete resistors are available with various power ratings, where the power rating scales with the size of the resistor itself. Figures 1.42 and 1.43 show (to scale) typical $\frac{1}{4}$-W and $\frac{1}{2}$-W resistors, respectively. Notice the bands along the length of each resistor. Discrete resistors are also available with typical power ratings of 1, 2, 5, 10 W, and larger. Many industrial power resistors are manufactured by winding wire, such as Nichrome, around a non-conducting core, such as ceramic, plastic, or fiberglass. Others are made of cylindrical sections of carbon. Power resistors are available in a variety of packages, such as cement or molded plastic, aluminum encasements with fins for wicking away heat, and enamel coatings. Typical power resistors are shown in Figure 1.44.

Figure 1.42 Typical $\frac{1}{4}$-W resistors (*Jim Kearns*)

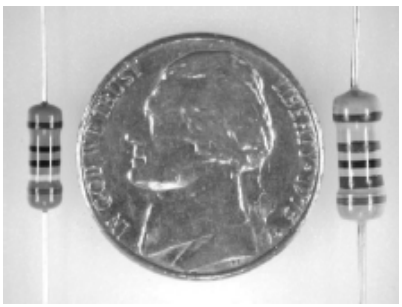


Figure 1.43 Typical $\frac{1}{2}$-W resistors (*Jim Kearns*)

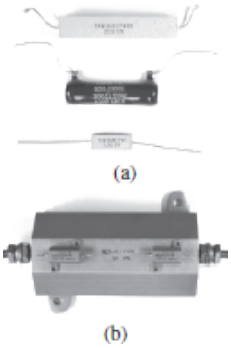


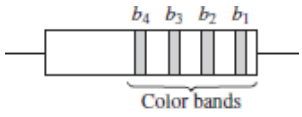
Figure 1.44 (a) 25-W, 20-W, and 5-W, and (b) two 5-W resistors sitting atop one 100-W resistor (*Jim Kearns*)

The value of a discrete resistor is determined by the resistivity, shape, and size of the conducting element. Table 1.2 lists the resistivity of many common materials.

Table 1.2 Resistivity of common materials at room temperature

Material	Resistivity (Ω-m)
Aluminum	2.733×10^{-8}
Copper	1.725×10^{-8}
Gold	2.271×10^{-8}
Iron	9.98×10^{-8}
Nickel	7.20×10^{-8}
Platinum	10.8×10^{-8}
Silver	1.629×10^{-8}
Carbon	3.5×10^{-5}

The nominal value and tolerance are often color-coded on a discrete resistor. Typically, discrete resistors have four color bands, where the first two designate a two-digit integer, the third designates a multiplier of 10, and the fourth designates the tolerance. Occasionally, discrete resistors have five bands, where the first three designate a three-digit integer, and the remaining two designate the multiplier and the tolerance. The value of each color band is decoded using the system displayed in Figure 1.45 and Table 1.3.



black	0	blue	6
brown	1	violet	7
red	2	gray	8
orange	3	white	9
yellow	4	silver	10%
green	5	gold	5%

Resistor value = $(b_1\, b_2) \times 10^{b_3}$;
b_4 = % tolerance in actual value

Figure 1.45 Resistor color code

$$(\text{Two- or three-digit integer}) \times 10^{\text{multiplier}}, \quad \text{in ohms } (\Omega)$$

Table 1.13 $b_1 b_2$ indicates the two-digit significand; b_3 indicates the multiplier

$b_1 b_2$	Code	b_3	Code	Ω	b_3	Code	kΩ	b_3	Code	kΩ	b_3	Code	kΩ
10	Brn-blk	1	Brown	100	2	Red	1.0	3	Orange	10	4	Yellow	100
12	Brn-red	1	Brown	120	2	Red	1.2	3	Orange	12	4	Yellow	120
15	Brn-grn	1	Brown	150	2	Red	1.5	3	Orange	15	4	Yellow	150
18	Brn-gry	1	Brown	180	2	Red	1.8	3	Orange	18	4	Yellow	180
22	Red-red	1	Brown	220	2	Red	2.2	3	Orange	22	4	Yellow	220
27	Red-vlt	1	Brown	270	2	Red	2.7	3	Orange	27	4	Yellow	270
33	Org-org	1	Brown	330	2	Red	3.3	3	Orange	33	4	Yellow	330
39	Org-wht	1	Brown	390	2	Red	3.9	3	Orange	39	4	Yellow	390
47	Ylw-vlt	1	Brown	470	2	Red	4.7	3	Orange	47	4	Yellow	470
56	Grn-blu	1	Brown	560	2	Red	5.6	3	Orange	56	4	Yellow	560
68	Blu-gry	1	Brown	680	2	Red	6.8	3	Orange	68	4	Yellow	680
82	Gry-red	1	Brown	820	2	Red	8.2	3	Orange	82	4	Yellow	820

For example, a resistor with four bands (yellow, violet, red, gold) has a nominal value of:

$$(\text{yellow})(\text{violet}) \times 10^{\text{red}} = 47 \times 10^2 = 4700\ \Omega = 4.7\ \text{k}\Omega$$

and a "gold" tolerance of ±5 percent 4.7 kΩ is often shortened in practice to 4K7, where the letter K indicates the placement of the decimal point as well as the unit of kΩ. Likewise, 3.3MΩ is often shortened to 3M3. Table 1.3 lists the standard nominal values established by the Electronic Industries Association (EIA) for a tolerance of 10 percent, commonly referred to as the E12 series. The number 12 indicates the number of logarithmic steps per decade of resistor values. Notice that the values in adjacent decades (columns) are different by a factor of 10.

Due to imperfect manufacturing the actual value of a discrete resistor is only approximately equal to its nominal value. The tolerance is a measure of the likely variation between the actual value and the nominal value. Other EIA series are E6, E24, E48, E96, and E192 for tolerances of 20%, 5%, 2%, 1%, and even finer tolerances, respectively.

Variable Resistors

The resistance of a variable resistor is not fixed but can vary with some other quantity. Examples of variable resistors are a photoresistor and a thermistor, in which the resistance varies with light intensity and temperature, respectively. Many useful sensors are based upon variable resistors.

Figure 1.46 shows a simple loop with a voltage source, a variable resistor R, and a fixed resistor $R0$. Apply KVL around the loop:

$$\begin{aligned} v_S &= iR + iR_0 = i(R + R_0) \\ &= iR + v_0 \end{aligned}$$

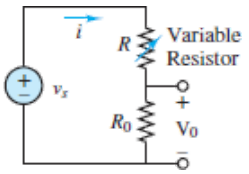


Figure 1.46 A variable resistor R in a series loop

Solve for i and substitute for it in the above equation:

$$i = \frac{v_S}{R + R_0} \qquad \text{and} \qquad v_0 = iR_0 = v_S \frac{R_0}{R + R_0}$$

Page 34

Now assume that the variable resistor has a range from 0Ω to some value R_{max} that is much larger than R_0. When $R = 0$:

$$v_0 = v_S \frac{R_0}{R + R_0} = v_S \frac{R_0}{R_0} = v_S \qquad (R = 0)$$

When $R = R_{\text{max}}$:

$$v_0 = v_S \frac{R_0}{R + R_0} = v_S \frac{R_0}{R_{\text{max}} + R_0} \approx v_S \frac{R_0}{R_{\text{max}}} \approx 0 \qquad (R = R_{\text{max}})$$

Thus, as R varies from 0 to R_{max}, $v0$ varies from v_S to 0. The changes in R can be observed as changes in $v0$. Imagine that the variable resistor in Figure 1.46 is a photoresistor, such as a cadmium sulfide (CdS) cell shown in Figure 1.47(a), that has a very small resistance when the incident light intensity is bright and has a very large resistance when the incident light intensity is dim or dark. The result is that under bright conditions, $v_0 \approx v_S$, while under dark conditions, $v_0 \approx 0$. A nightlight, such as that shown in Figure 1.47(b), is a device that turns on when $v_0 \ll v_{\text{ref}}$ and turns off when $v_0 \gg v_{\text{ref}}$, where v_{ref} is some appropriate reference voltage, such as $vS/2$.

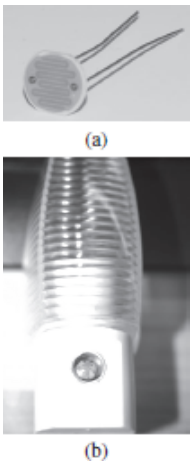


Figure 1.47 (a) A typical cadmium sulfide (CdS) cell. (b) A nightlight relies on a CdS cell to detect dark conditions. (*Jim Kearns*)

Figure 1.48 shows a typical thermistor, which can be used in exactly the same manner as a CdS cell but which responds to changes in temperature.

Figure 1.48 A typical negative temperature coefficient (NTC) thermistor (*Jim Kearns*)

Potentiometers

A potentiometer is a three-terminal device. Figure 1.49 depicts a potentiometer and its circuit symbol. A potentiometer has a fixed resistance R_0, typically formed by a tightly wound coil of wire, between terminals *A* and *C*. Terminal *B* is connected to a *wiper* that slides along the coil as the knob is turned. The arrow in the circuit symbol represents the position of the slider along the length of the coil R_0. The resistance from terminal B to the other two terminals is determined by the wiper position. As

R_{BA} increases, R_{BC} decreases, and vice versa, such that the sum $R_{BA} + R_{BC}$ always equals R_0.

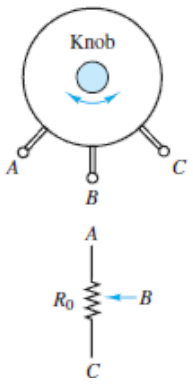


Figure 1.49 A potentiometer is a three-terminal resistive device with a fixed resistance *R*0 between terminals *A* and *C*. The resistances between terminal *B* (the "wiper") and the other two terminals is set by the knob.

Figure 1.50(a) illustrates the use of a potentiometer symbol in a simple circuit. Figure 1.50(b) is an equivalent representation of the circuit, where the resistance between terminals *A* and *B* and that between terminals *B* and *C* are depicted as discrete resistors.

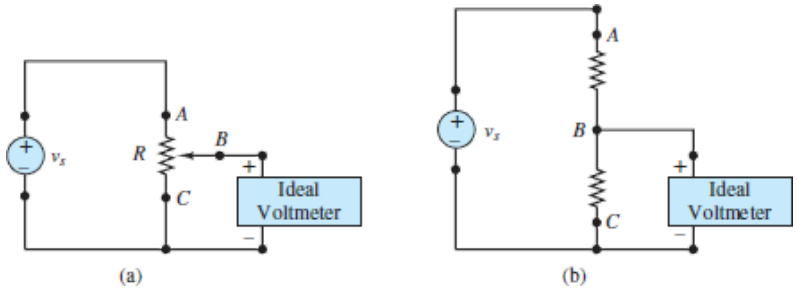


Figure 1.50 (a) A potentiometer in a simple circuit; (b) an equivalent circuit of (a), where $R = R_{AB} + R_{BC} = R_{AC}$

The ideal voltmeter reading v_bc can be calculated in a manner similar to that used in the preceding section on variable resistors. Apply KVL around the loop Page 35containing the voltage source and the two discrete resistors, using Ohm's law to express the change in voltage across each resistor. The result is

$$v_{BC} = v_S \frac{R_{BC}}{R_{BC} + R_{AB}}$$

This important result for two resistors in series is an example of *voltage division*. When the wiper is turned all the way to terminal C, $R_{BC} = 0$ and so $v_{BC} = 0$. When the wiper is turned all the way to terminal A, $R_{AB} = 0$ and so $v_{BC} = v_S$. In general, as the wiper is turned from terminal A to terminal C, the voltage across terminals B and C falls continuously from vS to 0.

Power Dissipation in Resistors

All discrete resistors have a power rating, which is not designated by a color band, but which tends to scale with the size of the resistor itself. Larger resistors typically have a larger power rating. The power consumed or dissipated by a resistor R is

$$P = vi = (iR)i = i^2R > 0$$
$$= v\left(\frac{v}{R}\right) = \frac{v^2}{R} > 0 \tag{1.15}$$

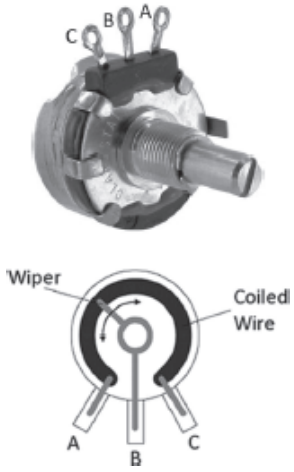


Figure 1.51 A typical $\frac{1}{2}$- potentiometer and its internal construction (*Jim Kearns*)

Remember that the voltage v and the current i are defined and linked by the *passive sign convention* and that power consumed by an element is positive. In the case of resistors, power is always positive and energy is dissipated as heat. The implication is that if the current through (or the voltage across) a resistor is too large, the power will exceed the resistor's rating and result in a smoking and/or burning resistor! The smell of an overheating resistor is well known to technicians and hobbyists alike.

Positive power is power dissipated (i.e., consumed) by an element.

EXAMPLE 1.12 Using Resistor Power Ratings

Problem

For a given voltage across a resistor, determine the minimum allowed resistance for a $\frac{1}{4}$W power rating.

Solution

Known Quantities: Resistor power rating 0.25 W. Voltages due to a battery across the resistor: 1.5 V and 3 V.

Find: The minimum allowed resistance for a $\frac{1}{4}$W resistor.

Schematics, Diagrams, Circuits, and Given Data: Figures 1.52 and 1.53.

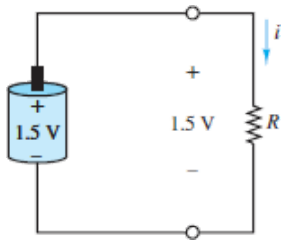


Figure 1.52

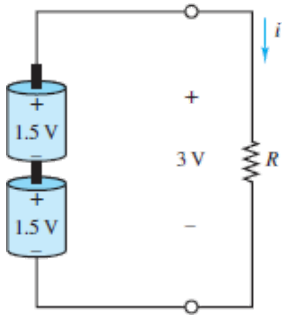


Figure 1.53

Analysis: The power dissipated by a resistor is

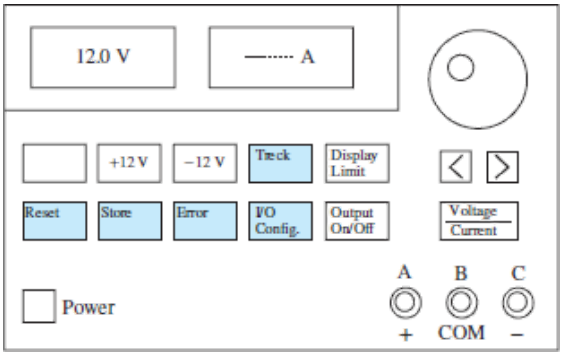
$$P_R = vi = v \cdot \frac{v}{R} = \frac{v^2}{R}$$

Setting P_R equal to the resistor power rating yields $v^2/R \leq 0.25$, or $R \geq v^2/0.25$. For a 1.5-V battery, the minimum size resistor will be $R = 1.5^2/0.25 = 9\Omega$. For a 3-V battery, the minimum size resistor will be $R = 3^2/0.25 = 36\ \Omega$.

Comments: Sizing resistors on the basis of power rating is very important since, in practice, resistors eventually fail when the power rating is exceeded. Notice that the minimum resistor size was quadrupled when the voltage was doubled. This result reflects the fact that power increases with the square of the voltage. Also notice that the power dissipated for the 3 -V battery is four times the power dissipated for the 1.5 -V battery. In other words, the power dissipated by R in Figure 1.53 cannot be computed by assuming that each of the two 1.5 -V batteries supplies the same amount of power as the single 1.5 -V battery in Figure 1.52. In fact, each battery in Figure 1.53 supplies twice as much power as the single battery in Figure 1.52. In mathematical terms, power is not linear.

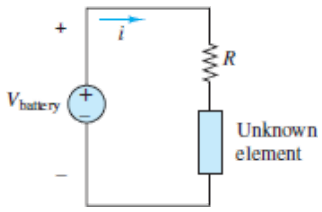
CHECK YOUR UNDERSTANDING

A typical three-terminal electronic power supply (see the first illustration) provides ±12 V, such that the change in voltage from terminal C to B is +12 V and that from terminals B to A is also +12 V. What is the minimum size (value) of a $\frac{1}{4}$-W resistor placed across terminals A and C? (*Hint:* The voltage from terminal C to A is +24 V.)

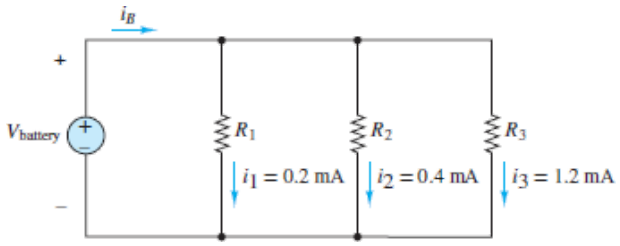


The single loop circuit in the illustration on the right contains a battery, a resistor, and an unknown circuit element.

1. If the voltage V_{batter} is 1.45 V and i = 5 mA, find the power supplied to or by the battery.
2. Repeat part 1 if i = −2 mA.



The battery in the triple mesh circuit shown below supplies power to resistors R_1, R_2, and R_3. Use KCL to determine the current i_B, and find the power supplied by the battery if $V_{battery} = 3$ V.



2,304 Ω; $P_1 = 7.25 \times 10^{-3}$ W (supplied by); $P_2 = 2.9 \times 10^{-3}$ W (supplied to); $i_B = 1.8$ mA; $P_B = 5.4$ mW

Page 37

MAKE THE CONNECTION

Thermal Systems

An analogy between electric circuits and thermal systems is often helpful for understanding both subjects. For example, the heat transfer that occurs when a heat-treated engine crankshaft is rapidly cooled ("quenched") in water has a useful "lumped parameter" model that is analogous to an electric circuit. This analogy includes the representation of two modes of heat transfer, conduction and convection, as electrical resistances, and the representation of the high-temperature crankshaft conducting heat to the surrounding water bath as a one-port network attached to a load. In this model, the temperature difference between the crankshaft and the

ambient water bath is analogous to a voltage difference (i.e., voltage source). Further, the heat transfer from crankshaft to water bath is analogous to an electric current from one terminal of a voltage source to the other. Finally, the lumped thermal conductivity of the crankshaft is analogous to the inverse of a Thevenin resistance while the convective heat transfer per degree of temperature difference at the surface of the crankshaft ($h\,A$) is analogous to an inverse load resistance.

The table below illustrates the analogy between various parameters and variables involved in this cooling process. As mentioned above, the difference in electrical potential across two nodes is analogous to the temperature difference between two bodies or surfaces. The second law of thermodynamics requires thermal energy to flow from high to low temperatures, unless external work is done to reverse the flow. Likewise, charge must flow from high to low electric potential, unless external work is done to reverse the flow. Heat is the rate at which thermal energy flows, just as current is the rate at which charge flows. Thus, heat transfer is analogous to current.

In general, heat transfer occurs in three different modes: conduction, convection, and radiation. For simplicity, only the first two are considered below.

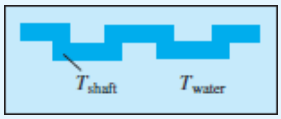
Thermal Resistance

Again, consider a heat-treated engine crankshaft that is to be quenched in a water bath at ambient temperature (see the figure below). Thermal energy at high temperatures flows from within the shaft to its surface, and then from the shaft surface to the water. This process continues until the shaft temperature approaches the water temperature.

The heat-transfer mechanism within the crankshaft itself is known as *conduction,* which occurs whenever there is a temperature gradient within a material body. Recall that temperature is the macroscopic expression of microscopic thermal vibrations. In conducting materials, such as metals, thermal energy is conveyed through the material by lattice vibrations and by *conduction band* electrons that are free to roam within the material matrix. These are the same electrons that account for the electric current in conductors. Their *mobility* is the principal contributor to the heat-transfer conduction coefficient k and the analogous electrical property of resistivity ρ.

Electrical variable	Thermal variable
Potential difference v, [V]	Temperature difference ΔT, [°C]
Current i, [A]	Heat flux q, [W]
Resistance R, [Ω]	Resistance R_t, [°C/W]
Resistivity ρ, [Ω/m]	Conduction heat-transfer coefficient $k, \left[\frac{\text{W}}{\text{m} - ^\circ\text{C}}\right]$
(No exact electrical analogy)	Convection heat-transfer coefficient, or film coefficient of heat transfer $h, \left[\frac{\text{W}}{\text{m}^2 - ^\circ\text{C}}\right]$

The heat-transfer mechanism at the surface of the crankshaft is known as *convection*, which occurs whenever there is fluid present to sweep away thermal energy at the surface of a body. Convection can be either forced (bulk fluid flow driven by some external means) or free (bulk fluid flow due to density gradients brought about by temperature gradients). In general, heat transfer to the fluid increases with fluid flow although the details of the fluid mechanics is often quite complicated. In a quenching process with large temperature gradients between a body and a fluid, free convection due to boiling can dominate the heat-transfer load seen by the body. The convective heat transfer between the shaft and the water bath is dependent on the product of the shared surface area between the shaft and the water bath A and the convective heat transfer coefficient h.



Thermal Circuit Model

The *conduction resistance* of the shaft is described by the following equation:

$$q = \frac{kA_t}{L}\Delta T$$

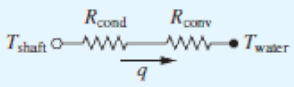
$$R_{cond} = \frac{\Delta T}{q} = \frac{L}{kA_t}$$

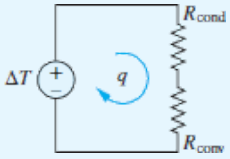
where A_1 is a cross sectional area and L is the distance from the inner core to the surface. The convection resistance is described by a similar equation, in which heat flow is described by the convective heat-transfer coefficient, h:

$$q = hA_2\Delta T$$
$$R_{conv} = \frac{\Delta T}{q} = \frac{1}{hA_2}$$

where A_2 is the surface area of the shaft in contact with the water. The equivalent thermal resistance and the overall circuit model of the crankshaft quenching process are shown in the figures below.

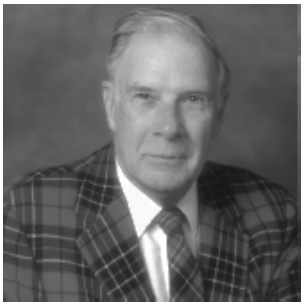
The following two figures represent the thermal resistance of the quenching process as the series summation of two analogous electrical resistances, R_{cond} and R_{conv}, and an analogous equivalent electric circuit representation of the quenching process itself.





1.7 THE NODE VOLTAGE METHOD

The **node voltage method** and the **mesh current method** are powerful computational tools for calculating voltages and currents in electrical circuits. Although relatively straightforward, these methods require practice to apply correctly. When applied to linear networks, both methods yield a system of *linearly independent* equations that can be solved easily by a computer. However, the methods themselves offer little insight to the fundamental nature and characteristics of an electrical network. Such insight, which is essential when attempting to modify or design a network, must be acquired through a careful and thoughtful examination of the data generated by these methods. Little learning of any importance will occur simply by applying these methods. The following quote by R.W. Hamming is well worth keeping in mind.



R.W. Hamming (*IEEE, Inc.*)

"The purpose of computation is insight, not numbers."

Page 39

The **node voltage method** (also known as nodal analysis) is based on defining the voltage at each node as an independent variable. One of the nodes is freely chosen as a **reference node**. Ohm's law is used to express resistor currents in terms of node voltages, such that *each branch current is expressed in terms of one or more node voltages.* Finally, KCL is applied to each non-reference node. Figures 1.54 and 1.55 illustrate how to apply Ohm's law and KCL in this method.

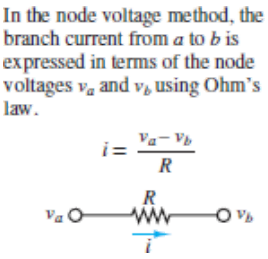


Figure 1.54 Branch current formulation in node analysis

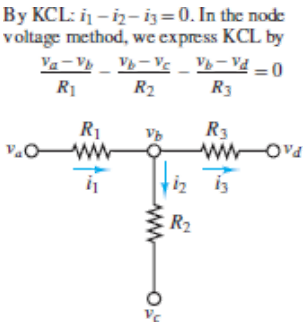


Figure 1.55 Use of KCL in node analysis

Once each branch current is defined in terms of the node voltages, Kirchhoff's current law is applied at each node:

$$\sum i = 0 \tag{1.16}$$

Consider the circuit shown in Figure 1.56. The directions of currents $i1$, $i2$, and $i3$may be selected arbitrarily; however, it is often helpful to select directions that conform with one's expectations. In this case, i_S is directed into node a and so one might guess that i_1 and i_2 should be directed out of that same node. Application of KCL at node a yields

$$i_S - i_1 - i_2 = 0 \tag{1.17}$$

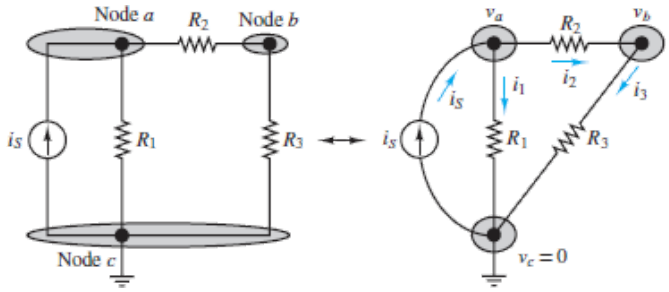


Figure 1.56 Illustration of node analysis

Page 40

whereas at node b

$$i_2 - i_3 = 0 \tag{1.18}$$

It is not necessary nor appropriate to apply KCL at the reference node since the resulting equation is dependent on the other two.

The branch currents in equations 1.17 and 1.18 can be expressed in terms of the node voltages using Ohm's law. For example:

$$i_1 = \frac{v_a - v_c}{R_1} \tag{1.19}$$

Similarly, for the other two branch currents

$$i_2 = \frac{v_a - v_b}{R_2} \tag{1.20}$$

$$i_3 = \frac{v_b - v_c}{R_3}$$

where v_c is the reference node voltage, freely chosen to be zero. These expressions for i_1, i_2, and i_3 can be substituted into equations 1.17 and 1.18 to obtain:

$$i_S - \frac{v_a}{R_1} - \frac{v_a - v_b}{R_2} = 0 \tag{1.21}$$

$$\frac{v_a - v_b}{R_2} - \frac{v_b}{R_3} = 0 \tag{1.22}$$

With a little practice, equations 1.21 and 1.22 can be obtained directly without introducing branch current variables. These equations can be reorganized and solved for v_a and v_b, assuming that i_S, R_1, R_2, and R_3 are known.

$$\left(\frac{1}{R_1} + \frac{1}{R_2}\right) v_a + \left(-\frac{1}{R_2}\right) v_b = i_S \tag{1.23}$$

$$\left(-\frac{1}{R_2}\right) v_a + \left(\frac{1}{R_2} + \frac{1}{R_3}\right) v_b = 0$$

Occasionally, it is helpful to redraw circuits in an equivalent but non-rectangular manner by viewing the circuit as a collection of circuit elements located between nodes. The right-hand portion of Figure 1.56 is constructed by drawing three node circles and then adding in the elements that sit between each pair of nodes. To successfully redraw a circuit it is imperative that the correct number of nodes is known. Thus, it is worthwhile to practice recognizing and counting nodes!

Page 41

FOCUS ON PROBLEM SOLVING

THE NODE VOLTAGE METHOD

1. Select a reference node. If the circuit has one or more voltage sources choose node that is attached to the largest number of voltage sources as the refere The voltage associated with each non-reference node will be relative to reference node, which is (for simplicity) assigned a value of 0 V.
2. Define voltage variables $v_1, v_2, \ldots, v_{n-1}$ for the remaining $n - 1$ nodes.
   - If the circuit contains m voltage sources, each of which is adjacent to a r with a known voltage (e.g., the reference node), the voltages at the c

adjacent nodes are known. Mark them as known.

- If the circuit contains ℓ additional voltage sources *not* adjacent to a r with a known voltage, create a "supernode" to enclose the nodes on sides of those voltage sources.

3. To generate $n - m - 1$ equations in the $n - m - 1$ unknown node voltages . . .
   - Apply KCL to each of the $n - m - 2\ell - 1$ nodes that are *not* part supernode.
   - Apply KCL to each of the ℓ supernodes.
   - Use the ℓ supernode voltage sources to write ℓ equations relating the unknown supernode voltages.
4. Collect coefficients for each of the $n - m - 1$ variables and solve the li system of $n - m - 1$ equations.

This procedure can be used to find a solution for any circuit. A good approach first practice solving circuits without any voltage sources and then learn to deal the added complexity of circuits with voltage sources.

EXAMPLE 1.13 Node Voltage Method: Solving for Branch Currents

Problem

Solve for the node voltages and the branch currents in the circuit of Figure 1.57.

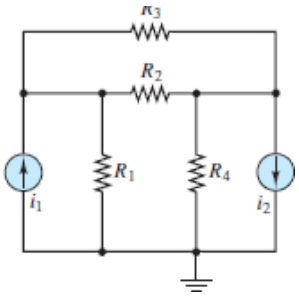


Figure 1.57

Solution

Given: Source currents, resistor values.

Find: All node voltages and branch currents.

Schematics, Diagrams, Circuits, and Given Data: $i_1 = 10$ mA; $i_2 = 50$ mA; $R_1 = 1$ kΩ; $R_2 = 2$ kΩ; $R_3 = 10$ kΩ; $R_4 = 2$ kΩ.

Analysis: Follow the steps outlined in the Focus on Problem Solving box "Node Voltage Method."

1. The node at the bottom of the circuit is chosen as the reference. There are no voltage sources in the circuit and each node is attached to four elements so any node would serve equally well as the reference.
2. The circuit of Figure 1.57 is shown again in Figure 1.58, with two non-reference nodes and the associated node voltage variables v_1 and v_2.
3. Apply KCL at each node and use Ohm's law to express branch currents in terms of node voltages to obtain:

$$i_1 - \frac{v_1 - 0}{R_1} - \frac{v_1 - v_2}{R_2} - \frac{v_1 - v_2}{R_3} = 0 \qquad \text{node 1}$$

$$\frac{v_1 - v_2}{R_2} + \frac{v_1 - v_2}{R_3} - \frac{v_2 - 0}{R_4} - i_2 = 0 \qquad \text{node 2}$$

4. Collect coefficients and reorganize the equations. Note that each resistance is given as kΩ and each current is given as mA so each voltage has the unit V.

$$\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) v_1 + \left(-\frac{1}{R_2} - \frac{1}{R_3}\right) v_2 = i_1 \qquad \text{node 1}$$

$$\left(-\frac{1}{R_2} - \frac{1}{R_3}\right) v_1 + \left(\frac{1}{R_2} + \frac{1}{R_3} + \frac{1}{R_4}\right) v_2 = -i_2 \qquad \text{node 2}$$

With some manipulation, the equations can be expressed as follows:

$$1.6 v_1 - 0.6 v_2 = 10$$
$$-0.6 v_1 + 1.1 v_2 = -50$$

These equations may be solved simultaneously to obtain

$$v_1 = -13.57 \text{ V}$$
$$v_2 = -52.86 \text{ V}$$

Knowing the node voltages, each branch current can be determined. For example, the current through R_3 (the 10-kΩ resistor) is given by

$$i_{R_3} = \frac{v_1 - v_2}{10{,}000} = 3.93 \text{ mA}$$

The positive value for *iR*3 indicates that the initial (arbitrary) choice of direction for this current is the same as its actual direction. Consider the current through R_1:

$$i_{R_1} = \frac{v_1}{1{,}000} = -13.57 \text{ mA}$$

Here, the value is negative, which indicates that the actual direction of this current is from ground to node 1, opposite of what was assumed, but as it must be, since the voltage at node 1 is negative with respect to ground. The branch-by-branch analysis may be continued to verify that $i_{R_2} = 19.65$ mA and $i_{R_4} = -26.43$ mA.

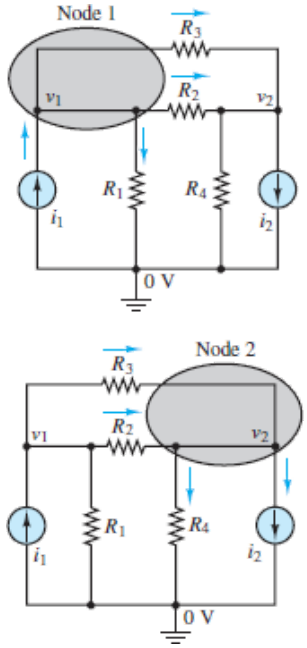


Figure 1.58

Page 43

EXAMPLE 1.14 Node Analysis: Solving for Node Voltages

Problem

Write the node equations and solve for the node voltages in the circuit of Figure 1.59.

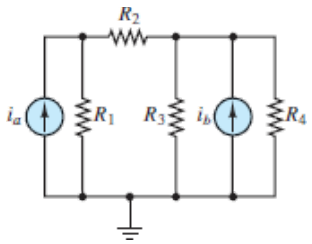


Figure 1.59

Solution

Given: Source currents, resistor values.

Find: All node voltages.

Schematics, Diagrams, Circuits, and Given Data: $i_a = 1$ mA; $i_b = 2$ mA; $R_1 = 1\,\text{k}\Omega$; $R_2 = 500\,\Omega$; $R_3 = 2.2\,\text{k}\Omega$; $R_4 = 4.7\,\text{k}\Omega$.

Analysis: Follow the steps outlined in the Focus on Problem Solving box "Node Voltage Method."

1. The bottom node is connected to more elements than either of the other two nodes so it is chosen as the reference node.
2. See Figure 1.60. There are two non-reference nodes, labeled v_a and v_b, in the circuit. There are no voltage sources in the circuit.

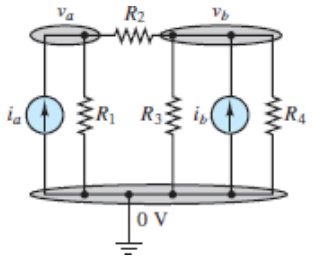


Figure 1.60

3. Apply KCL at each node and use Ohm's law to express branch currents in terms of node voltages to obtain:

$$i_a - \frac{v_a}{R_1} - \frac{v_a - v_b}{R_2} = 0 \qquad \text{node } a$$

$$\frac{v_a - v_b}{R_2} + i_b - \frac{v_b}{R_3} - \frac{v_b}{R_4} = 0 \qquad \text{node } b$$

4. Collect coefficients and reorganize the equations. Note that each resistance is given as kΩ and each current is given as mA so each voltage has the unit V.

$$\left(\frac{1}{R_1}+\frac{1}{R_2}\right)v_a + \left(-\frac{1}{R_2}\right)v_b = i_a$$
$$\left(-\frac{1}{R_2}\right)v_a + \left(\frac{1}{R_2}+\frac{1}{R_3}+\frac{1}{R_4}\right)v_b = i_b$$

Numerical values can be plugged into these equations to find:

$$3v_a - 2v_b = 1$$
$$-2v_a + 2.67v_b = 2$$

Multiply the second equation by $\frac{3}{2}$ and add the result to the first equation to find $v_b = 2$ V.. Plug v_b into either equation to find $v_a = 1.667$ V.

EXAMPLE 1.15 Using MatLab® to Solve a 3 × 3 System of Linear Equations

Problem

Use the node voltage analysis to determine the voltage v in the circuit of Figure 1.61. Assume that $R_1 = 2\Omega$, $R_2 = 1\Omega$, $R_3 = 4\Omega$, $R_4 = 3\Omega$, $i_1 = 2$ A, and $i_2 = 3$ A.

Page 44

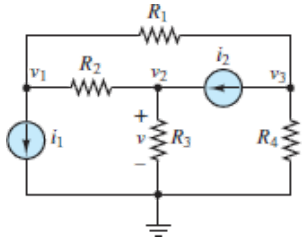


Figure 1.61 Circuit for Example 1.15

Solution

Given: Values of the resistors and the current sources.

Find: Voltage across R_3.

Analysis: Refer to Figure 1.61 and the steps in the Focus on Problem Solving box "Node Voltage Method."

1. Select one node as the reference and label it. There are no voltage sources in the circuit and each node is attached to three elements so any node would serve equally well as the reference.
2. Define node voltages v_1, v_2, v_3 for the three non-reference nodes.
3. Apply KCL at each of the $n - 1$ nodes, using Ohm's law to express the current through a resistor as the difference between the two adjacent node voltages divided by the resistance.

$$\begin{aligned}
\frac{v_3 - v_1}{R_1} + \frac{v_2 - v_1}{R_2} - i_1 &= 0 \\
\frac{v_1 - v_2}{R_2} - \frac{v_2}{R_3} + i_2 &= 0 \\
\frac{v_1 - v_3}{R_1} - \frac{v_3}{R_4} - i_2 &= 0
\end{aligned}$$

4. Collect the coefficients of each node voltage and reorganize the equations.

$$\begin{aligned}
-\left(\frac{1}{R_1}+\frac{1}{R_2}\right)v_1 + \left(\frac{1}{R_2}\right)v_2 + \left(\frac{1}{R_1}\right)v_3 &= i_1 \\
\left(\frac{1}{R_2}\right)v_1 - \left(\frac{1}{R_2}+\frac{1}{R_3}\right)v_2 &= -i_2 \\
\left(\frac{1}{R_1}\right)v_1 - \left(\frac{1}{R_1}+\frac{1}{R_4}\right)v_3 &= i_2
\end{aligned}$$

Multiply both sides of each equation by the common denominator on the left side. The common denominators are R_1R_2 for node 1, R_2R_3 for node 2, and R_1R_4 for node 3. Plug in values for the resistors and current sources.

$$\begin{aligned}
(-1-2)\,v_1 + 2v_2 + 1v_3 &= 4 \quad \text{node 1} \\
4v_1 + (-1-4)\,v_2 + 0v_3 &= -12 \quad \text{node 2} \\
3v_1 + 0v_2 + (-2-3)\,v_3 &= 18 \quad \text{node 3}
\end{aligned}$$

By including zero coefficients explicitly, all three voltage variables are now present in each equation. The resulting system of three equations in three unknowns can be solved by many handheld calculators. An alternative is Matlab. To solve using Matlab it is necessary to write the equations in matrix form.

$$\begin{bmatrix} -3 & 2 & 1 \\ 4 & -5 & 0 \\ 3 & 0 & -5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -12 \\ 18 \end{bmatrix}$$

In general, these equations can be written using compact notation as

$$Ax = b$$

where x is a 3 × 1 column vector whose elements are the node voltages v_1, v_2, and v_3. In Matlab the 3 × 3 A matrix and the 3 × 1 b column vector are entered as shown in Figure 1.62.

$$A = [-3\ \ 2\ \ 1\ ;\ 4\ -5\ \ 0\ ;\ 3\ \ 0\ -5]$$

$$b = [4\ ;\ -12\ ;\ 18] \quad (\text{or } b = [4\quad -12\quad 18]')$$

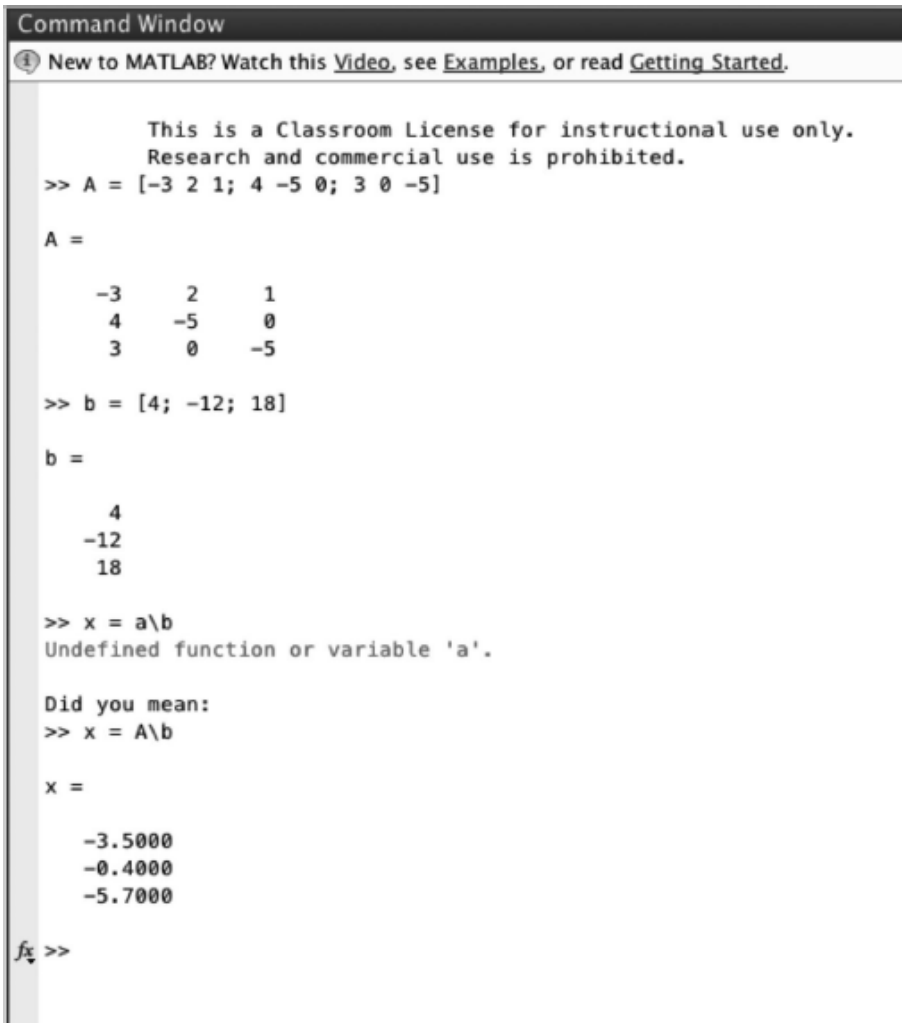


Figure 1.62 Typical Matlab command window. User-entered data follows the ≫ prompt. Note that Matlab is case sensitive, as shown at the fourth prompt. (*The MathWorks, Inc.*)

The apostrophe at the far right of the above equation is the Matlab transpose operator. It is used here to change a 1 × 3 row matrix into a 3 × 1 column vector. The solution for x is computed in Matlab by writing $x = A\backslash b$ to yield

$$x = [-3.5\text{V} \quad -0.4\text{V} \quad -5.7\text{V}]'$$

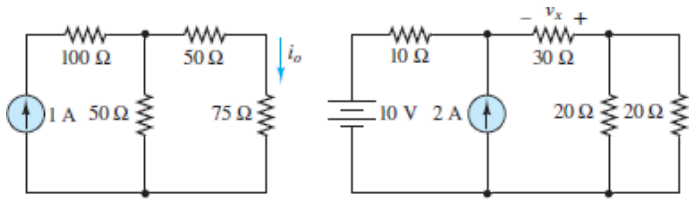
which are the three node voltages $[v_1 \quad v_2 \quad v_3]'$. The solution for the voltage drop v across R_3 is

$$v = v_2 = -0.4 \text{ V}$$

Page 46

CHECK YOUR UNDERSTANDING

Find i_o and v_x in the circuits on the left and right, respectively, using the node voltage method.



Answer: 0.2857 A; −18 V

CHECK YOUR UNDERSTANDING

In Example 1.14, use the two node voltages to verify that KCL is indeed satisfied at each node.

CHECK YOUR UNDERSTANDING

Repeat Example 1.15 when the directions of the current sources are opposite those shown in Figure 1.61. Find v.

Answer: $v = 0.4$ V

The Node Voltage Method with Voltage Sources

The circuits in the preceding examples did not contain voltage sources. However, in practice, they are quite common. To illustrate how the node voltage method is applied to such circuits, consider the circuit in Figure 1.63. Verify that this circuit has $n = 4$ total nodes.

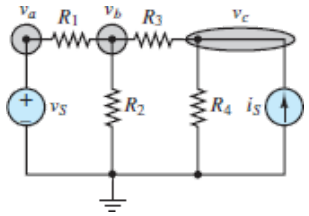


Figure 1.63 Node analysis with voltage sources

When voltage sources are present, it is advantageous to pick the reference node so that at least one of those voltage sources is attached to it. In Figure 1.63, the reference node, denoted by the ground symbol, is assumed to have a value of 0 V.

The remaining three (4 − 1 = 3) node voltages are labeled v_a, v_b, and v_C as shown Figure 1.63. Since node v_a is adjacent to the voltage source its value $v_a = V_S$ is known relative to the reference node. The only two unknown node voltages are v_b and v_c. Apply KCL at those two nodes.

At node *b:*

$$\frac{v_a - v_b}{R_1} - \frac{v_b - 0}{R_2} - \frac{v_b - v_c}{R_3} = 0 \tag{1.24a}$$

At node *c:*

$$\frac{v_b - v_c}{R_3} - \frac{v_c}{R_4} + i_S = 0 \tag{1.24b}$$

Substitute for v_a in equation 1.24a.

$$\frac{v_S - v_b}{R_1} - \frac{v_b}{R_2} - \frac{v_b - v_c}{R_3} = 0 \tag{1.25}$$

Finally, collect the coefficients of the two unknown node voltages.

$$\begin{aligned}\left(\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}\right)v_b &+\left(-\frac{1}{R_3}\right)v_c=\frac{1}{R_1}v_s \\ \left(-\frac{1}{R_3}\right)v_b &+\left(\frac{1}{R_3}+\frac{1}{R_4}\right)v_c=i_S\end{aligned} \tag{1.26}$$

The resulting system of two equations in two unknowns can now be solved.

EXAMPLE 1.16 Solution when a Voltage Source Is Not Adjacent to the Reference Node

Problem

Use the Node Voltage Method to determine the node voltages and the current i through the voltage source v_{S2} in the circuit of Figure 1.64. Assume that $R_1 = 2\Omega, R_2 = 2\Omega, R_3 = 4\Omega, R_4 = 3\Omega$, $v_{S1} = 2$ V, and $v_{S2} = 3$ V.

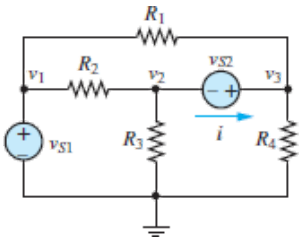


Figure 1.64 Circuit for Example 1.16

Solution

Given: Resistance values; current and voltage source values.

Find: The current i through the voltage source.

Analysis: Refer to Figure 1.64 and the steps in the Focus on Problem Solving box "Node Voltage Method."

1. Select a reference node and label it. There are two ($m = 2$) voltage sources in the circuit. Each node is attached to one voltage source and two resistors so any node would serve equally well as the reference.
2. Define three non-reference node voltages v_1, v_2, and v_3. The voltage source v_{S1} is adjacent to a node with a known voltage. (The reference node voltage is freely chosen to be zero.) Node v_1 is the other node adjacent to that voltage source.

Thus, $v_1 = v_{S1} + 0 = v_{S1}$ is known relative to the reference node. The only two unknown node voltages are v_2 and v_3.

The voltage source v_{S2} is not adjacent to a node with a known voltage. Create a "supernode" enclosing v_{S2} and the nodes v_2 and v_3 as shown in Figure 1.65.

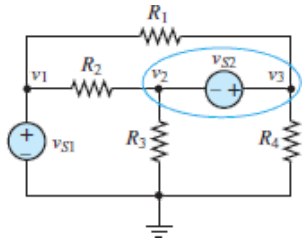


Figure 1.65 Circuit for Example 1.16 with "supernode"

3. Assume all currents are entering the supernode and apply KCL.

$$\frac{v_1 - v_2}{R_2} + \frac{0 - v_2}{R_3} + \frac{v_1 - v_3}{R_1} + \frac{0 - v_3}{R_4} = 0 \qquad \text{KCL at supernode boundary}$$

Use the voltage source v_{S2} to relate the two unknown node voltages v_2 and v_3.

$$v_3 = v_2 + v_{S2} \qquad \text{Within supernode}$$

Page 48

4. Collect coefficients of the unknown node voltages and substitute values for the known parameters.

$$\begin{aligned} 9v_2 + 10v_3 &= 24 \qquad \text{KCL at supernode boundary} \\ -v_2 + v_3 &= 3 \qquad \text{Within supernode} \end{aligned}$$

This system of two equations in two unknowns can be solved analytically by multiplying the second equation by 9 and adding the result to the first equation to eliminate v_2. The result is

$$19v_3 = 51 \qquad \text{or} \qquad v_3 = \frac{51}{19} \approx 2.68 \text{ V}$$

Use this result to find v_2.

$$v_2 = v_3 - 3 \qquad \text{or} \qquad v_2 = \frac{-6}{19} \approx -0.32 \text{ V}$$

Apply KCL at node v_3 to find current i through the voltage source v_{S2}.

$$i = \frac{v_3 - v_1}{R_1} + \frac{v_3}{R_4} \approx \frac{2.68 - 2}{2} + \frac{2.68}{3} \approx 1.24 \text{ A}$$

Comments: Knowing all three node voltages, the current through each resistor can be computed as follows: $i_1 = |v_3 - v_1|/R_1$ (to left), $i_2 = |v_1 - v_2|/R_2$ (to right), $i_3 = |v_2|/R_3$ (upward), and $i_4 = |v_3|/R_4$ (downward).

CHECK YOUR UNDERSTANDING

Repeat the exercise of Example 1.16 when the direction of the voltage source v_{S1} is opposite that shown in Figure 1.64. Find the node voltages and i.

Answer: $v_1 = -2$ V, $v_2 \approx -2.84$ V, $v_3 \approx 0.16$ V, and $i \approx 1.13$ A

1.8 THE MESH CURRENT METHOD

Another method of circuit analysis employs **mesh currents**. The objective, similar to that of the node voltage method, is to generate a system of *n linearly independent* equations in *n* unknown mesh currents. In this method, each mesh in a circuit is assigned a mesh current and Kirchhoff's voltage law (KVL) is applied around each mesh.

It is important to recall that mesh currents are not the same as branch currents. The perspective taken in the mesh current method is that there is one current circulating within each mesh and that branch currents in the circuit are comprised of these mesh currents. Specifically, when a branch is part of only one mesh, the branch current is the same as that mesh current. However, when a branch is shared by two meshes, the branch current is the sum or difference of the two mesh currents.

In the mesh current method it is necessary to assume a direction for the circulation of each mesh current. A helpful convention is to assume that *all mesh currents circulate in the clockwise (CW) direction.* With this convention, when a branch is shared by two meshes, the branch current is equal to the difference of two mesh currents. This result is illustrated in Figure 1.66 where Page 49the current through resistor R_2 is the difference of i_1 and i_2 since i_1 and i_2 are oppositely directed through R_2.

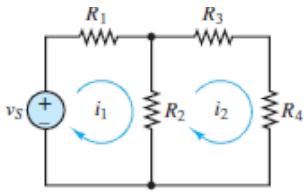


Figure 1.66 Two meshes and two mesh currents

It is helpful to apply KVL around a mesh in the same direction (e.g., CW) used to define the mesh current. Ohm's law implies that the *net* current through a resistor is directed from high to low voltage, as shown in Figure 1.67. Thus, when KVL is applied to mesh i_1 in Figure 1.68 the resulting Ohm's law expressions are

$$v_1 = i_1 R_1 \tag{1.27}$$

and

$$v_2 = (i_1 - i_2) R_2 \tag{1.28}$$

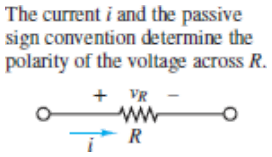


Figure 1.67 Ohm's law implies that current is directed from high (+) to low (−) potential

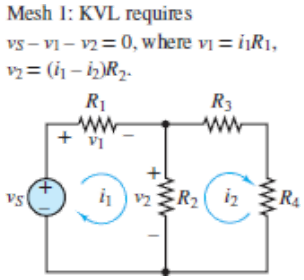


Figure 1.68 Assignment of currents and voltages around mesh 1

Notice that the *net* current through R_2 in the direction of mesh current i_1 is $(i_1 - i_2)$.

Thus, the KVL equation for mesh i_1 is

$$v_S - i_1 R_1 - (i_1 - i_2) R_2 = 0 \qquad \text{mesh 1} \tag{1.29}$$

When KVL is applied to mesh i_2, the result (see Figure 1.69) is

$$(i_1 - i_2)R_2 - i_2R_3 - i_2R_4 = 0 \qquad \text{mesh 2} \tag{1.30}$$

Mesh 2: KVL requires

$$v_2 - v_3 - v_4 = 0$$

where

$$v_2 = (i_1 - i_2)R_2$$

$$v_3 = i_2R_3$$

$$v_4 = i_2R_4$$

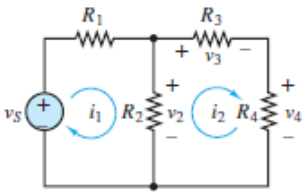


Figure 1.69 Assignment of voltages around mesh 2

Multiply both sides of the mesh 2 equation by −1. Then, collect coefficients of i_1 and i_2 in each equation to yield the following system of equations:

$$(R_1 + R_2)i_1 - R_2i_2 = v_S \tag{1.31}$$

$$-R_2i_1 + (R_2 + R_3 + R_4)i_2 = 0 \tag{1.32}$$

These two equations can be solved simultaneously for the two independent mesh current variables i_1 and i_2. The branch current through R_2 can then be found as well. If the resulting numerical answer for a mesh current is negative, then the actual direction for that mesh current is opposite of the defined direction. A careful determination of the voltage drops around each mesh, one mesh at a time, and in accord with the passive sign convention for Ohm's law, is necessary for success.

LO

FOCUS ON PROBLEM SOLVING

THE MESH CURRENT METHOD

1. Choose a circulation convention (either CW or CCW) for the mesh currents KVL.
2. Define mesh current variables $i_1, i_2, \ldots, i_n$ for each of the n meshes.
   - If the circuit contains no current sources, apply KVL around each mes generate n independent KVL equations in the n mesh current varia Jump ahead to step 7.
   - If the circuit contains m current sources, there will be $n - m$ KVL equat and m current source equations. Proceed to step 3.

Page 50

3. If there are ℓ meshes that contain no current source, apply KVL around such mesh to generate ℓ KVL equations. Use Ohm's law to express each res voltage drop in terms of the mesh currents.
4. Generate one current source equation $i_j = \pm i_S$ for each current source i_S borders only one mesh, where i_j is the mesh current.
5. Generate one current source equation $i_j - i_k = \pm i_S$ for each current source i_S borders two meshes, where i_j and i_k are the two mesh currents.
6. Define $n - m - \ell$ "supermeshes" such that each current source is conta within a supermesh but is not on the boundary of a supermesh. Apply K around each supermesh to generate $n - m - \ell$ additional KVL equations. Ohm's law to express each resistor voltage drop in terms of the adjacent n currents.
7. Collect coefficients for each of the n variables and solve the linear system equations.
   - Each current source equation $i_j = \pm i_S$ can be used to reduce the total nur of equations and variables by direct substitution.
8. Use the known mesh currents to solve for any or all branch currents in circuit. Any voltage drop can be found by applying Ohm's law and, w necessary, KVL.

EXAMPLE 1.17 The Mesh Current Method: Solving for Mesh Currents in a Circuit with Two Meshes

Problem

Find the mesh currents in the circuit of Figure 1.70.

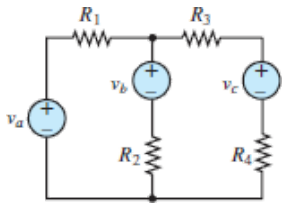


Figure 1.70

Solution

Given: Source voltages; resistor values.

Find: The Mesh currents.

Schematics, Diagrams, Circuits, and Given Data: $v_a = 10$ V; $v_b = 9$ V; $v_c = 1$ V; $R_1 = 5\,\Omega$; $R_2 = 10\,\Omega$; $R_3 = 5\,\Omega$; $R_4 = 5\,\Omega$.

Page 51

Analysis: Refer to Figures 1.70 and 1.71 and the steps in the Focus on Problem Solving box "The Mesh Current Method."

1. Select a clockwise circulation convention.
2. Note that there are two meshes in the circuit and define clockwise mesh current variables i_1 and i_2. There are no current sources in the circuit.
3. Apply KVL to each mesh and use Ohm's law to express each resistor voltage drop in terms of the mesh currents to generate two equations.

$$\begin{aligned} v_a - R_1 i_1 - v_b - R_2(i_1 - i_2) = 0 \quad & \text{mesh 1} \\ R_2(i_1 - i_2) + v_b - R_3 i_2 - v_c - R_4 i_2 = 0 \quad & \text{mesh 2} \end{aligned}$$

4. Collect coefficients and enter parameter values to yield the following system of linear equations:

$$\begin{aligned} 15 i_1 - 10 i_2 = v_a - v_b = 1 \quad & \text{mesh 1} \\ -10 i_1 + 20 i_2 = v_b - v_c = 8 \quad & \text{mesh 2} \end{aligned}$$

Multiply the mesh 1 equation by 2 and add the result to the mesh 2 equation to find i_1. Substitute for i_1 in either equation to find i_2. The results are

$$i_1 = 0.5\,\text{A} \quad \text{and} \quad i_2 = 0.65\,\text{A}$$

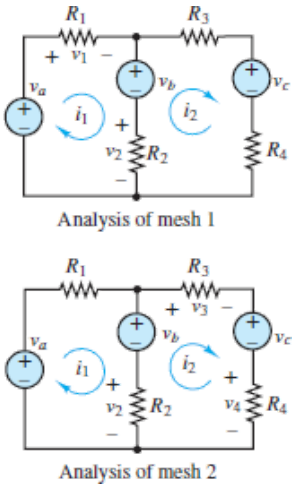


Figure 1.71

EXAMPLE 1.18 The Mesh Current Method: Using MatLab to Solve for Mesh Currents in a Circuit with Three Meshes

Problem

The circuit of Figure 1.72 is a simplified DC circuit model of a three-wire electrical distribution service to residential and commercial buildings. The two ideal sources and the resistances R_4 and R_5 represent the equivalent network of the distribution system; R_1 and R_2 represent 110-V lighting and utility loads rated at 800 and 300 W, respectively. Resistance R_3 represents a 220-V heating load rated at 3 kW. Determine the voltages across the three loads.

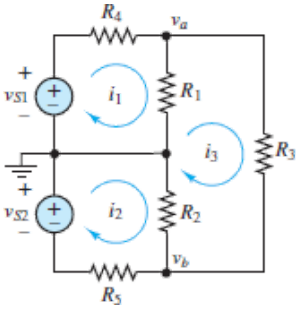


Figure 1.72

Solution

Given: The values of the voltage sources and resistors in the circuit of Figure 1.72 are $v_{S1} = v_{S2} = 110$ V; $R_4 = R_5 = 1.3\,\Omega$; $R_1 = 15\,\Omega$; $R_2 = 40\,\Omega$; $R_3 = 16\,\Omega$..

Find: i_1, i_2, i_3, v_a and v_b.

Analysis: Refer to Figure 1.72 and the steps in the Focus on Problem Solving box "The Mesh Current Method."

1. Select a clockwise circulation convention.
2. Note that there are three meshes in the circuit and define clockwise mesh current variables i_1, i_2, and i_3 as shown in Figure 1.72. There are no current sources in the circuit.
3. Apply KVL to each mesh and use Ohm's law to represent the voltage drop across each resistor in terms of the mesh currents directly.

$$\begin{aligned} &\text{Mesh 1:} && v_{S1} - R_4 i_1 - R_1(i_1 - i_3) = 0 \\ &\text{Mesh 2:} && v_{S2} - R_2(i_2 - i_3) - R_5 i_2 = 0 \\ &\text{Mesh 3:} && -R_1(i_3 - i_1) - R_3 i_3 - R_2(i_3 - i_2) = 0 \end{aligned}$$

Page 52

4. Collect coefficients to obtain the following system of three equations in three unknown mesh currents.

$$\begin{aligned} -(R_1 + R_4) i_1 \quad &+ \quad R_1 i_3 = -v_{S1} \\ -(R_2 + R_5) i_2 \quad &+ \quad R_2 i_3 = -v_{S2} \\ R_1 i_1 + R_2 i_2 \quad &- \quad (R_1 + R_2 + R_3) i_3 = 0 \end{aligned}$$

Enter numerical values for the parameters and express the equations in matrix form.

$$\begin{bmatrix} -16.3 & 0 & 15 \\ 0 & -41.3 & 40 \\ 15 & 40 & -71 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} -110 \\ -110 \\ 0 \end{bmatrix}$$

This form can be more simply represented as the product of a resistance matrix $[R]$ and a mesh current vector $[I]$ set equal to a voltage source vector $[V]$.

$$[R][I] = [V]$$

with a solution of

$$[I] = [R]^{-1}[V]$$

The solution for the mesh current vector can be found using an analytic or numerical technique. In this problem, Matlab was used to compute the inverse

$[R]^{-1}$ of the 3×3 $[R]$ matrix.

$$[R]^{-1} = \begin{bmatrix} -0.1072 & -0.0483 & -0.0499 \\ -0.0483 & -0.0750 & -0.0525 \\ -0.0499 & -0.0525 & -0.0542 \end{bmatrix}$$

The value of each mesh current is now determined.

$$[I] = [R]^{-1}[V] = \begin{bmatrix} -0.1072 & -0.0483 & -0.0499 \\ -0.0483 & -0.0750 & -0.0525 \\ -0.0499 & -0.0525 & -0.0542 \end{bmatrix} \begin{bmatrix} -110 \\ -110 \\ 0 \end{bmatrix} = \begin{bmatrix} 17.11 \\ 13.57 \\ 11.26 \end{bmatrix}$$

Therefore, we find

$$i_1 = 17.11 \text{ A} \quad i_2 = 13.57 \text{ A} \quad i_3 = 11.26 \text{ A}$$

The two unknown node voltages V_a and V_b are easily calculated using Ohm's law and the mesh currents. Notice the passive sign convention used in the following calculations!

$$\begin{aligned} v_a - 0 &= R_1(i_1 - i_3) \\ v_a &= 87.75 \text{ V} \\ v_b - 0 &= R_2(i_3 - i_2) \\ v_b &= -92.40 \text{ V} \end{aligned}$$

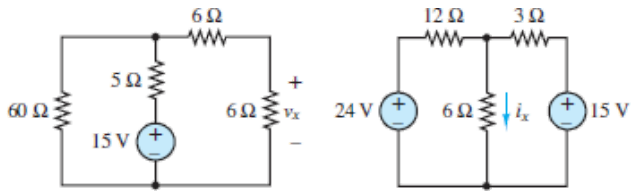
The values of the node voltages v_a and v_b are relative to the reference node. Verify that KVL holds for each mesh to check your understanding.

Comments: The inverse matrix computation is numerically inefficient compared to the Matlab left division computation used in Example 1.15.

Page 53

CHECK YOUR UNDERSTANDING

Use the mesh current method to find the unknown voltage v_x in the circuit on the left.



Use the mesh current method to find the unknown current i_x in the circuit on the right.

Answer: 5 V; 2 A

CHECK YOUR UNDERSTANDING

Repeat the exercise of Example 1.18, using the Node Voltage Method instead of the Mesh Current Method.

The Mesh Current Method with Current Sources

The circuits in the preceding examples contained no current sources. However, it is common, in practice, to encounter current sources in circuits. The relevant steps found in the *Focus on Problem Solving* section are listed below with added comments.

Step 1: *Choose a circulation convention (either CW or CCW) for the mesh currents and KVL.*

- In this book all mesh currents are chosen to have a clockwise orientation.

Step 2: *Define mesh current variables $i_1 i_2, \ldots, i_n$ for each of the n meshes. If the circuit contains m current sources, there will be $n - m$ KVL equations and m current source equations.*

- There are two meshes in the circuit shown in Figure 1.73. Notice that $n = 2$ and $m = 1$ such that there will be 1 KVL equation and 1 current source equation. Two mesh currents are defined as $i1$ and $i2$.

Step 3: *If there are ℓ meshes that contain no current source, apply KVL around each such mesh to generate ℓ KVL equations. Use Ohm's law to express each resistor voltage drop in terms of the mesh currents.*

- In Figure 1.73 only the i_1 mesh does not contain a current source, so $\ell = 1$. Apply KVL around that mesh.

$$v_S - R_1 i_1 - R_2 (i_1 - i_2) = 0 \tag{1.33}$$

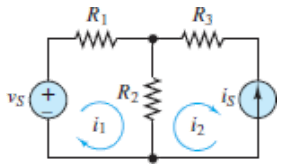


Figure 1.73 Mesh analysis with current sources

Step 4: *Generate one current source equation $i_j = \pm i_S$ for each current source i_S that borders only one mesh, where i_j is the mesh current.*

- In Figure 1.73, the i_2 mesh contains a current source that does not also border mesh i_1. Thus:

$$i_2 = -i_S \tag{1.34}$$

Step 5: *Generate one current source equation $i_j - i_k = \pm i_S$ for each current source i_S that borders two meshes, where i_j and i_k are the two mesh currents.*

- There are no current sources in Figure 1.73 that border two meshes.

Step 6: *Define $n - m - \ell$ independent "supermeshes" such that each current source is contained within a supermesh but is not on the boundary of a supermesh. Apply KVL around each supermesh to generate $n - m - \ell$ additional independent KVL equations. Use Ohm's law to express each resistor voltage drop in terms of the adjacent mesh currents.*

- There is no need for a supermesh in the circuit shown in Figure 1.73 since $n - m - \ell = 0$. In other words, there are already two linearly independent equations for the two mesh currents i_1 and i_2.

Step 7: *Collect coefficients for each of the n variables and solve the linear system of n equations. Each current source equation $i_j = \pm i_S$ can be used to reduce the total number of equations and variables by direct substitution.*

- Use equation 1.34 to substitute for i_2 in equation 1.33. The result is

$$i_1 = \frac{v_S - i_S R_2}{R_1 + R_2} \tag{1.35}$$

Step 8: *Use the known mesh currents to solve for any or all branch currents in the circuit. Any voltage drop can be found by applying Ohm's law and, when necessary, KVL.*

- For the circuit in Figure 1.73, the current through R_1 is i_1 and the current through R_3 is i_S. The current through R_2 is $i_1 - i_2$. The change in voltage across the current source is given by KVL as:

$$i_S R_3 + (i_1 - i_2) R_2$$

EXAMPLE 1.19 Mesh Analysis: Three Meshes and One Current Source

Problem

Find the mesh currents in the circuit of Figure 1.74.

Solution

Given: Source and resistor values.

Find: Mesh currents.

Page 55

Schematics, Diagrams, Circuits, and Given Data: $i_S = 0.5$ A; $v_S = 6$ V; $R_1 = 3\Omega$; $R_2 = 8\Omega$; $R_3 = 6\Omega$; $R_4 = 4\Omega$.

Analysis: Refer to Figure 1.74 and the steps in the Focus on Problem Solving box "The Mesh Current Method."

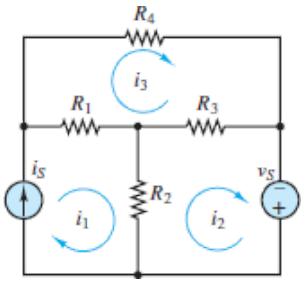


Figure 1.74

Step 1: *Choose a circulation convention (either CW or CCW) for the mesh currents and KVL.*

- In this book all mesh currents are chosen to have a clockwise orientation.

Step 2: *Define mesh current variables $i_1, i_2, \ldots, i_n$ for each of the n meshes. If the circuit contains m current sources, there will be $n - m$ KVL equations and m current source equations.*

- There are three meshes and one current source in the circuit shown in Figure 1.74. Thus, $n = 3$ and $m = 1$ such that there will be two KVL equations and one current source equation. Three mesh currents are defined as i_1, i_2, and i_3.

Step 3: *If there are ℓ meshes that contain no current source, apply KVL around each such mesh to generate ℓ KVL equations. Use Ohm's law to express each resistor voltage drop in terms of the adjacent mesh currents.*

- In Figure 1.74 the i_2 and i_3 meshes do not contain a current source, so $\ell = 2$. Apply KVL around each of those meshes.

$$-R_2(i_2 - i_1) - R_3(i_2 - i_3) + v_S = 0 \quad \text{mesh 2}$$
$$-R_1(i_3 - i_1) - R_4 i_3 - R_3(i_3 - i_2) = 0 \quad \text{mesh 3}$$

Step 4: *Generate one current source equation $i_j = \pm i_S$ for each current source i_S that borders only one mesh, where ij is the mesh current.*

- In Figure 1.74, the i_1 mesh contains a current source that does not also border the other meshes. Thus:

$$i_1 = i_S \quad \text{mesh 1}$$

Step 5: *Generate one current source equation $i_j - i_k = \pm i_S$ for each current source i_S that borders two meshes, where i_j and i_k are the two mesh currents.*

- There are no current sources in Figure 1.74 that border two meshes.

Step 6: *Define $n - m - \ell$ independent "supermeshes" such that each current source is contained within a supermesh but is not on the boundary of a supermesh. Apply KVL around each supermesh to generate $n - m - \ell$ additional independent KVL equations. Use Ohm's law to express each resistor voltage drop in terms of the adjacent mesh currents.*

- There is no need for a supermesh in the circuit shown in Figure 1.74 since $n - m - \ell = 0$. In other words, there are already three linearly independent equations for the three mesh currents i_1, i_2, and i_3.

Step 7: *Collect coefficients for each of the n variables and solve the linear system of n equations. Each current source equation* $i_j = \pm i_S$ *can be used to reduce the total number of equations and variables by direct substitution.*

- Use the mesh 1 equation to substitute for i_1 in the mesh 2 and 3 equations.

$$14i_2 - 6i_3 = 10$$
$$-6i_2 + 13i_3 = 1.5$$
$$i_2 = 0.95\,\text{A} \qquad i_3 = 0.55\,\text{A}$$

Page 56

Step 8: *Use the known mesh currents to solve for any or all branch currents in the circuit. Any voltage drop can be found by applying Ohm's law and, when necessary, KVL.*

- For the circuit in Figure 1.74, the current through R_1 is $i_1 - i_3$ and the current through R_2 is $i_1 - i_2$. The current through R_3 is $i_2 - i_3$ and the current through R_4 is i_3. The change in voltage across the current source is given by KVL as:

$$(i_1 - i_3)R_3 + (i_1 - i_2)R_2$$

CHECK YOUR UNDERSTANDING

Use the mesh currents to find the branch currents in Example 1.19. Apply KCL at each node to validate.

1.9 THE NODE VOLTAGE AND MESH CURRENT METHODS WITH DEPENDENT SOURCES

When a dependent source is present in a circuit, node or mesh equations can be written by treating it in the same manner as an independent source. The value of the dependent source may appear in circuit diagrams as an additional unknown variable. In this case, the dependence upon another current or voltage in the circuit is expressed by a **constraint equation**. Typically, the constraint equation is simple and can be directly substituted into the node or mesh equations to eliminate the unknown source variable.

Consider, for example, the circuit of Figure 1.75, which is a simplified *model* of an amplifier based upon a bipolar junction transistor. This circuit has three meshes

and three nodes, including the reference node. Therefore, node analysis will yield two equations for two unknown node voltages. Mesh analysis will also yield only two equations for two unknown mesh currents because the left-most mesh current is determined by i_S. Furthermore, notice that the right-most mesh current is determined by βi_b. Let's try both approaches.

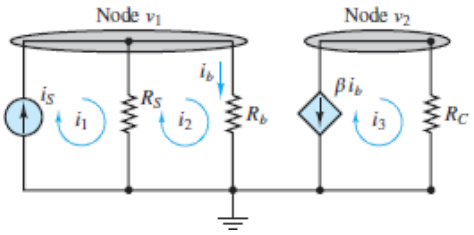


Figure 1.75 Circuit with dependent source

Apply the Mesh Current method and label each mesh i_1, i_2, and i_3 from left to right, such that

$$i_1 = i_S \qquad i_2 = i_b \qquad i_3 = -\beta i_b \tag{1.36}$$

Page 57

KVL around mesh 2 yields:

$$-(i_2 - i_1)R_S - i_2 R_b = 0 \tag{1.37}$$

which can be solved directly for i_2.

$$i_2 = i_1 \frac{R_S}{R_S + R_b} = i_S \frac{R_S}{R_S + R_b} \tag{1.38}$$

This equation is simply the result given by current division at node 1. Finally,

$$i_3 = -\beta i_b = -\beta i_2 = -\beta i_S \frac{R_S}{R_S + R_b} \tag{1.39}$$

The value of the dependent current source βi_b is explicit in the circuit diagram. Thus, there is no need for a separate constraint equation.

The Node Voltage method can also be used to solve this problem. Apply KCL at node v_1.

$$i_S = \frac{v_1}{R_S} + i_b = \frac{v_1}{R_S} + \frac{v_1}{R_b} \tag{1.40}$$

Apply KCL at node v_2.

$$\beta i_b + \frac{v_2}{R_C} = \beta \frac{v_1}{R_b} + \frac{v_2}{R_C} = 0 \tag{1.41}$$

where, by current division,

$$i_b = i_S \frac{R_S}{R_b + R_S} \tag{1.42}$$

Finally, the result is:

$$v_1 = i_S \frac{R_S R_b}{R_S + R_b} \tag{1.43}$$

$$v_2 = -\beta i_S \frac{R_S R_C}{R_S + R_b}$$

EXAMPLE 1.20 The Node Voltage Method with Dependent Sources

Problem

Find the node voltages in the circuit of Figure 1.76.

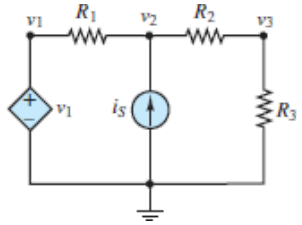


Figure 1.76

Solution

Given: Source current; resistor values; dependent voltage source constraint equation.

Find: Unknown node voltage v_2.

Schematics, Diagrams, Circuits, and Given Data: $i_S = 0.5$ A; $R_1 = 5\Omega$; $R_2 = 2\Omega$; $R_3 = 4\Omega$. Dependent source constraint equation: $v_1 = 2v_3$.

Analysis: Refer to Figure 1.76 and the steps in the Focus on Problem Solving box "The Node Voltage Method."

1. There are four nodes in the circuit. Select the bottom node of the circuit as the reference node.
2. Label the three non-reference nodes with node voltages v_1, v_2 and v_3. The dependent voltage source sets the node voltage v_1 equal to $2v_3$. As a result, consider v_1 determined (known) and so do not apply KCL at that node.
3. Apply KCL at nodes v_2 and v_3. Use Ohm's law to represent branch currents in terms of node voltages.

$$\frac{v_1 - v_2}{R_1} + i_S - \frac{v_2 - v_3}{R_2} = 0 \quad \text{node } v_2$$

$$\frac{v_2 - v_3}{R_2} - \frac{v_3 - 0}{R_3} = 0 \quad \text{node } v_3$$

   The constraint equation $v_1 = 2v_3$ can be used to substitute for $v1$ in the node v_2 equation.
4. Collect coefficients of v_2 and v_3.

$$\left(\frac{1}{R_1} + \frac{1}{R_2}\right) v_2 + \left(-\frac{2}{R_1} - \frac{1}{R_2}\right) v_3 = i_S$$

$$\left(\frac{1}{R_2}\right) v_2 - \left(\frac{1}{R_2} + \frac{1}{R_3}\right) v_3 = 0$$

   Enter numerical values to obtain

$$0.7\, v_2 - 0.9\, v_3 = 0.5$$
$$0.5\, v_2 - 0.75\, v_3 = 0$$

   Multiply the second equation by 7/5 and subtract the result from the first equation to find v_3 = 10/3 = 3.33 V. Substitute this value into either of the previous equations to find v_2 = 5. Finally, $v_1 = 2v_3$ = 6.66 V.

EXAMPLE 1.21 The Mesh Current Method with a Dependent Source Problem

Determine the voltage *gain* $G_v = v_2/v_1$ in the circuit of Figure 1.77.

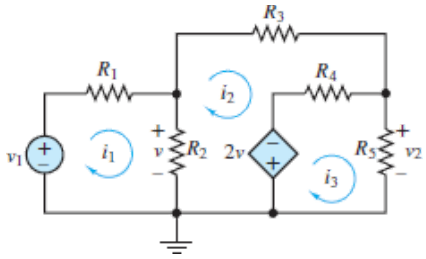


Figure 1.77 Circuit containing dependent voltage source

Page 59

Solution

Given: The values of the resistors are $R_1 = 1\Omega$; $R_2 = 0.5\Omega$; $R_3 = 0.25\Omega$; $R_4 = 0.25\Omega$; $R_5 = 0.25\Omega$.

Find: $G_v = v_2/v_1$.

Analysis: Refer to Figure 1.77 and the steps in the Focus on Problem Solving box "The Mesh Current Method" to find v_2 in terms of v_1.

Step 1: *Choose a circulation convention (either CW or CCW) for the mesh currents and KVL.*

- In this book all mesh currents are chosen to have a clockwise orientation.

Step 2: *Define mesh current variables $i_1, i_2, \ldots, i_n$ for each of the n meshes. If the circuit contains m current sources, there will be $n - m$ KVL equations and m current source equations.*

- There are three meshes in the circuit shown in Figure 1.77. Notice that $n = 3$ and $m = 0$ such that there will be 3 KVL equations and no current source equation. Three mesh currents are defined as i_1, i_2, and i_3.

Step 3: *If there are ℓ meshes that contain no current source, apply KVL around each such mesh to generate ℓ KVL equations. Use Ohm's law to express each resistor voltage drop in terms of the adjacent mesh currents.*

- Since there are no current sources in the circuit of Figure 1.77, $\ell = 3$. Apply KVL around each mesh. Note that Ohm's law is used to express the voltage drop across each resistor.

$$\begin{aligned} v_1 - R_1 i_1 - R_2(i_1 - i_2) &= 0 && \text{mesh 1} \\ -R_2(i_2 - i_1) - R_3 i_2 - R_4(i_2 - i_3) + 2v &= 0 && \text{mesh 2} \\ -2v - R_4(i_3 - i_2) - R_5 i_3 &= 0 && \text{mesh 3} \end{aligned}$$

Step 4: *Generate one current source equation $i_j = \pm i_S$ for each current source i_S that borders only one mesh, where i_j is the mesh current. It is important to correctly account for the direction of the mesh current relative to the current source.*

- There are no current sources in the circuit shown in Figure 1.77.

Step 5: *Generate one current source equation $i_j - i_k = \pm i_S$ for each current source i_S that borders two meshes, where i_j and i_k are the two mesh currents. It is important to correctly account for the direction of the mesh currents relative to the current source.*

- There are no current sources in the circuit shown in Figure 1.77.

Step 6: *Define $n - m - \ell$ independent "supermeshes" such that each current source is contained within a supermesh but is not on the boundary of a supermesh. Apply KVL around each supermesh to generate $n - m - \ell$ additional independent KVL equations. Use Ohm's law to express each resistor voltage drop in terms of the adjacent mesh currents.*

- There is no need for a supermesh in the circuit shown in Figure 1.77 since $n - m - \ell = 0$. In other words, there are already three linearly independent equations for the three mesh currents i_1, i_2, and i_3.

Page 60

Step 7: *Collect coefficients for each of the n variables and solve the linear system of n equations. Each current source equation $i_j = \pm i_S$ can be used to reduce the total number of equations and variables by direct substitution.*

- The constraint equation for the dependent voltage source is $2v = 2(i_1 - i_2)R_2$. Plug in for $2v$ in the mesh 2 and 3 equations to rewrite the system of equations.

$$\begin{aligned} (R_1 + R_2)i_1 - R_2 i_2 &= v_1 \\ (3R_2)i_1 - (3R_2 + R_3 + R_4)i_2 + (R_4)i_3 &= 0 \\ -2R_2 i_1 + (2R_2 + R_4)i_2 - (R_4 + R_5)i_3 &= 0 \end{aligned}$$

- This system of equations can be written in matrix form.

$$\begin{bmatrix} (R_1+R_2) & -R_2 & 0 \\ 3R_2 & -(3R_2+R_3+R_4) & R_4 \\ -2R_2 & (2R_2+R_4) & -(R_4-R_5) \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix}$$

- Enter numerical values for the resistors.

$$\begin{bmatrix} 1.5 & -0.5 & 0 \\ 1.5 & -2 & 0.25 \\ -1 & 1.25 & -0.5 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix}$$

- This system can be written in compact linear algebra notation as shown below, where the indices represent the number of rows and columns. The solution for i can be computed using the "left division" operator in Matlab

$$R_{33} i_{31} = v_{31}$$

- The results are

$$i_1 = 0.88 v_1$$
$$i_2 = 0.64 v_1$$
$$i_3 = -0.16 v_1$$

Step 8: *Use the known mesh currents to solve for any or all branch currents in the circuit. Any voltage drop can be found by applying Ohm's law and, when necessary, KVL.*

- v_2 can be found from i_3 by applying Ohm's law at R_5.

$$v_2 = R_5 i_3 = R_5(-0.16 v_1) = 0.25(-0.16 v_1) = -0.04 v_1$$
$$G_v = \frac{v_2}{v_1} = \frac{-0.04 v_1}{v_1} = -0.04$$

Comments: The Matlab commands required to solve this problem are listed below.

```
v = [1; 0; 0];
R = [1.5 -0.5 0; 1.5 -2 0.25; -1 1.25 -0.5];
i = R\v;
G = i(3)*0.25
```

Notice that v_1 was set to 1 in the Matlab computation so that $G = v_2/v_1 = v_2 = i_3 R_5$.

Remarks on Node Voltage and Mesh Current Methods

The methods of node and mesh analysis find applications beyond resistive circuits and should be viewed as general techniques for the analysis of any linear circuit. These methods provide systematic and effective means of obtaining the minimum number of equations necessary to solve a network problem. Since they are based on the fundamental laws of circuit analysis, KVL and KCL, they also apply to electric circuits containing non-linear circuit elements. You should master both methods as early as possible. Proficiency in these methods will enhance the learning process for more advanced concepts.

However, proficiency in these methods is not enough to understand and master circuit behavior. Except in the simplest examples, these methods do not produce solutions amenable to interpretation, generalization, and abstraction. They do provide an excellent means of generating useful numerical data. As such, they lend more insight when used to generate data, for a range of parameter (e.g., a resistor) values, that can be plotted and interpreted.

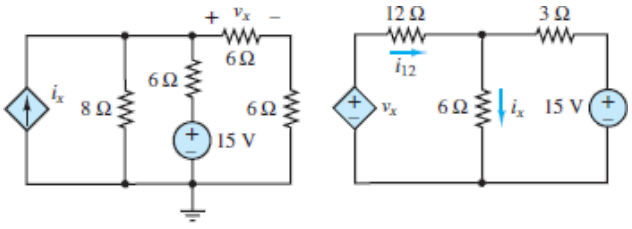
CHECK YOUR UNDERSTANDING

Solve Example 1.20 when $v_1 = 2i_S$.

Answer: $v_2 = \frac{21}{11}$ V; $v_3 = \frac{14}{11}$ V

CHECK YOUR UNDERSTANDING

For the figure on the left, assume $v_x = 3i_x$ and find the voltage v across the 8-Ω resistor by node analysis.



For the figure on the right assume $v_x = 2i_{12}$ and find the unknown current i_x using the mesh current method.

Answer: 12 V; 1.39 A

CHECK YOUR UNDERSTANDING

Determine the number of independent equations required to solve Example 1.21 using node analysis. Compare its efficiency in this problem to that of mesh analysis.

Conclusion

This chapter introduced the fundamentals students need in later chapters in the book to successfully analyze electric circuits. Upon successful completion of this chapter, a student will have learned to:

1. Identify the principal *features of electric circuits or networks*: nodes, loops, meshes, and branches. *Section 1.1*.
2. Apply definitions of charge, current and voltage. *Section 1.2*.
3. Identify *sources* and their *i-v characteristics*. *Section 1.3*.
4. Apply the *passive sign convention* to compute the power consumed or supplied by circuit elements. *Section 1.4*.
5. Apply *Kirchhoff's laws* to simple electric circuits. *Section 1.5*.
6. Apply *Ohm's law* to calculate unknown voltages and currents in simple circuits. *Section 1.6*.
7. Apply the *Node Voltage method* to solve for unknown voltages and currents in resistive networks. *Section 1.7*.
8. Apply the *Mesh Current method* to solve for unknown voltages and currents in resistive networks. *Section 1.8*.
9. Apply the Node Voltage and Mesh Current methods to solve for unknown voltages and currents in resistive networks with *dependent sources*. *Section 1.9*.

HOMEWORK PROBLEMS

Section 1.2: Charge, Current and Voltage

1.1 A free electron has an initial potential energy per unit charge (*voltage*) of 17 kJ/C and a velocity of 93 Mm/s. Later, its potential energy per unit charge is 6

kJ/C. Determine the change in velocity of the electron.

1.2 The units for voltage, current, and resistance are the volt (V), the ampere (A), and the ohm (Ω), respectively. Express each unit in fundamental MKS units.

1.3 A particular fully charged battery can deliver $2.7 \cdot 10^6$ coulombs of charge.

a. What is the capacity of the battery in ampere-hours?

b. How many electrons can be delivered?

1.4 The charge cycle shown in Figure P1.4 is an example of a three-rate charge. The current is held constant at 30 mA for 6 h. Then it is switched to 20 mA for the next 3 h. Find:

a. The total charge transferred to the battery.

b. The energy transferred to the battery.

Hint: Energy w is the integral of power, or $P = dw/dt$.

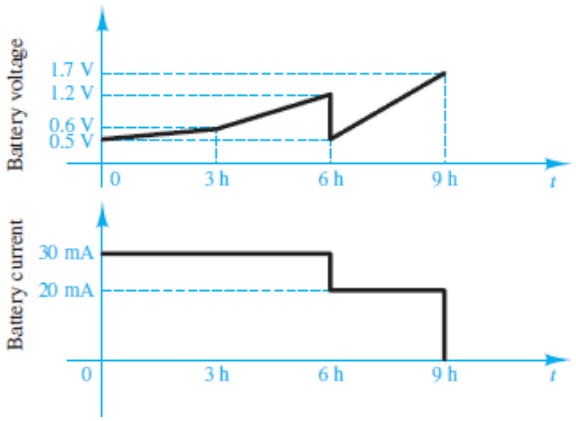


Figure P1.4

1.5 Batteries (e.g., lead-acid batteries) store and release chemical potential energy. Batteries do not store electric charge. During discharge, electrons exit the *cathode* terminal and reenter the battery at the *anode* terminal having done work on some external device (e.g., a lightbulb). The chemical energy stored in the battery is used to replenish the potential energy of those same electrons. It is convenient to think of positive carriers flowing in the opposite direction, that is, conventional current, and exiting at a higher voltage. (Benjamin Franklin caused this mess!) For a battery rated at 12 V and 350 A-h, determine:

a. The rated chemical energy stored in the battery.

b. The total charge that can be supplied at the rated voltage.

1.6 What determines:

a. The current *through* an ideal voltage source?

b. The voltage *across* an ideal current source?

1.7 An automotive battery is rated at 120 A-h. This means that under certain test conditions it can output 1 A at 12 V for 120 h (under other test conditions, the battery may have other ratings).

a. How much total energy is stored in the battery?

b. If the headlights are left on overnight (8 h), how much energy will still be stored in the battery in the morning? (Assume a 150-W total power rating for both headlights together.)

1.8 A car battery kept in storage in the basement needs recharging. Assume the voltage and the current provided by the charger during a charge cycle are shown in Figure P1.8.

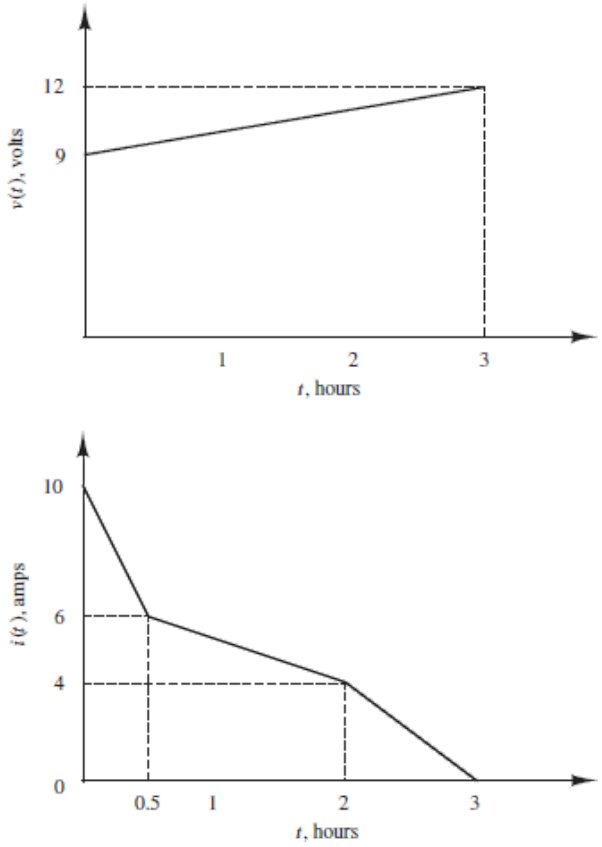


Figure P1.8

a. Find the total charge transferred to the battery.

b. Find the total energy transferred to the battery.

1.9 Suppose the current through a wire is given by the curve shown in Figure P1.9.

a. Find the amount of charge q that flows through the wire between $t_1 = 0$ and $t_2 = 1$ s.

b. Repeat part a for $t_2 = 2, 3, 4, 5, 6, 7, 8, 9$, and 10 s.

c. Sketch $q(t)$ for $0 \leq t \leq 10$s.

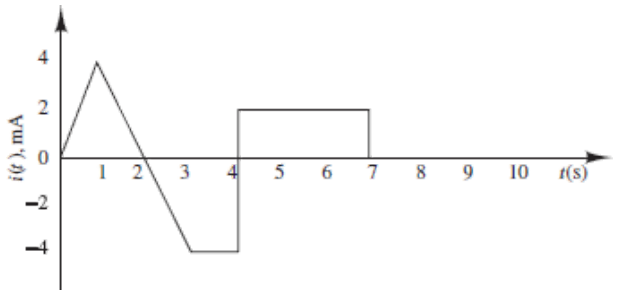


Figure P1.9

1.10 The charge cycle shown in Figure P1.10 is an example of a two-rate charge. The current is held constant at 70 mA for 1 h. Then it is switched to 60 mA for the next 1 h. Find:

a. The total charge transferred to the battery.

b. The total energy transferred to the battery.

Hint: Energy w is the integral of power, or $P = dw/dt$. Let

$$v_1(t) = 5 + e^{t/5194.8}\ \text{V}$$

$$v_2(t) = \left(6 - \frac{4}{e^{1h} - 1}\right) + \frac{4}{e^{2h} - e^{1h}} \cdot e^{t}\ \text{V}$$

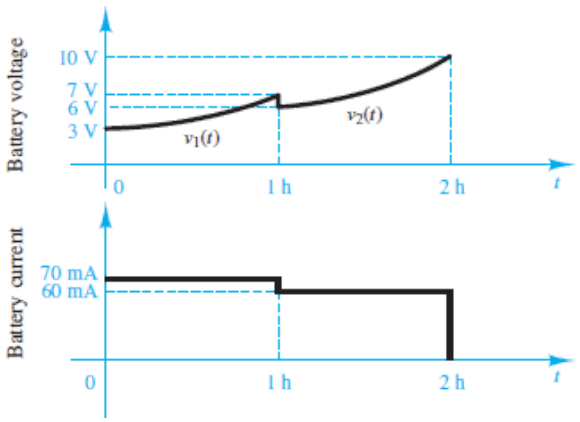


Figure P1.10

Page 64

1.11 The charging scheme used in Figure P1.11 is an example of a constant-current charge cycle. The charger voltage is controlled such that the current into the battery is held constant at 40 mA, as shown in Figure P1.11. The battery is charged for 6 h. Find:

a. The total charge delivered to the battery.

b. The energy transferred to the battery during the charging cycle.

Hint: Recall that the energy w is the integral of power, or $P = dw/dt$.

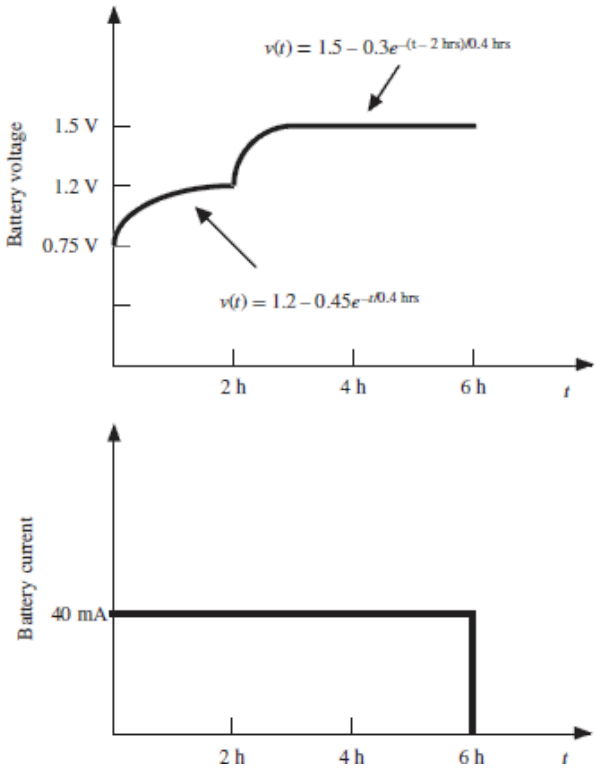


Figure P1.11

1.12 The charging scheme used in Figure P1.12 is called a *tapered-current charge cycle*. The current starts at the highest level and then decreases with time for the entire charge cycle, as shown. The battery is charged for 12 h. Find:

a. The total charge delivered to the battery.

b. The energy transferred to the battery during the charging cycle.

Hint: Recall that the energy w is the integral of power, or $P = dw/dt$.

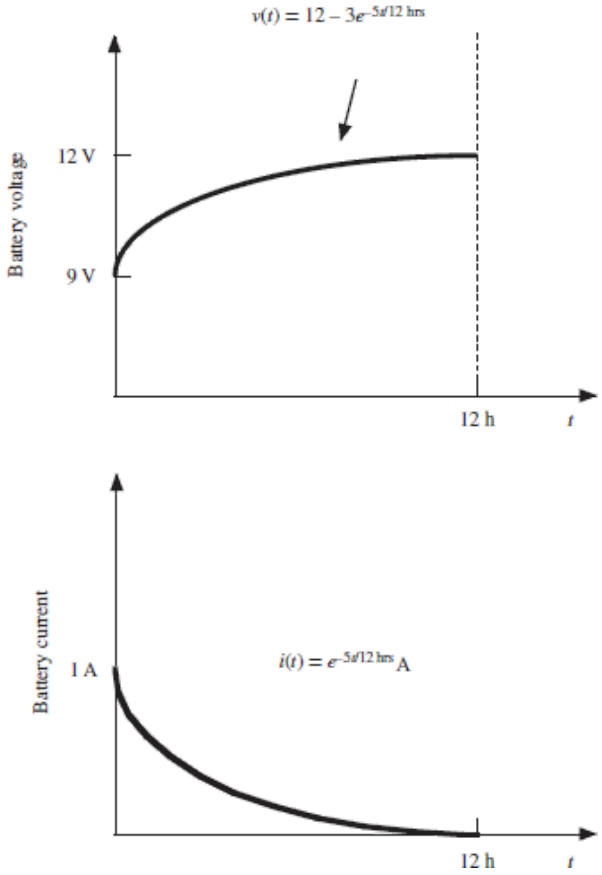


Figure P1.12

Section 1.4: Power and the Passive Sign Convention

1.13 Find the power delivered by the source in Figure P1.13.

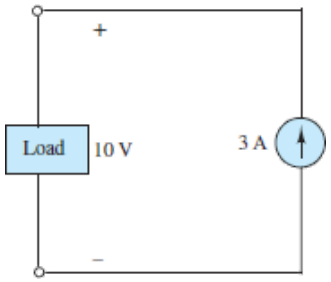


Figure P1.13

1.14 Find the power delivered by the source in Figure P1.14.

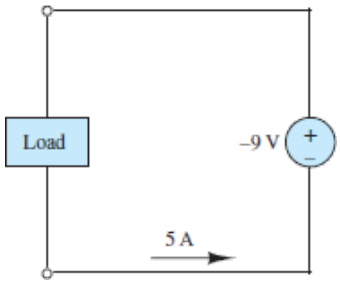


Figure P1.14

1.15 Determine whether each element in Figure P1.15 is supplying or dissipating power, and how much.

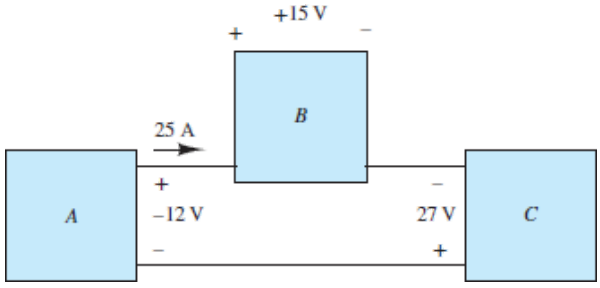


Figure P1.15

1.16 In the circuit of Figure P1.16, determine the power absorbed by the resistor R_4 and the power delivered by the current source.

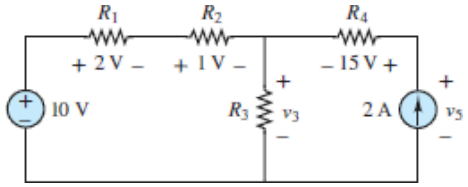


Figure P1.16

1.17 For the circuit shown in Figure P1.17:

a. Determine whether each component is absorbing or delivering power.

b. Is conservation of power satisfied? Explain your answer.

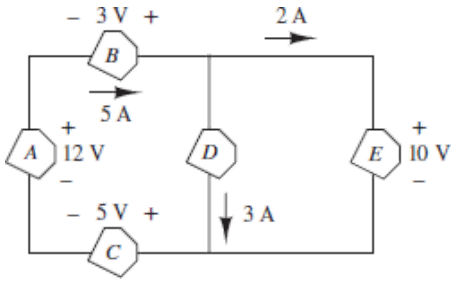


Figure P1.17

1.18 For the circuit shown in Figure P1.18, determine which components are supplying power and which are dissipating power. Also determine the amount of power dissipated and supplied.

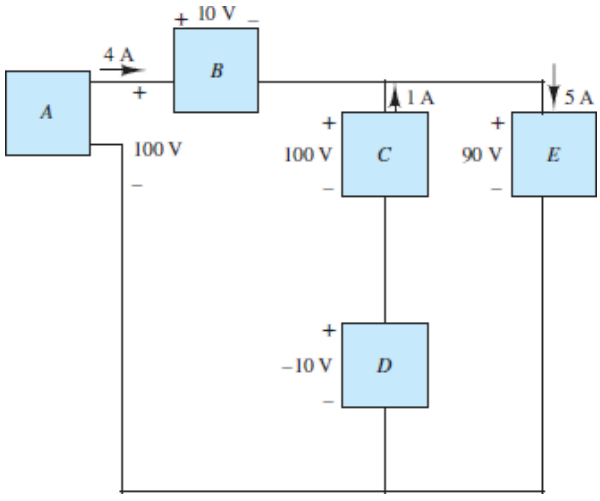


Figure P1.18

1.19 For the circuit shown in Figure P1.19, determine which components are supplying power and which are dissipating power. Also determine the amount of power dissipated and supplied.

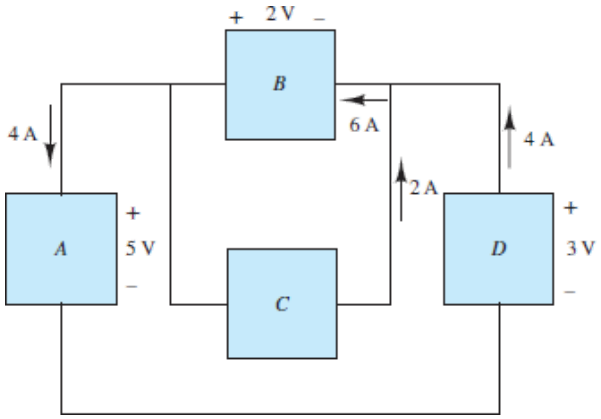


Figure P1.19

1.20 If an electric heater requires 23 A at 110 V, determine

a. The power it dissipates as heat or other losses.

b. The energy dissipated by the heater in a 24-h period.

c. The cost of the energy if the power company charges at the rate 6 cents/kWh.

Page 66

Section 1.5: Kirchhoff's Laws

1.21 For the circuit shown in Figure P1.21, determine the power absorbed by the 5-Ω resistor. *Hint:* The voltage across the 5-Ω resistor is 5 V.

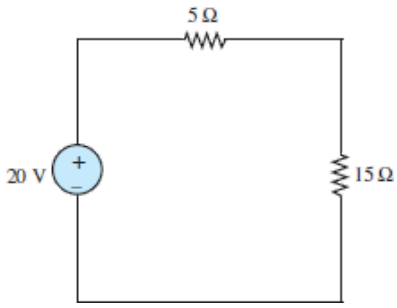


Figure P1.21

1.22 Use KCL to determine the unknown currents in the circuit of Figure P1.22. Assume $i_0 = 2$ A and $i_2 = -7$ A.

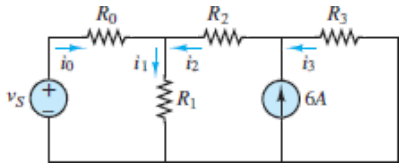


Figure P1.22

1.23 Use KCL to find the currents i_1 and i_2 in Figure P1.23. Assume that $i_a = 3$ A, $i_b = -2$ A, $i_c = 1$ A, $i_d = 6$ A., and $i_e = -4$ A..

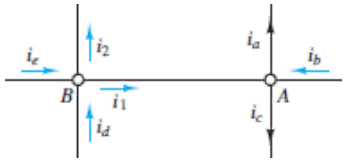


Figure P1.23

1.24 Use KCL to find the currents i_1, i_2, and i_3 in the circuit of Figure P1.24. Assume that $i_a = 2$ mA, $i_b = {}_7$ mA, and $i_c = {}_4$ mA.

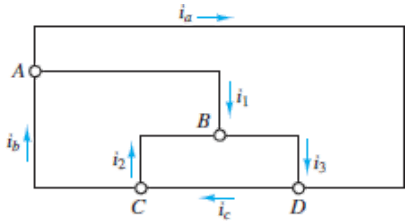


Figure P1.24

1.25 Use KVL to find the voltages v_1, v_2, and v_3 in Figure P1.25. Assume that $v_a = {}_2$ V, $v_b = {}_4$ V, and $v_c = {}_5$ V.

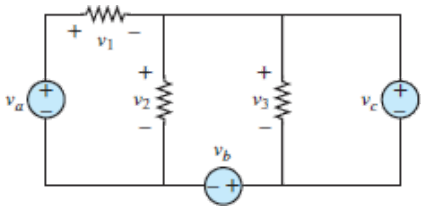


Figure P1.25

1.26 Use KCL to determine the currents i_1, i_2, i_3, and i_4 in the circuit of Figure P1.26. Assume that $i_a = -2$ A, $i_b = 6$ A, $i_c = 1$ A, and $i_d = -$A.

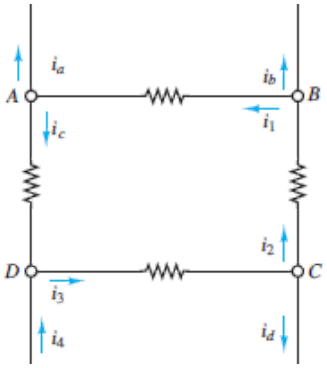


Figure P1.26

Section 1.6: Resistance and Ohm's Law

1.27 In the circuit shown in Figure P1.27, determine the terminal voltage v_T of the source, the power absorbed by R_o, and the efficiency of the circuit. Efficiency is defined as the ratio of load power to source power.

$$v_S = 12\,\text{V} \qquad R_S = 5\,\text{k}\Omega \qquad R_o = 7\,\text{k}\Omega$$

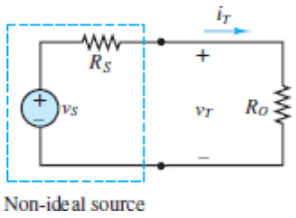


Figure P1.27

1.28 A 24-V automotive battery is connected to two headlights that are in parallel, similar to that shown in Figure 1.12. Each headlight is intended to be a 75-W Page 67load; however, one 100-W headlight is mistakenly installed. What is the resistance of each headlight? What is the total current supplied by the battery?

1.29 What is the total current supplied by the battery of Problem 1.28 if two 15-W tail lights are added (in parallel) to two 75-W headlights?

1.30 For the circuit shown in Figure P1.30, determine the power absorbed by the variable resistor R, ranging from 0 to 30 Ω. Plot the power absorption as a function of R. Assume that v_S = 15V, R_S = 10Ω. (*Hint:* Apply KVL around the circuit, using Ohm's law to express the voltage drop across each resistance.)

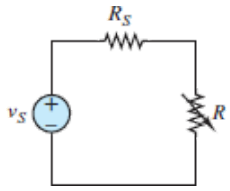


Figure P1.30

1.31 Refer to Figure P1.27 and assume that $v_S = 15$ V and $R_S = 100\Omega$. For $i_T = 0$, 10, 20, 30, 80, and 100 mA:

a. Find the total power supplied by the ideal source.

b. Find the power dissipated within the non-ideal source.

c. How much power is supplied to the load resistor?

d. Plot the terminal voltage v_T and power supplied to the load resistor as a function of terminal current i_T.

1.32 In the circuit of Figure P1.32, assume $v_2 = v_S/6$ and the power delivered by the source is 150 mW. Also assume that $R_1 = 8$ kΩ, $R_2 = 10$ kΩ, $R_3 = 12$ kΩ. Find R, v_S, v_2, and i. (*Hint:* Apply KVL around the circuit, using Ohm's law to express the voltage drop across each resistance.)

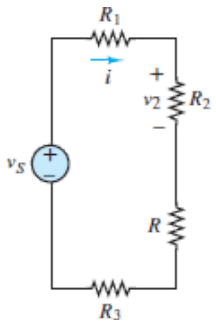


Figure P1.32

1.33 A GE SoftWhite Longlife lightbulb is rated as follows:

P_R = rated power = 60 W

P_{OR} = rated optical power = 820 lumens (lm) (average)

$$1 \text{ lumen} = \frac{1}{680} \text{ W}$$

Operating life = 1,500 h (average)

V_R = rated operating voltage = 115 V

The resistance of the filament of the bulb at room temperature, measured with a standard multimeter, is 16.7 Ω. When the bulb is operating at the rated values, determine:

a. The resistance of the filament.

b. The efficiency of the bulb.

1.34 An incandescent lightbulb rated at 100 W is designed to dissipate 100 W as heat and light when connected across a 110-V ideal voltage source. Determine the resistance of the lightbulb when operated as designed.

1.35 An incandescent lightbulb rated at 60 W is designed to dissipate 60 W as heat and light when connected across a 110-V ideal voltage source. A 100-W bulb is designed to dissipate 100 W when connected across the same source. Determine the resistance of each lightbulb when operated as designed.

1.36 Refer to Figure P1.36, and assume that v_S = 12 V, R_1 = 5Ω, R_2 = 3Ω, R_3 = 4Ω, and R_4 = 5Ω. Apply KVL and Ohm's law as often as necessary to find:

a. The voltage v_{ab}.

b. The power dissipated in R_2.

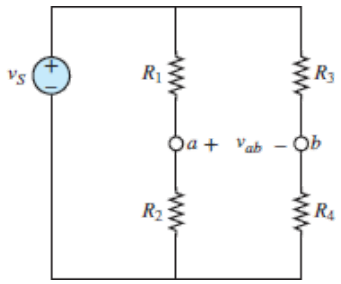


Figure P1.36

Page 68

1.37 Refer to Figure P1.37, and assume that v_S = 7 V, i_S = 3 A, R_1 = 20Ω, R_2 = 12Ω, and R_3 = 10Ω. Apply Kirchhoff's laws and Ohm's law to find:

a. The currents i_1 and i_2.

b. The power supplied by the source v_S.

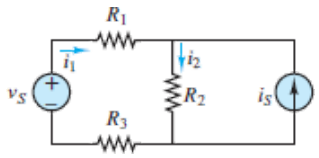


Figure P1.37

1.38 Refer to Figure P1.38, and assume $v_1 = 15$ V, $v_2 = 6$ V, $R_1 = 18\Omega$, $R_2 = 10\Omega$. Apply Kirchhoff's laws and Ohm's law to find:

a. The currents i_1, i_2.

b. The power delivered by the sources v_1 and v_2.

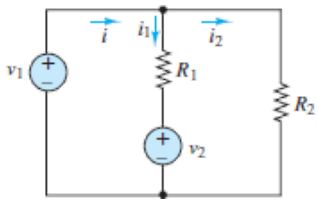


Figure P1.38

1.39 Consider NiMH hobbyist batteries depicted in Figure P1.39.

a. If $V_1 = 12.0$ V, $R_1 = 0.15\Omega$, and $R_o = 2.55\Omega$, find the current I_o and the power dissipated by R_o.

b. If battery 2 with $V_2 = 12.0$ V and $R_2 = 0.28\Omega$ is placed in parallel with battery 1, will the current I_o increase or decrease? Will the power dissipated by R_o increase or decrease? By how much?

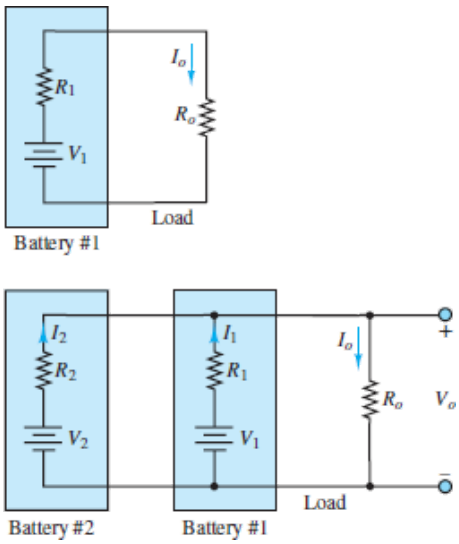


Figure P1.39

1.40 The open-circuit voltage across the terminals of a particular power supply is 50.8 V. When a 10-W lightbulb is attached, the voltage drops to 49 V. This result is explained by modeling the power supply as a non-ideal source, such as that shown in Figure P1.27.

a. Determine v_s and R_s for this non-ideal source.

b. What voltage would be measured across the terminals when a 15-Ω resistor is attached?

c. How much current could be drawn from this power supply under short-circuit conditions?

1.41 A 220-V electric heater has two heating coils that can be switched such that either coil can be used independently or the two can be connected in series or parallel, for a total of four possible configurations. If the warmest setting corresponds to 2,000-W power dissipation and the coolest corresponds to 300 W, determine the resistance of each coil.

1.42 For the circuits of Figure P1.42, determine the resistor values (including the power rating) necessary to achieve the indicated voltages. Resistors are available in ⅛-, ¼-, ½-, and 1-W ratings.

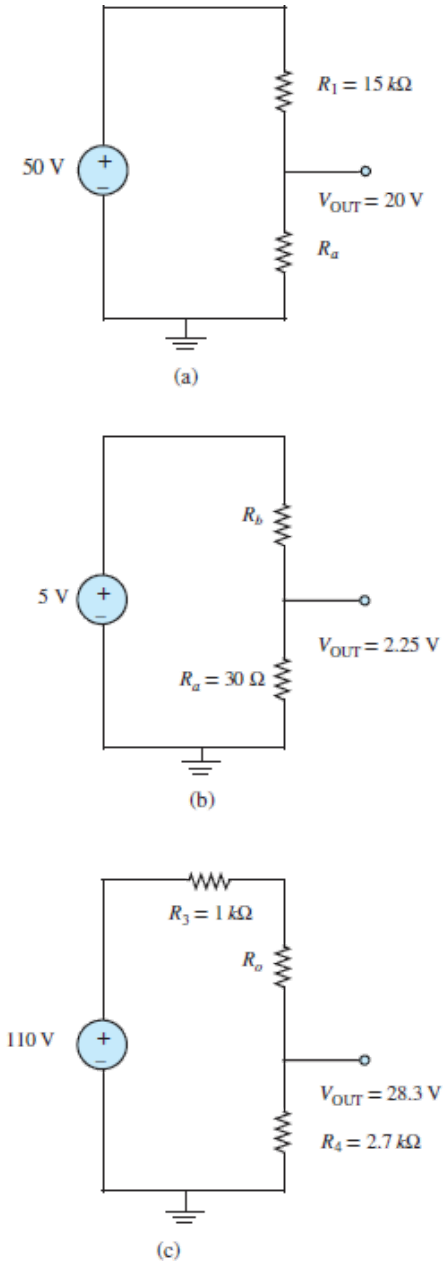


Figure P1.42

1.43 At an engineering site, a 1-hp motor is placed a distance d from a portable generator, as depicted in Figure P1.43. The generator can be modeled as an ideal DC source V_G = 110V. The nameplate on the Page 69motor gives the following rated voltages and full-load currents:

$$V_{M\,\min} = 105\ \text{V} \rightarrow I_{M\,\text{FL}} = 7.10\ \text{A}$$
$$V_{M\,\max} = 117\ \text{V} \rightarrow I_{M\,\text{FL}} = 6.37\ \text{A}$$

If d = 150 m and the motor must deliver its full-rated power, determine the minimum AWG conductors that must be used in a rubber-insulated cable. Assume that losses occur only in the wires.

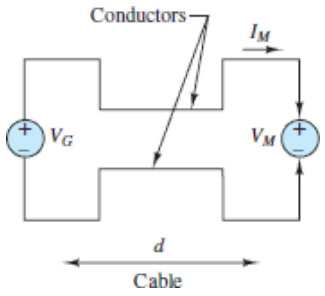


Figure P1.43

1.44 Cheap resistors are fabricated by depositing a thin layer of carbon onto a non-conducting cylindrical substrate (see Figure P1.44). If such a cylinder has radius a and length d, determine the thickness of the film required for a resistance R if:

$$a = 1\,\text{mm} \qquad R = 33\,\text{k}\Omega$$
$$\sigma = \frac{1}{\rho} = 2.9\,\text{M}\frac{\text{S}}{\text{m}} \qquad d = 9\,\text{mm}$$

Neglect the end surfaces of the cylinder and assume that the thickness is much smaller than the radius.

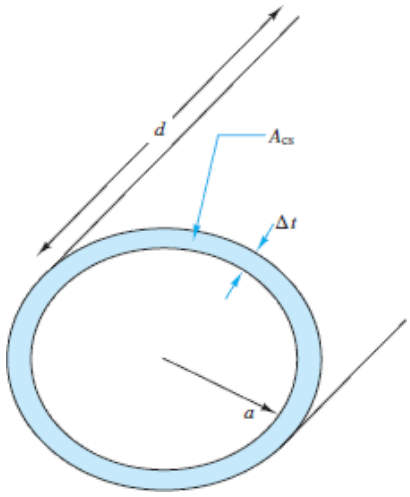


Figure P1.44

1.45 The resistive elements of fuses, lightbulbs, heaters, etc., are non-linear (i.e., the resistance is dependent on the current through the element). Assume the resistance of a fuse (Figure P1.45) is given by Page 70 $R = R_0[1 + A(T - T_0)]$

where: $T - T_0 = kP$; $T_0 = 25°C$; $A = 0.7[°C]^{-1}$; $k = 0.35°C/W$; $R_0 = 0.11\Omega$; and P is the power dissipated in the resistive element of the fuse. Determine the rated current at which the fuse will melt (that is, "blow") and thus act as an open-circuit. (*Hint:* The fuse blows when R becomes infinite.)

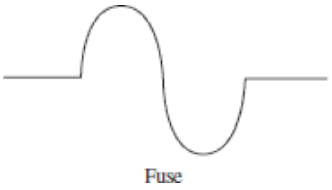


Figure P1.45

1.46 Refer to Figure P1.22. Assume $R_0 = 1\Omega$, $R_1 = 2\Omega$, $R_2 = 3\Omega$, $R_3 = 4\Omega$, and $v_S = 10$ V. Use KCL and Ohm's law to find the unknown currents.

1.47 Refer to Figure P1.47 and assume $R_0 = 1\,\Omega, R_1 = 2\,\Omega, R_2 = 3\,\Omega, R_3 = 4\,\Omega$, and $V_S = 12\ V$. Use KVL and Ohm's law to find:

a. The mesh currents i_a, i_b, and i_c. (Mesh currents circulate within meshes and are an alternative to branch currents. For example, the branch current through R_1 is the difference between i_a and i_b.)

b. The current through each resistor.

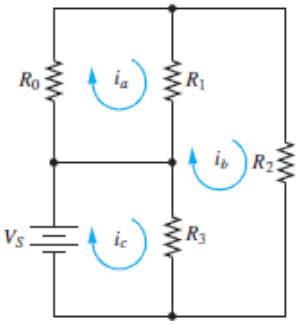


Figure P1.47

1.48 Refer to Figure P1.47 and assume $R_0 = 2\Omega$, $R_1 = 2\Omega$, $R_2 = 5\Omega$, $R_3 = 4\Omega$, and $VS = 24$ V. Use KVL and Ohm's law to find:

a. The mesh currents i_a, i_b, and i_c.

b. The voltage across each resistor.

1.49 Assume that the voltage source in Figure P1.47 is now replaced by a DC current source I_S, and $R_0 = 1\Omega$, $R_1 = 3\Omega$, $R_2 = 2\Omega$, $R_3 = 4\Omega$, and $IS = 12$ A, directed

positively upward. Use KVL and Ohm's law to determine the voltage across each resistor.

1.50 The voltage divider network of Figure P1.50 is designed to provide $v_{out} = v_S/2$. However, in practice, the resistors may not be perfectly matched; that is, their tolerances are such that the resistances are unlikely to be identical. Assume $v_S =$ 10 V and nominal resistance values of $R_1 = R_2 = 5$ k Ω.

a. If the resistors have ±10 percent tolerance, find the expected range of possible output voltages.

b. Find the expected output voltage range for a tolerance of ±5 percent.

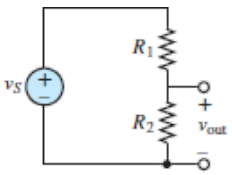


Figure P1.50

Section 1.7: The Node Voltage Method

1.51 Use the node voltage method to find the voltages V_1 and V_2 for the circuit of Figure P1.51.

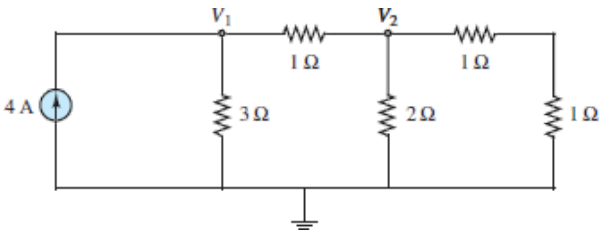


Figure P1.51

1.52 Using the node voltage method find the voltages V_1 and V_2 for the circuit of Figure P1.52.

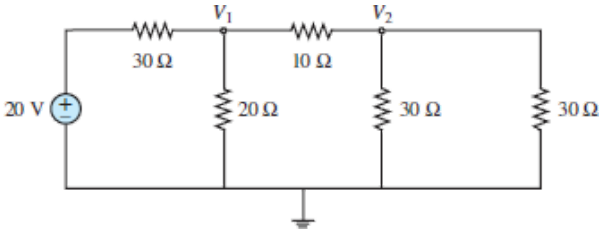


Figure P1.52

Page 71

1.53 Using the node voltage method in the circuit of Figure P1.53, find the voltage v across the 0.25-Ω resistance.

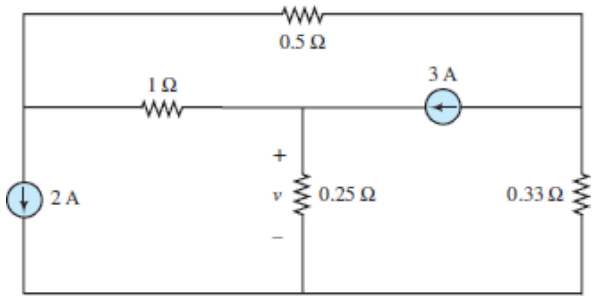


Figure P1.53

1.54 Using the node voltage method in the circuit of Figure P1.54, find the current i through the voltage source.

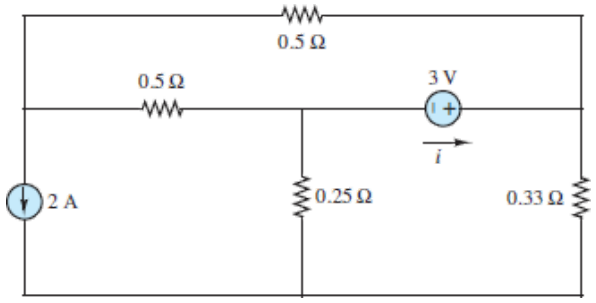


Figure P1.54

1.55 Use the node voltage method in the circuit of Figure P1.55 to find V_a. Let R_1 = 12Ω, R_2 = 6Ω, R_3 = 10Ω, V_1 = 4 V, V_2 = 1 V.

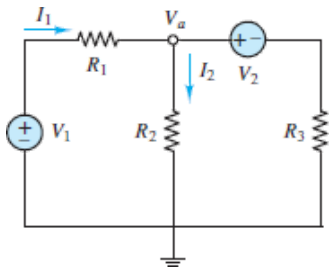


Figure P1.55

1.56 Use the node voltage method in the circuit of Figure P1.56 to find v_1, v_2, and v_3. Let $R_1 = 10\Omega$, $R_2 = 8\Omega$, $R_3 = 10\Omega$, $R_4 = 5\Omega$, $i_S = 2$ A, $v_S = 1$ V.

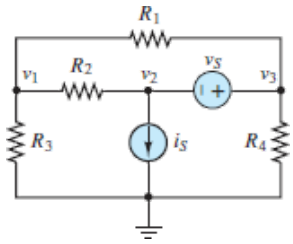


Figure P1.56

1.57 Use the node voltage method in the circuit of Figure P1.57 to find the voltages at nodes A, B, and C. Let $V_1 = 12$ V, $V_2 = 10$ V, $R_1 = 2\Omega$, $R_2 = 8\Omega$, $R_3 = 12\Omega$, $R_4 = 8\Omega$.

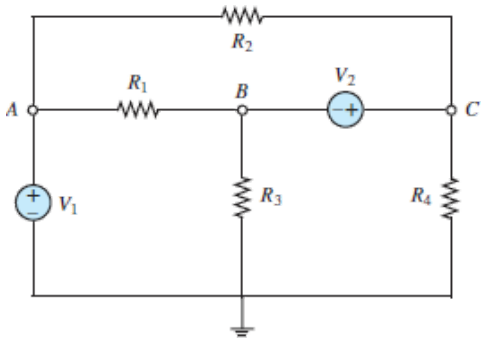


Figure P1.57

1.58 Use the node voltage method in the circuit of Figure P1.58 to find V_a and V_b. Let $R_1 = 10\Omega$, $R_2 = 4\Omega$, $R_3 = 6\Omega$, $R_4 = 6\Omega$, $V_1 = 2$ V, $V_2 = 4$ V, $i_1 = 2$ A.

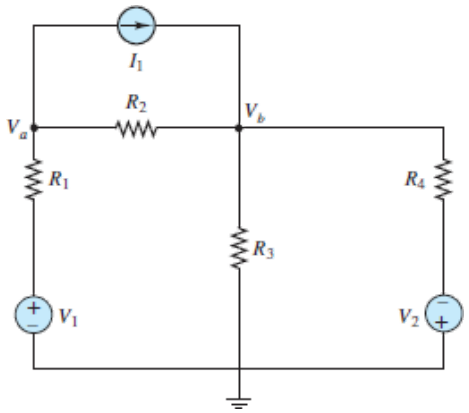


Figure P1.58

Page 72

1.59 Find the power delivered to the resistor R_0 for the circuit of Figure P1.59, using the node voltage method, given that $R_1 = 2\Omega$, $RV = R_2 = R_0 = 4\Omega$, $VS = 4$ V, and $I_S = 0.5$ A.

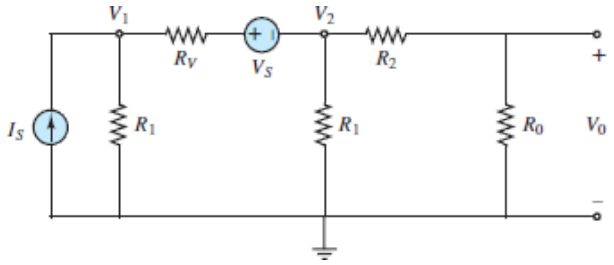


Figure P1.59

1.60 For the circuit of Figure P1.60, write the node equations necessary to find voltages V_1, V_2, and V_3. Note that $G = 1/R$ = conductance. From the results, note the interesting form that the matrices $[G]$ and $[I]$ have taken in the equation $[G][V] = [I]$ where

$$[G] = \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & \cdots & g_{2n} \\ g_{31} & & \ddots & & \\ \vdots & & & \ddots & \\ g_{n1} & g_{n2} & \cdots & \cdots & g_{nn} \end{bmatrix} \quad \text{and} \quad [I] = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ \vdots \\ I_n \end{bmatrix}$$

Write the matrix form of the node voltage equations again, using the following formulas:

$g_{ii} = \sum$ conductances connected to node i

$g_{ij} = -\sum$ conductances *shared* by nodes i and j

$I_i = \sum$ all *source* currents into node i

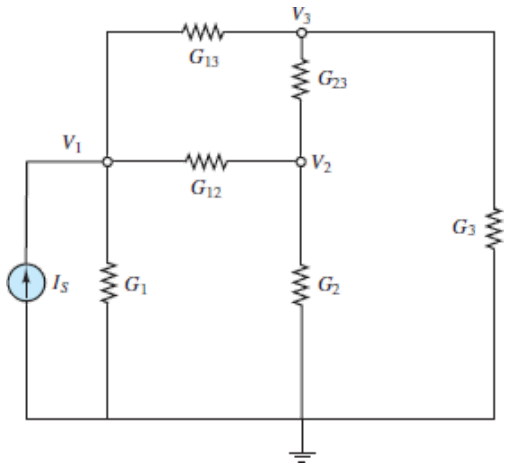


Figure P1.60

1.61 In the circuit in Figure P1.61, assume the source voltage and source current and all resistances are known.

a. Write the node equations required to determine the node voltages.

b. Write the matrix solution for each node voltage in terms of the known parameters.

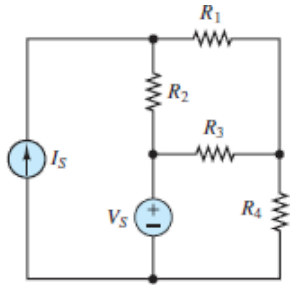


Figure P1.61

1.62 For the circuit of Figure P1.62 determine:

a. The voltage across R_1 using the node voltage method.

b. The voltage across R_3 using the node voltage method.

$V_{S1} = V_{S2} = 110\,\text{V}$
$R_1 = 500\,\text{m}\Omega \qquad R_2 = 167\,\text{m}\Omega$
$R_3 = 700\,\text{m}\Omega$
$R_4 = 200\,\text{m}\Omega \qquad R_5 = 333\,\text{m}\Omega$

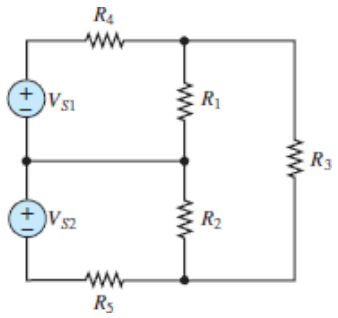


Figure P1.62

1.63 Figure P1.63 represents a temperature measurement system, where temperature T is linearly related to the voltage source $VS2$ by a transduction constant k. Use the node voltage method to determine the temperature.

$$V_{S2} = kT \qquad k = 10\,\text{V/°C}$$
$$V_{S1} = 24\,\text{V} \qquad R_S = R_1 = 12\,\text{k}\Omega$$
$$R_2 = 3\,\text{k}\Omega \qquad R_3 = 10\,\text{k}\Omega$$
$$R_4 = 24\,\text{k}\Omega \qquad V_{ab} = -2.524\,\text{V}$$

In practice, V_{ab} is used as the measure of temperature, which is introduced to the circuit through a temperature sensor modeled by the voltage source V_{S2} in series with R_S.

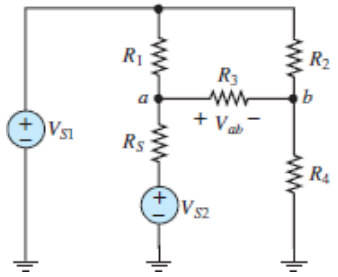


Figure P1.63

1.64 Use the node voltage method to find node voltages $V1$, $V2$, and V_3 in Figure P1.64. Let $R_1 = 10\Omega$, $R_2 = 6\Omega$, $R_3 = 7\Omega$, $R_4 = 4\Omega$, $i_1 = 2$ A, $i_2 = 1$ A.

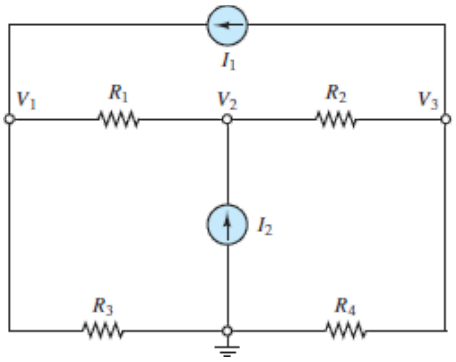


Figure P1.64

1.65 Use the node voltage method to find the current through R_4 in Figure P1.65. Let $R_1 = 10\Omega$, $R_2 = 6\Omega$, $R_3 = 4\Omega$, $R_4 = 3\Omega$, $R_5 = 2\Omega$, $R_6 = 2\Omega$, $i_1 = 2$ A, $i_2 = 3$ A, $i_3 = 5$ A.

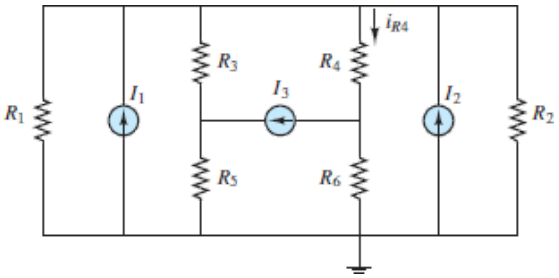


Figure P1.65

1.66 The circuit shown in Figure P1.66 is a simplified DC version of an AC three-phase wye-wye (Y-Y) electrical distribution system commonly used to supply industrial loads, particularly rotating machines.

$$V_{S1} = V_{S2} = V_{S3} = 170\text{ V}$$
$$R_{w1} = R_{w2} = R_{w3} = 0.7\,\Omega$$
$$R_1 = 1.9\,\Omega \qquad R_2 = 2.3\,\Omega$$
$$R_3 = 11\,\Omega$$

a. Determine the number of non-reference nodes.

b. Determine the number of unknown node voltages.

c. Compute v_1', v_2', v_3', and v_n'.

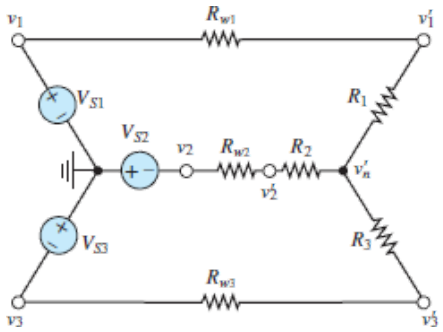


Figure P1.66

1.67 Apply the node voltage method to the circuit of Figure P1.67 to find the three indicated node voltages and the current i. Assume: $R_1 = 10\Omega$, $R_2 = 20\Omega$, $R_3 = 20\Omega$, $R_4 = 10\Omega$, $R_5 = 10\Omega$, $R_6 = 10\Omega$, $R_7 = 5\Omega$, $V_1 = 20$ V, $V_2 = 20$ V.

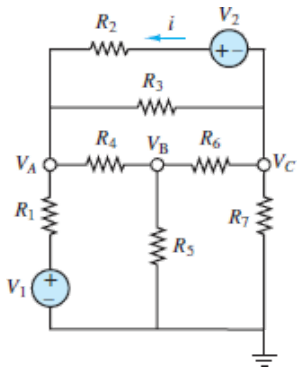


Figure P1.67

Page 74

Section 1.8: The Mesh Current Method

1.68 In the circuit shown in Figure P1.68, the mesh currents are

$$I_1 = 5\,\text{A} \quad I_2 = 3\,\text{A} \quad I_3 = 7\,\text{A}$$

Determine the branch currents through:

a. R_1. b. R_2. c. R_3.

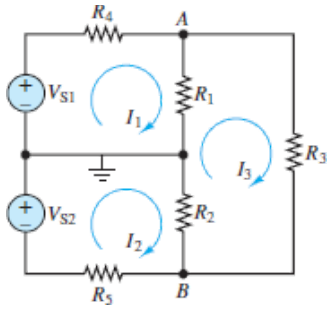


Figure P1.68

1.69 In the circuit shown in Figure P1.68, the source and node voltages are

$$V_{S1} = V_{S2} = 110\ \text{V}$$
$$V_A = 103\ \text{V} \quad V_B = -107\ \text{V}$$

Use the mesh current method to determine the voltage across each of the five resistors.

1.70 Use the mesh current method in the circuit of Figure P1.55 to find V_a. Let $R_1 = 12\Omega$, $R_2 = 6\Omega$, $R_3 = 10\Omega$, $V_1 = 4$ V, $V_2 = 1$ V.

1.71 Using the mesh current method, find the currents i_1 and i_2 for the circuit of Figure P1.71.

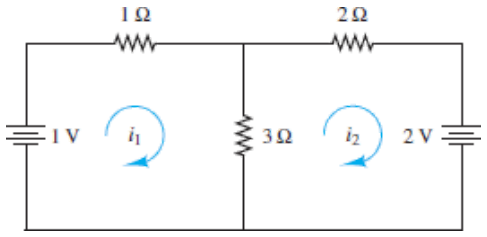


Figure P1.71

1.72 Using the mesh current method, find the currents i_1 and i_2 and the magnitude and polarity of the voltage across the upper 10-Ω resistor in the circuit of Figure P1.72.

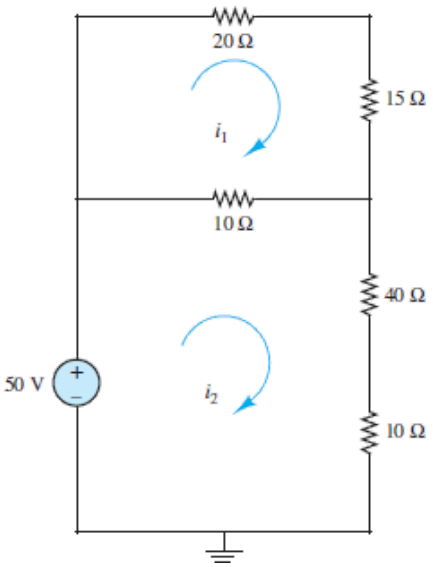


Figure P1.72

1.73 Using the mesh current method, find the voltage v across the 3-Ω resistor in the circuit of Figure P1.73.

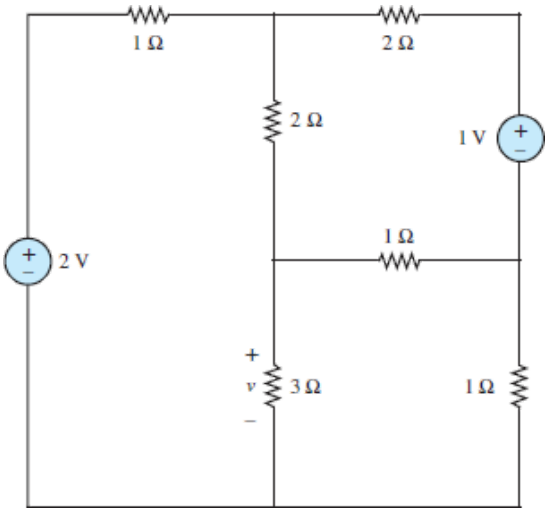


Figure P1.73

Page 75

1.74 Using the mesh current method, find the currents i_1, i_2, and i_3 in the circuit of Figure P1.74.

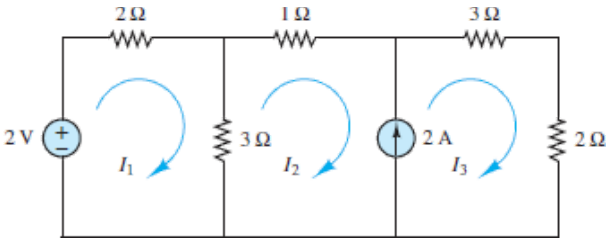


Figure P1.74

1.75 For the circuit of Figure P1.75, write the mesh equations in matrix form. Notice the form of the $[R]$ and $[V]$ matrices in $[R][I] = [V]$, where

$$[R] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & \cdots & r_{2n} \\ r_{31} & & \ddots & & \\ \vdots & & & \ddots & \\ r_{n1} & r_{n2} & \cdots & \cdots & r_{nn} \end{bmatrix} \quad \text{and} \quad [V] = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ \vdots \\ V_n \end{bmatrix}$$

Write the matrix form of the mesh equations again by using the following formulas:

$$r_{ii} = \sum \text{resistances around loop } i$$
$$r_{ij} = -\sum \text{resistances } \textit{shared} \text{ by loops } i \text{ and } j$$
$$V_i = \sum \text{source voltages around loop } i$$

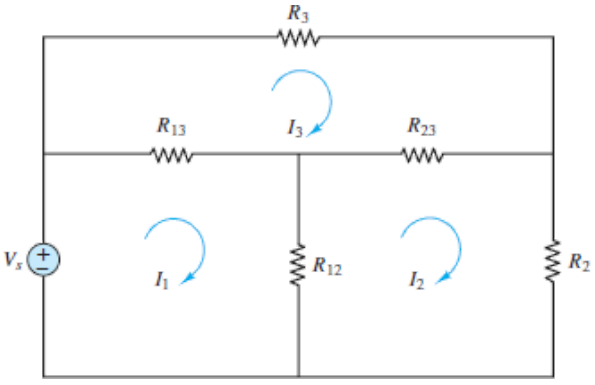


Figure P1.75

1.76 For the circuit of Figure P1.76, use the mesh current method to find four equations in the four mesh currents. Collect coefficients and solve for the mesh currents.

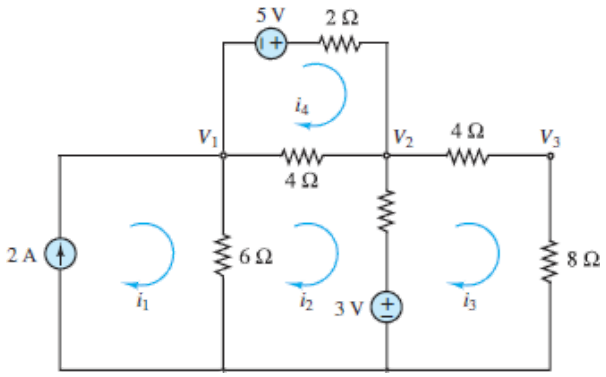


Figure P1.76

1.77 Use the mesh current method to find the mesh currents in Figure P1.77. Let R_1 = 10Ω, R_2 = 5Ω, V_1 = 2 V, V_2 = 1 V, I_s = 2 A.

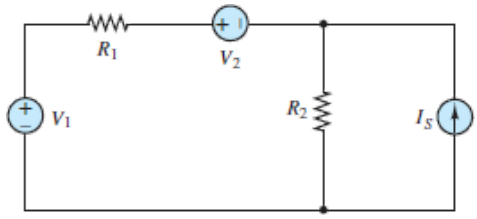


Figure P1.77

1.78 Use the mesh current method to find the mesh currents in Figure P1.78. Let R_1 = 6Ω, R_2 = 3Ω, R_3 = 3Ω V_1 = 4 V, V_2 = 1 V, V_3 = 2 V.

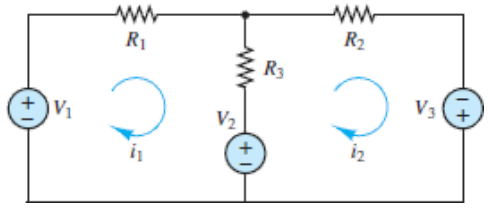


Figure P1.78

Page 76

1.79 Use the mesh current method to find V_4 in Figure P1.79. Let R_2 = 6Ω, R_3 = 3Ω, R_4 = 3Ω, R_5 = 3Ω, vS = 4 V, iS = 2 A.

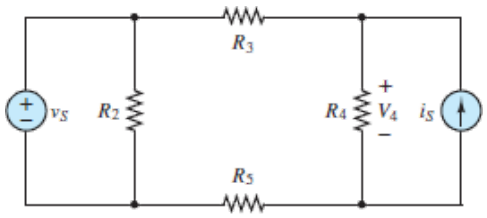


Figure P1.79

1.80 Use the mesh current method to find mesh currents in Figure P1.80. Let R_1 = 8Ω, R_2 = 3Ω, R_3 = 5Ω, R_4 = 2Ω, R_5 = 4Ω, R_6 = 3Ω, V_1 = 4 V, V_2 = 2 V, V_3 = 1 V, V_4 = 2 V, V_5 = 3 V, V_6 = 2 V.

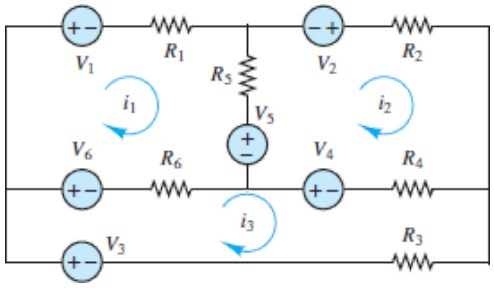


Figure P1.80

1.81 Use the mesh current method to find the current i in Figure P1.81. Assume i_S = 2 A.

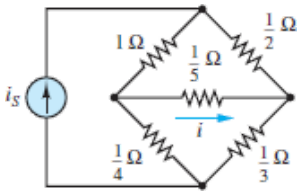


Figure P1.81

1.82 Use the mesh current method to find the currents through every branch in Figure P1.82. Let R_1 = 10Ω, R_2 = 5Ω, R_3 = 4Ω, R_4 = 1Ω, V_1 = 5 V, V_2 = 2 V.

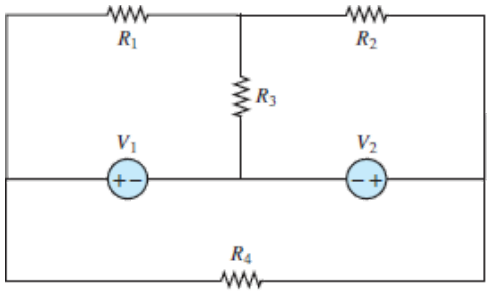


Figure P1.82

1.83 Using the data of Problem 1.66 and Figure P1.66,

a. Determine the number of meshes.

b. Compute the mesh currents.

c. Use the mesh currents to determine v_{π}'.

Section 1.9: The Node Voltage and Mesh Current Methods with Dependent Sources

1.84 Use the node voltage method on the circuit in Figure P1.84 to determine the voltage V_4. Note that one source is a dependent (controlled) voltage source! Let $V_S = 5\text{V}$; $A_V = 70$; $R_1 = 2.2\text{k}\Omega$; $R_2 = 1.8\text{k}\Omega$; $R_3 = 6.8\text{k}\Omega$; $R_4 = 220\Omega$.

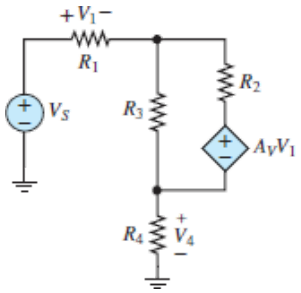


Figure P1.84

1.85 Use the mesh current method to find the current i in Figure P1.85. Let $v_S = 5.6$ V; $R_1 = 50\Omega$; $R_2 = 1.2\text{k}\Omega$; $R_3 = 330\Omega$; $g_m = 0.2$ S; $R_4 = 440\Omega$.

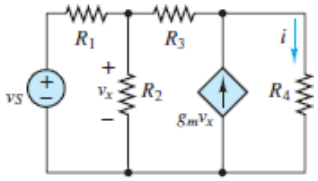


Figure P1.85

Page 77

1.86 Use the mesh current method to find the voltage gain $G_v = v_2/v_s$ in Figure P1.86.

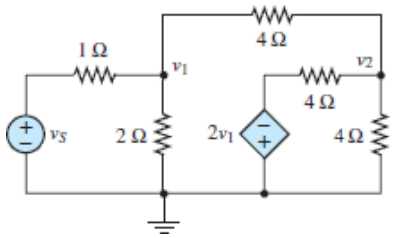


Figure P1.86

1.87 Apply the mesh current method to find the power supplied by the voltage source in Figure P1.87. Assume $k = 0.25$ A/A^2.

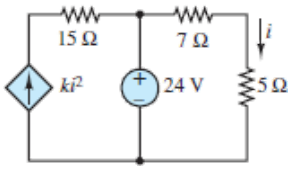


Figure P1.87

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]Gustav Robert Kirchhoff (1824–1887), a German scientist, published the first systematic description of the laws of circuit analysis. His contribution—though not original in terms of its scientific content—forms the basis of all circuit analysis.

Page 78

C H A P T E R

2

EQUIVALENT NETWORKS

Chapter 2 begins by introducing the critically important concepts of equivalent resistance, voltage division, and current division. A thorough understanding of these basic concepts is essential to the successful development of network analysis and design skills. Next, the *principle of superposition* is introduced to decompose a network into multiple, simpler perspectives. Both the principle of superposition and the concept of equivalent resistance represent network simplification techniques.

The chapter then expands upon the concept of equivalent resistance to introduce the more general concept of *equivalent one-port networks*. A *one-port network* is a network that can be accessed through two terminals. In the context of electrical networks, the term *equivalent* does not mean identical. Instead, two one-port networks are equivalent when the *i-v* characteristics of each pair of terminals are the same such that the impact of either network on a third one-port network is the same when connected. The implication is that any one-port network may be replaced by a simpler one-port network as long as the two networks are equivalent. Thus, it may be possible (as it often is) to simplify an electrical network by dividing it into two connected one-port networks and replacing one or both of them with simpler equivalent networks. This process is introduced generally as the *source-load perspective* followed by network simplification methods known as *source transformations* and *Thévenin* and *Norton* equivalent networks.

The chapter concludes with the concept of maximum power transfer and applications of equivalent networks to practical sources and meters as well as nonlinear elements.

Throughout this chapter students are encouraged to apply a simple two-step method to problem solving, with the first step being network simplification followed by the second step of solving the problem.

Learning Objectives

Students will learn to...

1. Apply *voltage and current division* to calculate unknown voltages and currents in simple series, parallel, and series-parallel circuits. *Sections 2.1–2.2*.
2. Correctly redraw a resistive network, as necessary, and compute the equivalent resistance between two nodes. *Section 2.3*.
3. Apply the *principle of superposition* to linear circuits containing independent and dependent sources. *Section 2.4*.
4. Apply the source-load perspective to find graphical solutions to circuit problems *Section 2.5*.

5. Apply source transformations to simplify and solve linear circuits containing independent and dependent sources. *Section 2.6*.

6. Determine *Thévenin and Norton equivalent circuits* for networks containing linear resistors and independent and dependent sources. *Section 2.7*.
7. Use equivalent-circuit ideas to compute the *maximum power transfer* between a source and a load. *Section 2.8*.
8. Understand the impact of internal resistance in practical models of voltage and current sources as well as of voltmeters, ammeters, and wattmeters. *Sections 2.9–2.10*.
9. Use the concept of equivalent circuits to determine voltage, current, and power for nonlinear loads by using *load-line analysis* and analytical methods. *Section 2.11*.

2.1 RESISTORS IN SERIES AND VOLTAGE DIVISION

It is common to find two or more circuit elements situated along a single current path or branch; that is, the elements are in *series*. When elements are in series, the voltage across the entire branch is divided among the elements in the branch. This important observation is known as **voltage division**.

The most fundamental instance of voltage division occurs when two resistors are in series, as shown in Figure 2.1. KVL applied around the series loop requires the voltage drop v_S across the source to be equal to the sum of the voltage drops v_1 and v_2 across the two resistors.

$$v_S = v_1 + v_2 \qquad \text{KVL}$$

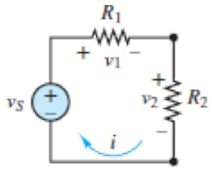


Figure 2.1 The current i flows through each of the three elements in the series loop. KVL requires $v_S = v_1 + v_2$.

Ohm's law can be applied to each resistor to find expressions for v_1 and v_2. (Notice the passive sign convention.)

$$v_1 = iR_1 \quad \text{and} \quad v_2 = iR_2$$

Page 81

Plug in for v_1 and v_2 to find:

$$v_S = iR_1 + iR_2 = i(R_1 + R_2) \equiv iR_{\text{EQ}}$$

This expression defines the *equivalent resistance* R_{EQ} of two resistors in series, where:

$$R_{\text{EQ}} = (R_1 + R_2) \quad \text{(two resistors in series)}$$

When three or more resistors are connected in series, the equivalent resistance is equal to the sum of all the resistances.

$$R_{\text{EQ}} = \sum_{n=1}^{N} R_n \quad \text{resistors in series} \tag{2.1}$$

Clearly, R_{EQ} is greater than any of the individual resistances in the series. It is often useful to replace a series of two or more resistances with a single equivalent resistance, as indicated in Figure 2.2. To do so correctly, remove the selected resistances in series along the branch and replace them with a single equivalent resistance along the same branch. This simple procedure illustrates a very important principle: From the perspective of whatever else is eventually attached to that branch (e.g., the voltage source in Figure 2.1), the resistances in series are *seen* as a single resistance R_{EQ}.

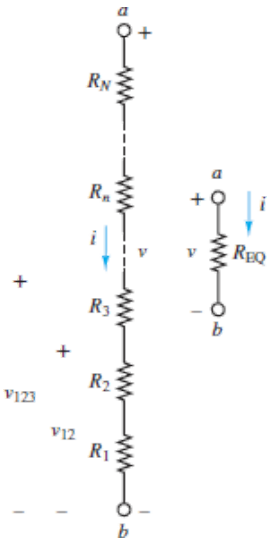


Figure 2.2 The equivalent resistance of three or more resistances in series equals the sum of those resistances.

An expression for how the voltage across the entire branch is divided among the individual resistances along that branch can be found by using Ohm's law and noting that the current is the same through each resistance. Consider the series loop in Figure 2.1:

$$i = \frac{v_1}{R_1} = \frac{v_2}{R_2} = \frac{v_S}{R_{EQ}}$$

which yields the following relationships:

$$\frac{v_1}{v_S} = \frac{R_1}{R_{EQ}} \quad \text{and} \quad \frac{v_2}{v_S} = \frac{R_2}{R_{EQ}} \quad \text{and} \quad \frac{v_1}{v_2} = \frac{R_1}{R_2} \tag{2.2}$$

These results, known as *voltage division,* indicate that *for resistors in series the ratio of voltages equals the ratio of the corresponding resistances.* The voltage drops v_1 and v_2 are fractions of the total voltage v_S because R_1 and R_2 are both less than R_{EQ}.

The ratio of the voltages across any two resistances in series equals the ratio of those resistances.

When series connections are encountered in circuit diagrams, one should immediately think of voltage division.

Series Connection ⇒ Voltage Division

It is important to realize that the voltage division rule applies to any two resistances in series, not just any two discrete resistors. For example, consider the series of Page 82resistors shown in Figure 2.2. The ratio of the voltage across $R_1 + R_2$ to the voltage across $R_1 + R_2 + R_3$ equals the ratio of $R_1 + R_2$ to $R_1 + R_2 + R_3$. That is:

$$\frac{v_{12}}{v_{123}} = \frac{R_1 + R_2}{R_1 + R_2 + R_3} \qquad \text{Voltage division}$$

EXAMPLE 2.1 Voltage Division

Problem

Determine the voltage v_3 in the circuit of Figure 2.3.

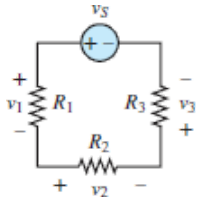


Figure 2.3

Solution

Known Quantities: Source voltage, resistance values.

Find: Unknown voltage v_3.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 10\Omega$; $R_2 = 6\Omega$; $R_3 = 8\Omega$; $v_S = 3$ V. Figure 2.3.

Analysis: The circuit is a simple series loop; that is, all the elements are along the same (the only) current pathway. Apply voltage division directly to solve for v_3:

$$\frac{v_3}{v_S} = \frac{R_3}{R_1 + R_2 + R_3} = \frac{8}{10 + 6 + 8} = \frac{1}{3}$$

Thus: $v_3 = v_S/3 = 1$ V.

Comments: The application of voltage division to a series of elements along a branch is fairly straightforward. However, one has to be careful to correctly determine the sign of the polarity. For example, note that the voltage source in Figure 2.3 guarantees that the upper left node is at a higher voltage than the upper right node. Consequently, the voltage drops v_1, v_2 and v_3 are all positive.

EXAMPLE 2.2 The Wheatstone Bridge

Problem

The **Wheatstone bridge** is a resistive circuit that is frequently encountered in a variety of measurement circuits. The general form of the bridge circuit is shown in Figure 2.4(a), where R_1, R_2, and R_3 are known and R_x is to be determined. The circuit can be redrawn, as shown in Figure 2.4(b), to clarify that R_1 and R_2 are in series, as are R_3 and R_x. The two branches from node c to the reference node are in parallel.

In the figures, v_a and v_b are *node voltages* relative to the common reference node. The value of the reference node can be chosen arbitrarily; however, it may be helpful to consider its value to be zero.

1. Find an expression for the voltage $v_{ab} = v_a - v_b$ in terms of the four resistances and the source voltage v_S.
2. Find the value of R_x when $R_1 = R_2 = R_3 = 1$ kΩ, $v_S = 12$ V, and $v_{ab} = 12$ mV.

Solution

Known Quantities: Source voltage, resistance values, bridge voltage.

Find: Unknown resistance R_x.

Schematics, Diagrams, Circuits, and Given Data: See Figure 2.4.

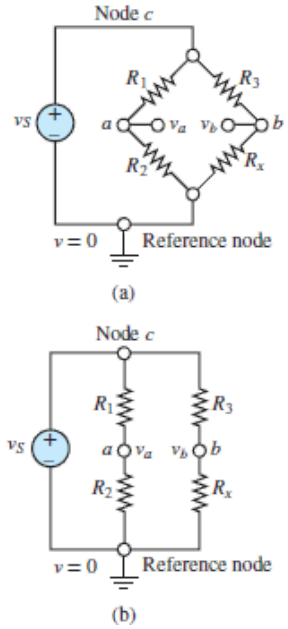


Figure 2.4 A Wheatstone bridge is a mixed series-parallel circuit.

$$R_1 = R_2 = R_3 = 1\ \text{k}\Omega;\ v_S = 12\ \text{V};\ v_{ab} = 12\ \text{mV}.$$

Analysis:

1. The circuit consists of three parallel branches: the voltage source v_S branch, the $R_1 + R_2$ branch, and the $R_3 + R_x$ branch. All three branches sit between node c and the reference node, with the same voltage v_S across each branch.

   In the analysis that follows it is important to keep in mind that all node voltages are understood to be relative to the reference node. That is, v_a is the voltage across R_2, v_b is the voltage across R_x, and $v_c = v_S$.

   Since R_1 and R_2 are in series, voltage division can be applied to find v_a in terms of v_c. Likewise, since R_3 and R_x are in series, voltage division can also be applied to find v_b in terms of v_c.

$$\frac{v_a}{v_c} = \frac{R_2}{R_1 + R_2} \quad \text{and} \quad \frac{v_b}{v_c} = \frac{R_x}{R_3 + R_x}$$

   Plug in $v_c = v_S$ to find that $v_{ab} = v_a - v_b$ is given by:

$$v_{ab} = v_S\left(\frac{R_2}{R_1 + R_2} - \frac{R_x}{R_3 + R_x}\right)$$

   This result is very useful and quite general.

2. Plug in numerical values for v_{ab}, v_S, R_1, R_2, and R_3 in the preceding equation to find:

$$0.012 = 12\left(\frac{1\,\text{k}\Omega}{2\,\text{k}\Omega} - \frac{R_x}{1\,\text{k}\Omega + R_x}\right)$$

Divide both sides by −12 and add 0.5 to both sides to find:

$$0.499 = \frac{R_x}{1\,\text{k}\Omega + R_x}$$

Multiply both sides by $1\text{k}\Omega + R_x$ to find:

$$0.499(1\,\text{k}\Omega + R_x) = R_x \quad \text{or} \quad 499.0 = 0.501R_x \quad \text{or} \quad R_x = 996\,\Omega$$

Comment: The Wheatstone bridge finds application in many measuring instruments.

CHECK YOUR UNDERSTANDING

Use the result of part 1 of Example 2.2 to find the relationship between R_x and the other three resistors such that $v_{ab} = 0$. Using the data in Example 2.2, what is the value of R_x that satisfies $v_{ab} = 0$, the so-called balanced condition for the bridge? Does the balanced bridge condition require that all four resistors be identical?

Answer: $R_1R_x = R_2R_3$; 1 kΩ; No

Page 84

FOCUS ON MEASUREMENTS

Resistive Throttle Position Sensor

Problem:

A typical **automotive resistive throttle position sensor** is shown in Figure 2.5(a). Figure 2.5(b) and (c) depict the geometry of the throttle plate and the equivalent circuit of the throttle sensor. A typical throttle plate has a useful measurement range of just under 90°, from closed throttle to wide-open throttle. The possible mechanical range of rotation of the sensor is usually somewhat greater. It is always necessary to *calibrate* any sensor to determine the actual relationship between the input variable (e.g., the throttle position) and the output variable (e.g., the sensor voltage). The following example illustrates such a procedure.

Figure 2.5(a) Typical throttle position sensor
(*Courtesy: CTS Corporation*)

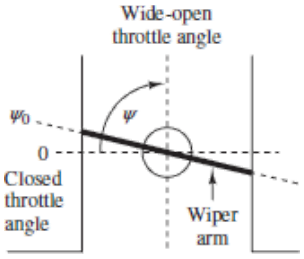


Figure 2.5(b) Throttle blade geometry

Solution:

Known Quantities: Functional specifications of throttle position sensor.

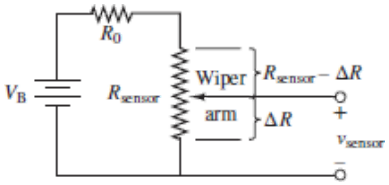


Figure 2.5(c) Throttle position sensor equivalent circuit

Find: Calibration of sensor in volts per degree of throttle plate opening.

Schematics, Diagrams, Circuits, and Given Data:

Functional specifications of throttle position sensor

Total resistance $= R_{sensor} + R_0$	12 kΩ
R_0	3 kΩ
Input V_B	5 V ± 4% regulated
Output V_{sensor}	5% to 95% V_B
Current draw I_S	≤ 20 mA
Recommended load R_L	≤ 220 kΩ
Electrical travel,[1] maximum	112°

[1]Note that in actual operation the sensor will only be actuated between 2° and 90°.

Page 85

Assumptions: Assume a nominal supply voltage of 5 V and total throttle plate travel of 88°, with a closed-throttle angle of 2° and a wide-open throttle angle of 90°.

Analysis: The equivalent circuit of the sensor is a series loop with a battery, a fixed resistor, and a potentiometer, as shown in Figure 2.5(c). The sensor output voltage is determined by the position of the *wiper arm,* whose actual displacement is angular; however, it is convention to depict all potentiometers in circuit diagrams as having straight line displacement, as shown in the figure. The range of the potentiometer

(see specifications above) is 2° to 112° for a resistance of 3 to 12 kΩ; thus, assuming a linear sensor response, the *calibration constant* of the potentiometer is:

$$k_{pot} = \frac{112 - 2}{12 - 3} = 12.22°/\text{k}\Omega, \quad \text{such that } \theta = k_{pot}\Delta R$$

Voltage division requires that the sensor voltage be proportional to the ratio of the series resistances.

$$\begin{aligned} v_{sensor} &= V_B\left(\frac{\Delta R}{R_0 + R_{sensor}}\right) = (5\text{ V})\left(\frac{\Delta R}{12}\right) \\ &= 0.417\ \Delta R \quad \text{V} \quad (\Delta R \text{ in k}\Omega) \\ &= 0.417\frac{\theta}{k_{pot}} \end{aligned}$$

The *calibration curve* for the sensor is shown in Figure 2.5(d).

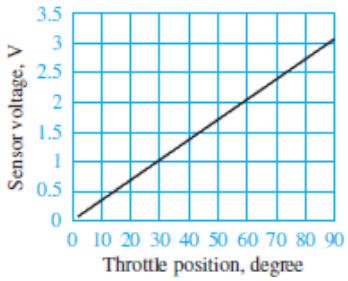


Figure 2.5(d) Calibration curve for throttle position sensor

When the throttle is closed,

$$v_{sensor} = 0$$

and when the throttle is wide open,

$$v_{sensor} = 0.417\ \Delta R = 0.417\frac{\theta}{k_{pot}} = 0.417\ \frac{\text{V}}{\text{k}\Omega}\ \frac{90°}{12.22°/\text{k}\Omega} = 3.07\text{ V}$$

Comments: The fixed resistor R_0 prevents the wiper arm from inadvertently connecting the + terminal of the battery directly to its − terminal, which would occur if the wiper were shorted to the lower node and θ = 112°. Note that the intended operational range of the sensor is from 2° to 90°, specifically to avoid a harmful short-circuit scenario.

FOCUS ON MEASUREMENTS

Resistance Strain Gauges

A **strain gauge** is a resistive element that has many applications in engineering measurements. A strain gauge contains one or more thin conductive strips, usually encased in an epoxy matrix. These strips shrink or stretch with the surface to which the strain gauge is bonded. Since the resistance of a thin conducting strip is dependent upon its geometry, it is possible to calibrate a strain gauge to relate changes in resistance to material *strain* along the surface. Surface strain can then be related to stress, force, torque, and pressure through various constitutive relations, such as Hooke's law. A variety of strain gauges are available to *transduce* the principal strains (extensional and shear) along a surface. The most versatile and popular strain gauge is a planar rosette, with which all three planar strains can be deduced simultaneously.

Recall that the resistance of a cylindrical conductor of cross-sectional area A, length L, and resistivity ρ is given by the expression

$$R = \rho \frac{L}{A}.$$

When the conductor is compressed or elongated, both the length L and (due to the Poisson effect) cross-sectional area A will change, and with them the resistance of the conductor. In particular, when the length of the conductor is increased, its cross-sectional area will decrease, with both changes causing its resistance to increase.

Likewise, when the length of the conductor is decreased, its cross-sectional area will increase, with both changes causing its resistance to decrease. The empirical

relationship between change in resistance and change in length is defined as the gauge factor GF:

$$\text{GF} = \frac{\Delta R/R}{\Delta L/L}$$

The fractional change in length of an object is defined as the strain ε:

$$\varepsilon = \frac{\Delta L}{L}$$

Using these definitions, the change in resistance due to an applied strain ε is given by

$$\Delta R = R_0 \, \text{GF} \varepsilon,$$

where R_0 is the *zero strain resistance.* The value of GF for metal foil resistance strain gauges is usually about 2.

Figure 2.6 depicts a typical foil strain gauge. The maximum strain that can be measured by a foil gauge is $\Delta L / L_{\text{max}} \approx 0.005$, which corresponds to a maximum change in resistance of approximately 1.2Ω for a 120-Ω gauge. Because of the small scale of the change in resistance, strain gauges are usually incorporated in a Wheatstone bridge, which increases the sensitivity of the resistance measurement.

Figure 2.6 Metal-foil resistance strain gauge

Comments: Resistance strain gauges are used in many measurement applications. One such application is the measurement of a force on a cantilever beam, which is discussed in the next example.

FOCUS ON MEASUREMENTS

The Wheatstone Bridge and Force Measurements

One of the simplest applications of a strain gauge is the measurement of a force applied to a cantilever beam, as illustrated in Figure 2.7.

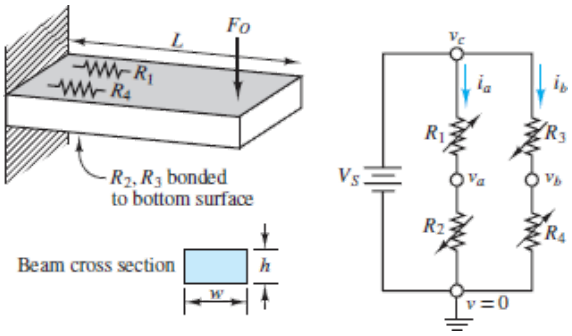


Figure 2.7 A force-measuring instrument

Four strain gauges are employed in this case, of which two are bonded to the upper surface of the beam at a distance L from the point where the external force F_0 is applied, and two are bonded on the lower surface, also at a distance L. Under the influence of the external force, the beam deforms and causes the upper gauges to extend and the lower gauges to compress. Thus, the resistance of the upper gauges will increase by an amount ΔR, and that of the lower gauges will decrease by an equal amount, assuming that the gauges are symmetrically placed. Let R_1 and R_4 be the upper gauges and R_2 and R_3 the lower gauges. Thus, under the influence of the external force, we have

$$R_1 = R_4 = R_0 + \Delta R$$

$R_2 = R_3 = R_0 - \Delta R$

where R_0 is the zero strain resistance of the gauges. It can be shown from elementary strength of materials and statics that the relationship between the strain ε and a force F_0 applied at a distance L for a cantilever beam is

$$\varepsilon = \frac{6LF_0}{wh^2Y}$$

where h and w are as defined in Figure 2.7 and Y is Young's modulus for the beam.

In the circuit of Figure 2.7, the voltages v_a and v_b are given by voltage division.

$$\frac{v_a}{V_S} = \frac{R_2}{R_1 + R_2} \quad \text{and} \quad \frac{v_b}{V_S} = \frac{R_4}{R_3 + R_4}$$

The bridge output voltage is defined as $v_o \equiv v_b - v_a$ such that:

$$\frac{v_o}{V_S} = \frac{R_4}{R_3 + R_4} - \frac{R_2}{R_1 + R_2}$$

Plug in $R_1 = R_4 = R_0 + \Delta R$ and $R_2 = R_3 = R_0 - \Delta R$ to find:

$$\begin{aligned}\frac{v_o}{V_S} &= \frac{R_0 + \Delta R}{R_0 + \Delta R + R_0 - \Delta R} - \frac{R_0 - \Delta R}{R_0 + \Delta R + R_0 - \Delta R} \\ &= \frac{1}{2R_0}[R_0 + \Delta R - (R_0 - \Delta R)] \\ &= \frac{\Delta R}{R_0} = (\mathrm{GF})\varepsilon\end{aligned}$$

Page 88

where GF is the gauge factor and $\Delta R/R_0 = (\mathrm{GF})\varepsilon$ was obtained in the previous example "Focus on Measurements: Resistance Strain Gauges." Thus, the relationship between the bridge output voltage and the force F_0 is

$$v_o = V_S(\mathrm{GF})\varepsilon = V_S(\mathrm{GF})\frac{6LF_0}{wh^2Y} = \frac{6V_S(\mathrm{GF})L}{wh^2Y}F_0$$

where $6V_S(\mathrm{GF})L/wh^2Y$ is the calibration constant for this force transducer.

Comments: **Strain gauge bridges** are commonly used in mechanical, chemical, aerospace, biomedical, and civil engineering applications to make measurements of force, pressure, torque, stress, or strain.

CHECK YOUR UNDERSTANDING

Compute the full-scale (i.e., largest) output voltage for the force-measuring apparatus of "Focus on Measurements: The Wheatstone Bridge and Force Measurements." Assume that the strain gauge bridge is to measure forces ranging from 0 to 500 newtons (N), $L = 0.3$ m, $w = 0.05$ m, $h = 0.01$ m, GF = 2, and Young's modulus for the beam is 69×10^9 N/m^2 (aluminum). The source voltage is 12 V. What is the calibration constant of this force transducer?

Answer: v_o (full scale) = 62.6 mV; $k = 0.125$ mV/N

2.2 RESISTORS IN PARALLEL AND CURRENT DIVISION

It is common to find two or more circuit elements situated between the same two nodes; that is, the elements are in *parallel.* When elements are in parallel, the current entering either of the two nodes is divided among the parallel elements. This important observation is known as **current division**.

The most fundamental instance of current division occurs when two resistors are in parallel, as shown in Figure 2.8. KCL applied at the upper node requires the current i_S through the source to be equal to the sum of the currents i_1 and i_2 through the two resistors.

$$i_S = i_1 + i_2 \quad \text{KCL}$$

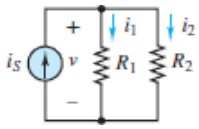


Figure 2.8 The voltage v is across each of the three elements in parallel. KCL requires $I_S = i_1 + i_2$.

Ohm's law can be applied to each resistor to find expressions for i_1 and i_2. (Notice the passive sign convention.)

$$i_1 = \frac{v}{R_1} \quad \text{and} \quad i_2 = \frac{v}{R_2}$$

Plug in for i_1 and i_2 to find:

$$i_S = \frac{v}{R_1} + \frac{v}{R_2} = v\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \equiv v\frac{1}{R_{EQ}}$$

Page 89

This expression defines the *equivalent resistance* R_{EQ} of two resistors in parallel, where:

$$\frac{1}{R_{EQ}} = \frac{1}{R_1} + \frac{1}{R_2} \quad \text{(two resistors in parallel)}$$

However, this inverted expression for the equivalent resistance is awkward and nonintuitive. Often, a more useful form is

$$R_{EQ} = R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2} \quad \text{(two resistors in parallel)}$$

It is easy to show that R_{EQ} is smaller than either R_1 or R_2. To do so, simply write R_{EQ} as:

$$R_{EQ} = R_1 \parallel R_2 = R_1\frac{R_2}{R_1 + R_2} = R_2\frac{R_1}{R_1 + R_2}$$

Both fractions in the above equation are less than 1; thus, $R_{EQ} < R_1$ and $R_{EQ} < R_2$ The notation $R_1 \mathbin{/\!/} R_2$ indicates a parallel combination of R_1 and R_2. The same notation can be used to indicate a parallel combination of three or more resistors by writing:

$$R_1 \parallel R_2 \parallel R_3 \ldots$$

When three or more resistors are connected in parallel, as shown in Figure 2.9, the inverse of the equivalent resistance is equal to the sum of the inverses of all the resistances.

$$\frac{1}{R_{EQ}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_N} \tag{2.3}$$

or

$$R_{EQ} = \frac{1}{1/R_1 + 1/R_2 + \cdots + 1/R_N} \quad \text{Equivalent parallel resistance} \tag{2.4}$$

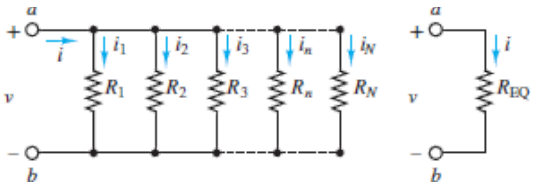


Figure 2.9 The inverse of the equivalent resistance of three or more resistances in parallel equals the sum of the inverses of those resistances.

The equivalent parallel resistance R_{EQ} is always smaller than any of the individual resistances in parallel. It is often useful to replace two or more resistances in parallel with a single equivalent resistance, as indicated in Figure 2.9. To do so correctly, remove all the resistances between nodes *a* and *b* and replace them with a single equivalent resistance attached between these same two nodes. This simple procedure illustrates a very important principle: From the perspective of whatever else is eventually attached to nodes *a* and *b* (e.g., a current source such as that Page 90shown in Figure 2.8), the parallel resistances are *seen* as a single resistance of value R_{EQ}.

An expression for how the current entering either of the two nodes is divided among the individual resistances in parallel can be found by using Ohm's law and noting that the voltage is the same across each resistance. Consider the parallel circuit in Figure 2.8:

$$v = i_1 R_1 = i_2 R_2 = i_S R_{EQ}$$

which yields the following relationships:

$$\frac{i_1}{i_S} = \frac{R_{EQ}}{R_1} \quad \text{and} \quad \frac{i_2}{i_S} = \frac{R_{EQ}}{R_2} \quad \text{and} \quad \frac{i_1}{i_2} = \frac{R_2}{R_1}$$

These results, known as *current division,* indicate that *for resistors in parallel the ratio of currents equals the inverse ratio of the corresponding resistances.* The currents i_1 and i_2 are fractions of the total current i_S because R_1 and R_2 are both greater than R_{EQ}.

> The ratio of the currents through any two resistances in parallel equals the inverse ratio of those resistances.

The current division results shown in the previous equation can be rewritten by plugging in for R_{EQ} to find:

$$\frac{i_1}{i_S} = \frac{R_2}{R_1 + R_2} \quad \text{and} \quad \frac{i_2}{i_S} = \frac{R_1}{R_1 + R_2} \quad \text{and} \quad \frac{i_1}{i_2} = \frac{R_2}{R_1} \tag{2.5}$$

In these forms, current division for two resistances in parallel is that the ratio of i_1 to i_S equals the ratio of the "other" resistance R_2 to the sum of the two resistances $R_1 + R_2$. Likewise, the ratio of i_2 to i_S equals the ratio of the "other" resistance R_1 to the sum of the two resistances $R_1 + R_2$. These forms may be appealing since they resemble the expressions used to compute voltage division for two resistances in series.

When parallel connections are encountered in circuit diagrams, one should immediately think of current division.

Parallel Connection ⇒ Current Division

It is important to realize that the current division rule applies to any two resistances in parallel, not just any two discrete resistors. For example, consider the parallel resistors shown in Figure 2.9. The ratio of the combined current through R_1 and R_2 to the current through R_3 equals the ratio of R_3 to $(R_{12})_{EQ}$. That is:

$$\frac{i_1 + i_2}{i_3} = \frac{R_3}{(R_{12})_{EQ}} \quad \text{Current division}$$

where

$$(R_{12})_{EQ} = \frac{R_1 R_2}{R_1 + R_2}$$

Likewise:

$$\frac{i_n}{i} = \frac{(R_{1\cdots N})_{EQ}}{R_n} \quad \text{Current division}$$

where

$$\frac{1}{(R_{1\cdots N})_{EQ}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_N}$$

These last two expressions can be combined to yield:

$$\frac{i_n}{i} = \frac{1/R_n}{1/R_1 + 1/R_2 + \cdots + 1/R_n + \cdots + 1/R_N} \quad \text{Current divider} \tag{2.6}$$

Many practical circuits contain resistors in parallel and in series. The concepts of equivalent resistance, voltage division, and current division are useful even in very complicated circuits.

EXAMPLE 2.3 Current Division

Problem

Determine the current i_1 in the circuit of Figure 2.10.

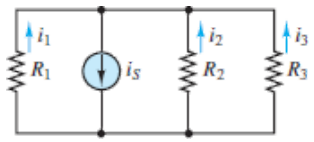


Figure 2.10

Solution

Known Quantities: Source current, resistance values.

Find: Unknown current i_1.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 10\Omega$; $R_2 = 2\Omega$; $R_3 = 20\Omega$; $i_S = 4$ A. Figure 2.10.

Analysis: Apply current division directly to find:

$$\frac{i_1}{i_S} = \frac{1/R_1}{1/R_1 + 1/R_2 + 1/R_3} = \frac{\frac{1}{10}}{\frac{1}{10} + \frac{1}{2} + \frac{1}{20}} = \frac{2}{13}$$

Thus:

$$i_1 = 4 \text{ A} \times \frac{2}{13} \approx 0.62 \text{ A}$$

An alternative approach is to find the equivalent resistance of $R_2 \mathbin{/\!/} R_3$ and then apply one of the simpler expressions for current division between two resistances in parallel.

$$R_2 \parallel R_3 = \frac{R_2 R_3}{R_2 + R_3} = \frac{(2)(20)}{2 + 20} \approx 1.82\ \Omega$$

(Notice that $R_2 \mathbin{/\!/} R_3$ is less than both R_2 and R_3.)

$$\frac{i_1}{i_S} = \frac{R_2 \parallel R_3}{R_2 \parallel R_3 + R_1} \approx \frac{1.82}{1.82 + 10} = \frac{2}{13}$$

The result is the same as that found by applying current division directly:

$$i_1 = 4\ \text{A} \times \frac{2}{13} \approx 0.62\ \text{A}$$

Comments: The application of current division to elements in parallel between two nodes is fairly straightforward. Occasionally, it may be difficult to determine which elements are, in fact, in parallel. This issue is explored in Example 2.4.

Page 92

EXAMPLE 2.4 Resistors in Series and Parallel

Problem

Determine the voltage v in the circuit of Figure 2.11.

Solution

Known Quantities: Source voltage, resistance values.

Find: Unknown voltage v.

Schematics, Diagrams, Circuits, and Given Data: See Figures 2.11, 2.12.

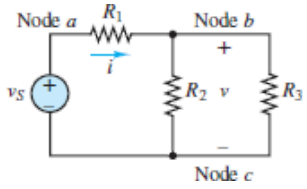


Figure 2.11

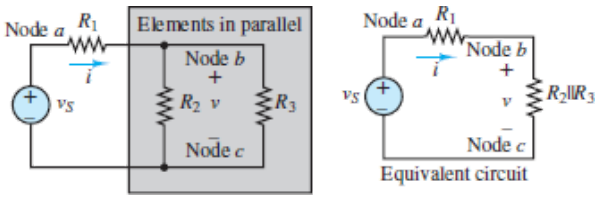


Figure 2.12

Analysis: The circuit of Figure 2.11 contains three resistors that are not completely in series nor in parallel with each other. This fact may not be apparent at first glance, but consider whether the conditions for series and parallel are met for all three resistors.

1. Are all three resistors in series? Are all three resistors situated along the same branch? Is there one common current through all three resistors? Clearly, the current i entering node b will be divided on its way to node c. Some of it will pass through R_2 while the rest will pass through R_3. Thus, there is *not* one common current through all three resistors; that is, they are *not* in series.
2. Are all three resistors in parallel? Are all three resistors situated between the same two nodes? R_1 sits between nodes a and b, while R_2 and R_3 sit between nodes b and c. Thus, the three resistors do *not* sit between the same two nodes; that is, they are *not* in parallel.

However, it is possible to simplify the circuit by noting, as mentioned above and depicted in Figure 2.12, that R_2 and R_3 sit between the same two nodes and are, therefore, in parallel. These two resistors can be removed from the circuit and replaced by the equivalent resistance between nodes b and c, which is:

$$R_{EQ} = R_2 \parallel R_3 = \frac{R_2 R_3}{R_2 + R_3}$$

An equivalent circuit can now be drawn as shown in Figure 2.12. The result is a simple series loop. Voltage division can be applied directly to solve for v:

$$v = \frac{R_2 \parallel R_3}{R_1 + R_2 \parallel R_3} v_S$$

The current can also be found:

$$i = \frac{v}{R_2 \parallel R_3} = \frac{v_S}{R_1 + R_2 \parallel R_3}$$

CHECK YOUR UNDERSTANDING

For the circuit in Figure 2.10, apply current division to find i_2 and i_3 and verify that KCL at either node is satisfied by the results. Also, verify that the ratio of any two branch currents equals the inverse ratio of their associated resistances. Finally, verify that $i_2 = 5 \times i1$ because $R_1 = 5 \times R_2$ and that $i_1 = 2 \times i_3$ because $R_3 = 2 \times R_1$. (These

results should not be a surprise since larger currents are expected through the smaller resistances.)

CHECK YOUR UNDERSTANDING

Consider the circuit of Figure 2.11, with resistor R_3 replaced by an open-circuit. Calculate the voltage v when the source voltage is $v_S = 5$ V and $R_1 = R_2 = 1$ kΩ.

Repeat when resistor R_3 is in the circuit and its value is $R_3 = 1$ kΩ.

Repeat when resistor R_3 is in the circuit and its value is $R_3 = 0.1$ kΩ.

Answer: $v = 2.50$ V; $v = 1.67$ V; $v = 0.4167$ V

2.3 EQUIVALENT RESISTANCE BETWEEN TWO NODES

The concept of *equivalent resistance* was introduced earlier as part of the discussions of resistors in series and resistors in parallel. In those discussions the following results were established:

1. For two resistors R_1 and R_2 in series between nodes a and b (see Figure 2.13) the equivalent resistance between those two nodes is

$$R_{EQ} = R_1 + R_2 \qquad \text{Two resistors in series}$$

2. For two resistors R_1 and R_2 in parallel between nodes a and b (see Figure 2.14) the equivalent resistance between those two nodes is

$$R_{EQ} = \frac{R_1 R_2}{R_1 + R_2} \qquad \text{Two resistors in parallel}$$

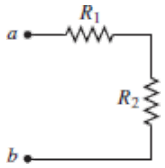


Figure 2.13 Two resistors in series between nodes a and b

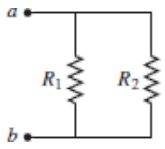


Figure 2.14 Two resistors in parallel between nodes *a* and *b*

In general, the equivalent resistance of multiple resistors in series is larger than the largest of the resistors; also, the equivalent resistance of multiple resistors in parallel is smaller than the smallest of the resistors.

The central idea in both cases is that two resistors in series or parallel *between two nodes* can be replaced by a single resistor whose value is *the equivalent resistance between those same two nodes*. In problem solving one can literally remove the pair of resistors and insert a single resistor in their place. Complicated resistor networks can often be simplified by repeated application of this process of substituting equivalent resistances for series and parallel connections.

Any two nodes or terminals have an equivalent resistance between them. Equivalent resistance is not a characteristic of a circuit, but of two nodes.

Page 94

In general, the equivalent resistance between two terminals (nodes) is defined by

$$R_{EQ} \equiv \frac{v_s}{i_s}$$

where v_s is a voltage applied across the two terminals and i_s is the resulting current into and out of those same terminals. This definition yields the correct equivalent resistance when applied to two resistors in series (see Figure 2.13) and to two resistors in parallel (see Figure 2.14). However, in many cases, it is not necessary to resort to the definition of equivalent resistance in order to calculate it. Instead, the series and parallel resistance results can be used to reduce a complicated network of resistors to a single equivalent resistance.

Redrawing a Resistive Network

By convention, electric circuits and networks are drawn in a rectilinear fashion, consisting of straight line segments drawn along the horizontal, the vertical, and

occasionally the diagonal. With practice, these networks can be interpreted correctly. However, some networks can appear complicated and even misleading, particularly to an analyst with little experience. In such cases it is often helpful to redraw the circuit with a focus on discerning which elements sit between the same two nodes and which sit along the same branch. Consider finding the equivalent resistance between the two open terminals shown in the resistive network of Figure 2.15(a). In the figures, K is shorthand for kilo-ohms.

1. Carefully count the number of nodes in the network. Mark each node with a boundary and a letter from the sequence *A*, *B*, *C*, ... The two open terminals should be labeled with the first and last letter in the sequence. Label adjacent nodes with adjacent letters. For instance, make sure node *B* is adjacent to nodes *A* and *C* when labeling the network. See Figure 2.15(b).
2. Begin the redrawn network by drawing, for each node, one small circle ○ equally spaced along a straight line, either horizontal or vertical. Be sure to allow adequate space between each circle to draw a resistor.
3. Label each circle in order using the sequence of letters *A*, *B*, *C*, ... These circles represent the nodes in the network.
4. One at a time, place each resistor present in the original network between the same two nodes (circles) in the redrawn network. See Figure 2.15(c), (d), and (e).

Page 95

5. Use the redrawn network and the definitions of series and parallel connections to identify resistors in series and resistors in parallel. Replace those resistors in series and those resistors in parallel with their equivalent resistance. See Figure 2.15(f).
6. Repeat step 5 as often as necessary until a single resistance sits between the first and last node in the sequence. This resistance is the equivalent resistance between those two nodes. See Figure 2.15(g), (h), and (i).

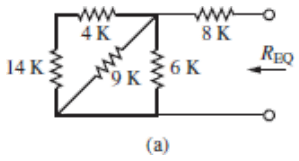
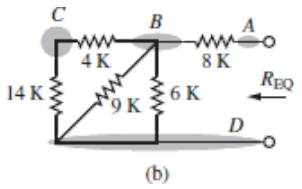
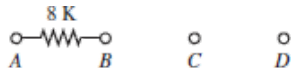


Figure 2.15(a)



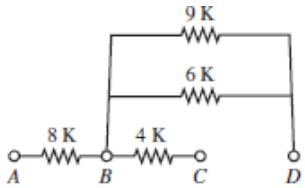
(b)

Figure 2.15(b)



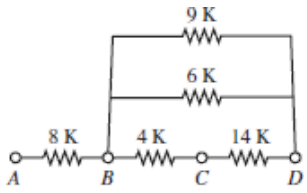
(c)

Figure 2.15(c)



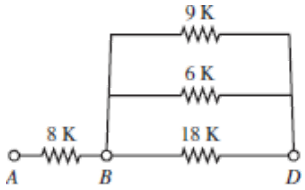
(d)

Figure 2.15(d)



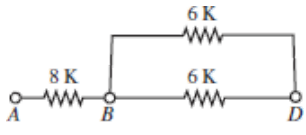
(e)

Figure 2.15(e)



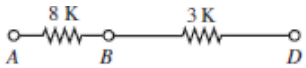
(f)

Figure 2.15(f)



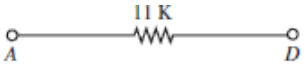
(g)

Figure 2.15(g)



(h)

Figure 2.15(h)



(i)

Figure 2.15(i)

An alternative layout for a network with three total nodes is to place the circles to form the three vertices of an equilateral triangle. Likewise, for a network with four total nodes, place the circles to form the four vertices of a square or diamond. These layouts may also work well. For networks with more than four nodes, the straight line layout is usually best.

Calculation and Approximation Tips

It is tempting to rely solely on an electronic calculator to produce quick equivalent resistance values, particularly for resistors in parallel. However, there are a few easy techniques that allow for quick calculation of equivalent resistance values by observation. Learning these techniques is valuable for inspecting resistive networks, particularly when someone else is working an example on a classroom board or discussing a resistive network that is part of a larger project one is working on. These techniques allow one to follow along with confidence and comfort and without being distracted by a calculator.

To begin, it is important to remember that the equivalent resistance of two or more resistors in series is *greater than* the largest individual resistor. Also, the equivalent resistance of two or more resistors in parallel is *less than* the smallest individual resistor. These two facts provide quick lower and upper bounds on the value of the equivalent resistance for series and parallel arrangements, respectively.

Of course, calculating the equivalent resistance of two or more resistors in series is easy since it is simply the sum of the individual resistors. On the other hand, calculating the equivalent resistance of two or more resistors in parallel is not quite so easy, at first glance. However, consider the formula for the equivalent resistance of two resistors in parallel:

$$R_{EQ} = R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2}$$

Assume that R_2 is N times larger than R_1 such that

$$R_2 = N \cdot R_1$$

Page 96

Plug in for R_2:

$$R_{EQ} = R_1 \parallel NR_1 = \frac{N \cdot R_1^2}{R_1 + N \cdot R_1} = \frac{N \cdot R_1}{1 + N} = \frac{R_2}{1 + N}$$

For example, when N is an integer:

$$R_2 = R_1 \qquad R_{EQ} = \frac{R_2}{2} \tag{2.7}$$

$$R_2 = 2R_1 \qquad R_{EQ} = \frac{R_2}{3} \tag{2.8}$$

$$R_2 = 3R_1 \qquad R_{EQ} = \frac{R_2}{4} \tag{2.9}$$

$$R_2 = 4R_1 \qquad R_{EQ} = \frac{R_2}{5} \tag{2.10}$$

In words, for two resistors in parallel, when R_2 equals R_1, the equivalent resistance is one-half of R_2; when R_2 is twice R_1, the equivalent resistance is one-third of R_2; when R_2 is thrice R_1, the equivalent resistance is one-fourth of R_2; when R_2 is four times R_1, the equivalent resistance is one-fifth of R_2; and so on with the same pattern.

This relationship $R_{EQ} = R_2/(1 + N)$ (where R_2 is the larger of the two resistors in parallel) holds for noninteger values of N as well, although the calculation is somewhat more difficult.

In practice, when $N \leq 10$, it is often acceptable to approximate the equivalent resistance of two resistors in parallel by the smaller of the two resistors. This practice is known as the 10:1 (ten-to-one) rule. It is easy to see from the above formula that the error associated with this approximation is less than 10 percent. One should be careful not to apply this approximation repeatedly to the same network because of the potential for the errors to accumulate.

The Wye-Delta Transformation

Occasionally, a circuit will contain resistors that are neither in series nor parallel with any other resistors. For example, consider the Wheatstone bridge networks shown in Figure 2.4(a) and (b) when a resistor R_5 is attached to terminals v_a and v_b. After a bit

of inspection, it becomes clear that *none* of the five resistors are in series or parallel. In such cases, the so-called *wye-delta* (or *Y*-Δ) transformation can be used to produce series or parallel connections that can then be simplified, as needed.

Figure 2.16(a) and (b) depict generic wye and delta resistor networks, respectively. Notice that each network has three exterior nodes *A*, *B*, and *C*. These two networks are interchangeable if the equivalent resistance between any pair of nodes in one network equals the equivalent resistance between the same pair of nodes in the other network. In general, to evaluate the equivalent resistance between a pair of nodes it is sufficient to attach an ideal source across those nodes and compute the following ratio:

$$R_{EQ} \equiv \frac{v_S}{i_S} \qquad \text{Definition of equivalent resistance}$$

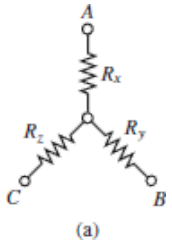


Figure 2.16(a)

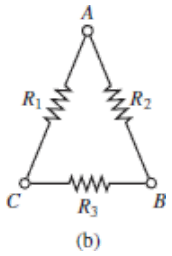


Figure 2.16(b)

Page 97

This approach was used earlier to determine the equivalent resistance of two resistors in series (Figure 2.1) and two resistors in parallel (Figure 2.8) when it was asserted that $v = iR_{EQ}$. To avoid unnecessary details, notice that when an ideal source is placed across two of the three nodes in either Figure 2.16(a) or (b), the third node is not affected. The result is that the equivalent resistance *seen* by the ideal source is found easily by inspection. For example, when an ideal source is placed across nodes *A*-*B*, the equivalent resistance seen by the source is simply:

$$R_{AB} = R_x + R_y \qquad \text{Wye network} \tag{2.11}$$

$$= R_2 \| (R_1 + R_3) = \frac{R_2(R_1 + R_3)}{R_1 + R_2 + R_3} \qquad \text{Delta network} \tag{2.12}$$

Likewise, when an ideal source is placed across nodes *A*-*C* and across nodes *B*-*C*, the equivalent resistance seen by the source is simply:

$$R_{AC} = R_x + R_z \qquad \text{Wye network} \tag{2.13}$$

$$= R_1 \| (R_2 + R_3) = \frac{R_1(R_2 + R_3)}{R_1 + R_2 + R_3} \qquad \text{Delta network} \tag{2.14}$$

$$R_{BC} = R_y + R_z \qquad \text{Wye network} \tag{2.15}$$

$$= R_3 \| (R_1 + R_2) = \frac{R_3(R_1 + R_2)}{R_1 + R_2 + R_3} \qquad \text{Delta network} \tag{2.16}$$

The above three equations for R_{AB}, R_{AC}, and R_{BC} relate the wye resistances R_x, R_y, and R_z to the delta resistances R_1, R_2, and R_3. These equations can be solved to yield the following results:

$$R_x = \frac{R_1 R_2}{R_1 + R_2 + R_3} \tag{2.17}$$

$$R_y = \frac{R_2 R_3}{R_1 + R_2 + R_3} \tag{2.18}$$

$$R_z = \frac{R_1 R_3}{R_1 + R_2 + R_3} \tag{2.19}$$

or

$$R_1 = \frac{R_x R_y + R_x R_z + R_y R_z}{R_y} \tag{2.20}$$

$$R_2 = \frac{R_x R_y + R_x R_z + R_y R_z}{R_z} \tag{2.21}$$

$$R_3 = \frac{R_x R_y + R_x R_z + R_y R_z}{R_x} \tag{2.22}$$

These two sets of equations can be used to transform a delta network into a wye network, or vice versa. The key to using the transformation correctly is to detach the one network at the three nodes *A*, *B*, and *C* and attach the other network at exactly the same three nodes, as indicated in Figure 2.16(a) and (b).

EXAMPLE 2.5 Equivalent Resistance Between Two Nodes

Problem

Determine the equivalent resistance between nodes $a \leftrightarrow b$ and $a \leftrightarrow c$ in the resistive network shown in Figure 2.17(a).

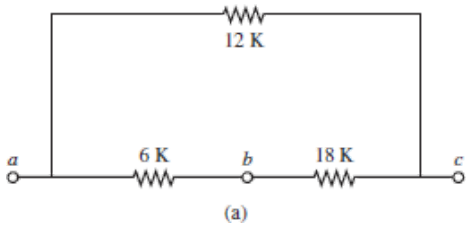


Figure 2.17(a)

Solution

Known Quantities: Resistance values in the network.

Find: Equivalent resistance between nodes $a \leftrightarrow b$ and $a \leftrightarrow c$.

Schematics, Diagrams, Circuits, and Given Data: Figure 2.17(a) to (e).

Analysis: Apply the equivalent resistance formulas for series and parallel resistors to reduce the network to a single equivalent resistance.

1. To find the equivalent resistance between nodes a and b note that there are two pathways from a to b: one pathway is directly through the 6K resistor; the other pathway is through the 12K and 18K resistors. Therefore, the latter two resistors are in series, as "seen" from a to b, and can be replaced with a single 30K resistor as shown in Figure 2.17(b). Also, the two pathways from a to b are in parallel; that is, the 6K and 30K resistances sit between the same two nodes, a and b. Their parallel equivalent resistance is $(6 \cdot 30)/(6 + 30)$ or 5K, which could also be calculated as (30/6)K because 30 is five times larger than 6. See Figure 2.17(c).

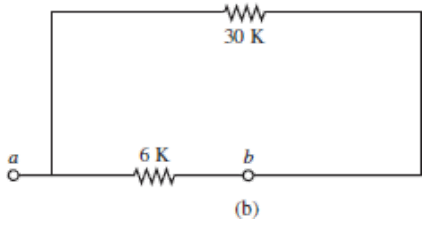


Figure 2.17(b)

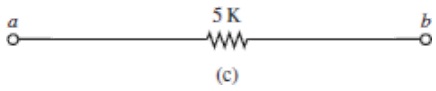


Figure 2.17(c)

2. To find the equivalent resistance between nodes *a* and *c* note that there are two pathways from *a* to *c*: one pathway is directly through the 12K resistor; the other pathway is through the 6K and 18K resistors. Therefore, the latter two resistors are in series, as "seen" from *a* to *c*, and can be replaced with a single 24K resistor as shown in Figure 2.17(d). Also, the two pathways from *a* to *c* are in parallel; that is, the 12K and 24K Page 99resistances sit between the same two nodes, *a* and *c*. Their parallel equivalent resistance is $(12 \cdot 24)/(12 + 24)$ or 8K, which could also be calculated as (24/3)K because 24 is two times larger than 12. See Figure 2.17(e).

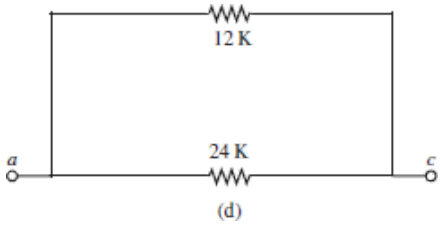


Figure 2.17(d)

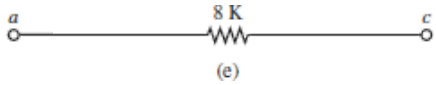


Figure 2.17(e)

Comments: Notice that the equivalent resistance between nodes $a \leftrightarrow b$ is not the same as that between nodes $a \leftrightarrow c$. This result demonstrates that networks do not have an equivalent resistance, but two nodes do have an equivalent resistance between them. For two different pairs of nodes in the same network it is quite possible to get two different values for the equivalent resistance between them.

EXAMPLE 2.6 Equivalent Resistance Between Two Nodes

Problem

Determine the equivalent resistance between nodes $A \leftrightarrow E$ in the resistive network shown in Figure 2.18(a).

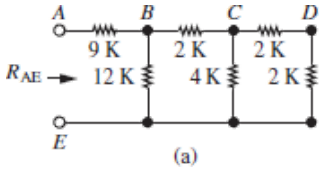


Figure 2.18(a)

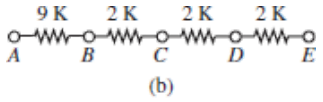


Figure 2.18(b)

Solution

Known Quantities: Resistance values in the network.

Find: Equivalent resistance between nodes $A \leftrightarrow E$.

Schematics, Diagrams, Circuits, and Given Data: Figure 2.18(a)-(g).

Analysis: Follow the procedure for redrawing a resistive network in a rectilinear fashion.

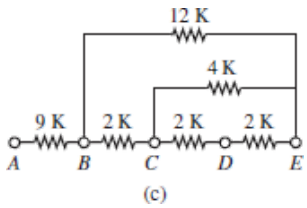


Figure 2.18(c)

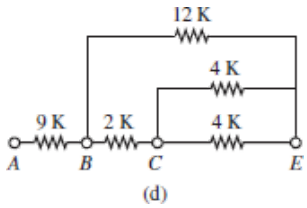


Figure 2.18(d)

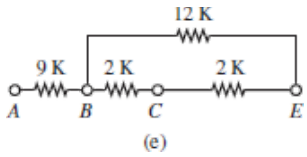


Figure 2.18(e)

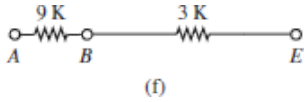


Figure 2.18(f)

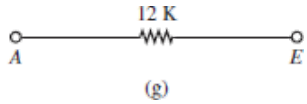


Figure 2.18(g)

1. Begin by noting that there are five nodes (*A* . . . *E*) in the original network. To the extent possible, adjacent nodes are marked with adjacent labels.
2. Lay out five circles along a straight line and label each circle with the *A* . . . *E* sequence.
3. For each pair of adjacent nodes, insert a resistor as appropriate. As shown in Figure 2.18(a), there is a 9K resistor between nodes *A* and *B* and there is a 2K resistor between each pair of nodes *BC*, *CD*, and *DE*. Figure 2.18(b) shows these resistors placed in the redrawn rectilinear network.

4. Two other resistors remain to be placed. The 4K and 12K resistors sit between nodes *CE* and *BE*, respectively. Figure 2.18(c) shows the completed redrawn rectilinear network with these two resistors in place.
5. One step at a time, simplify the network by replacing resistors in series and resistors in parallel with their equivalent resistance. In Figure 2.18(c) notice that the pair of 2K resistors between nodes *CD* and *DE* are in series. Replace them with their equivalent 4K resistance as shown in Figure 2.18(d).
6. Next, notice that there are now two 4K resistances between nodes *C* and *E*. Replace them with their 2K parallel equivalent resistance as shown in Figure 2.18(e).
7. Now there are two 2K resistors in series between nodes *B* and *E*. Replace them with their 4K series equivalent resistance. Immediately notice that this 4K resistance is in parallel with the 12K resistance that is also between nodes *B* and

E. Replace both resistances with their 3K parallel equivalent resistance as shown in Figure 2.18(f).

8. Finally, the 9K resistance is in series with the 3K resistance and can be replaced with its 12K equivalent resistance as shown in Figure 2.18(g).

Comments: This rather long and detailed procedure will become shorter and more efficient with practice. Try it out on the following Check Your Understanding problems.

CHECK YOUR UNDERSTANDING

Find the equivalent resistance between nodes *b* and *c* in Figure 2.17(a).

Answer: 9K

CHECK YOUR UNDERSTANDING

Find the equivalent resistance between nodes *A* and *C* in Figure 2.18(a).

Answer: 10.75K

2.4 LINEAR NETWORKS AND THE PRINCIPLE OF SUPERPOSITION

In general, the criteria for a linear function are

Superposition: If $y_1 = f(x_1)$ and $y_2 = f(x_2)$, then $y_1 + y_2 = f(x_1 + x_2)$.
Homogeneity: If $y = f(x)$, then $\alpha y = f(\alpha x)$.

where x is the function input and y is the function output.

Linear networks obey the same rules. Superposition implies that each source (e.g., x_1 and x_2) makes its own independent contribution (e.g., y_1 and y_2) to each current and voltage in a network, and that the total value of each current and voltage is the sum (e.g., $y_1 + y_2$) of these contributions.

Page 101

Homogeneity implies that the contribution due to any one source scales linearly with the value of the source. For example, if the contribution due to source x_1 is y_1, then the contribution due to the same source doubled $2x_1$ is also doubled $2y_1$, where $\alpha = 2$ is an example scaling factor.

In general, to determine if a network is linear, it is necessary to verify that these two criteria are satisfied for all possible inputs, or at least to verify a range of inputs within which the network is linear. Luckily, it is not always necessary to verify superposition and homogeneity directly. A sufficient, but not necessary, alternative condition is

Any network composed of linear elements only is itself linear. Common linear elements are ideal sources, resistors, capacitors, and inductors.

The *principle of superposition* is a valid, and frequently used, analytic tool for any linear circuit. It is also a powerful conceptual aid for understanding the behavior of circuits with multiple sources.

For any linear circuit, the principle of superposition states that each *independent* source contributes to each voltage and current present in the circuit. Moreover, the contributions of one source are independent of those from the other sources. In this way, each voltage and each current in a circuit of N independent sources is the sum of N component voltages and N component currents, respectively.

As a problem-solving tool, the principle of superposition permits a problem to be decomposed into two or more simpler problems. The efficiency of this “divide and

conquer" tactic depends upon the particular problem being solved. However, solutions generated using the principle of superposition reveal the contribution of each independent source in the overall circuit.

The method is to turn off (set to zero) all independent sources but one, and then solve for voltages and currents due to the lone remaining independent source. This procedure may be repeated successively for each source until the contributions due to all the sources have been computed. The components for a particular voltage or current can be summed to find its value in the original complete circuit.

A zero voltage source is equivalent to a short-circuit, and a zero current source is equivalent to an open-circuit. When using the principle of superposition, it is necessary, and helpful, to replace each zero source with its equivalent short- or open-circuit and thus simplify the circuit. These substitutions are summarized in Figure 2.19.

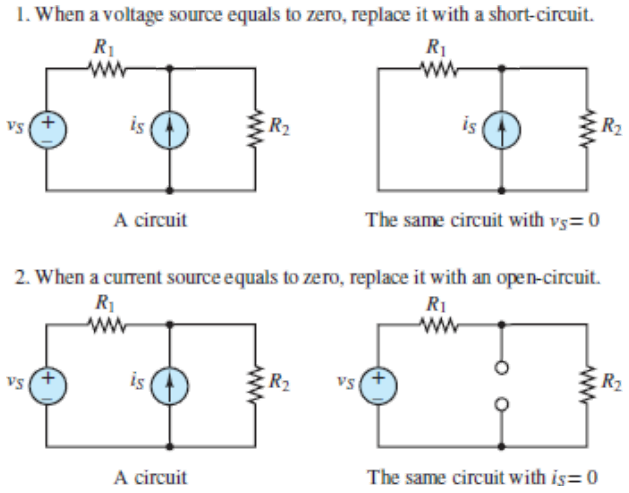


Figure 2.19 Zeroing voltage and current sources

In general, if a circuit is decomposed into multiple simpler circuits, each independent source must contribute once and only once to the components of any voltage or current. In other words, it is not necessary that only one independent source be on at a time. However, it is necessary that each independent source be on only once throughout the entire procedure.

Superposition may be applied to circuits containing *dependent* sources; however, the dependent sources must not be set to zero. They are not independent sources and must not be treated as such.

FOCUS ON PROBLEM SOLVING

SUPERPOSITION

1. Define the voltage v or current i to be solved in the circuit.
2. For each of the N sources, define a component voltage v_k or current i_k such th

$$v = v_1 + v_2 + \cdots + v_N \quad \text{or} \quad i = i_1 + i_2 + \cdots + i_N$$

3. Turn off all sources except source S_k and solve for the component voltage
current i_k. Find components for all k where $k = 1,2,\ldots,N$.
4. Find the complete solution for the voltage v or current i by summing all o
components as defined in step 2.

Details and Examples

An elementary application of the principle is to find the current in a single loop with two sources connected in series, as shown in Figure 2.20.

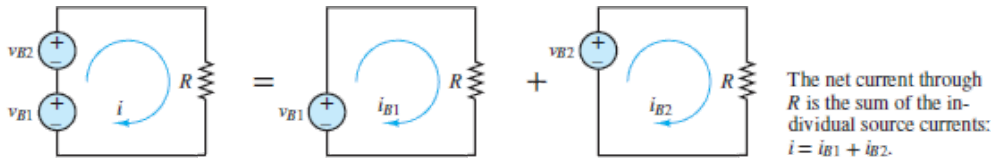


Figure 2.20 The principle of superposition

The current in the far left circuit of Figure 2.20 is easily found by a direct application of KVL and Ohm's law.

$$v_{B1} + v_{B2} - iR = 0 \quad \text{or} \quad i = \frac{v_{B1} + v_{B2}}{R} \tag{2.23}$$

Figure 2.20 depicts the far left circuit as being equivalent to the combined effects of two component circuits, each containing a single source. In each of these two

circuits, one battery (which is a DC voltage source) has been set to zero and replaced with a short-circuit.

KVL and Ohm's law can be applied directly to each of these component circuits.

$$i_{B1} = \frac{v_{B1}}{R} \quad \text{and} \quad i_{B2} = \frac{v_{B2}}{R} \tag{2.24}$$

According to the principle of superposition

$$i = i_{B1} + i_{B2} = \frac{v_{B1}}{R} + \frac{v_{B2}}{R} = \frac{v_{B1} + v_{B2}}{R} \tag{2.25}$$

Voila! The complete solution is found, as expected. This simple example illustrates the essential method; however, more challenging examples are needed to reinforce it.

EXAMPLE 2.7 Principle of Superposition

Problem

Determine the current i_2 in the circuit of Figure 2.21a), using the principle of superposition.

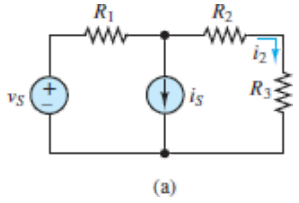


Figure 2.21(a) Circuit for the illustration of the principle of superposition

Solution

Known Quantities: Voltage and current values of each source; resistor values.

Find: Unknown current i_2.

Given Data: $v_S = 10$ V; $i_S = 2$ A; $R_1 = 5\,\Omega$; $R_2 = 2\,\Omega$; $R_3 = 4\,\Omega$..

Analysis: Refer to Figure 2.21(a) and the steps in the Focus on Problem Solving box "Superposition."

1. The objective is to find current i_2.
2. There are two independent sources in the circuit, so there will be two components of i_2.

$$i_2 = i_2' + i_2''$$

3. *Part 1:* Turn off the current source and replace it with an open-circuit. The resulting circuit is a simple series loop shown in Figure 2.21(b). Here, i_2' is the same as the loop current because of the open-circuit. The total series resistance is $5 + 2 + 4 = 11\Omega$, such that $i_2' = 10\text{ V}/11\ \Omega = 0.909\text{ A}$..

   Part 2: Turn off the voltage source and replace it with a short-circuit. The resulting circuit consists of three parallel branches, as shown in Figure 2.21(c): i_S, R_1, and $R_2 + R_3$. Apply current division.

$$i_2'' = (-i_S)\frac{R_1}{R_1 + R_2 + R_3} = (-2\text{A})\frac{5}{5+2+4} = -0.909\text{ A}$$

4. The complete i_2 is found to be

$$i_2 = i_2' + i_2'' = 0.909\text{ A} - 0.909\text{ A} = 0\text{ A}$$

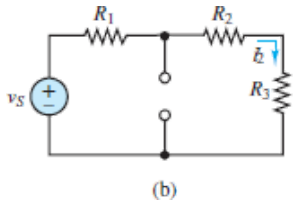


Figure 2.21(b) Circuit with current source set to zero

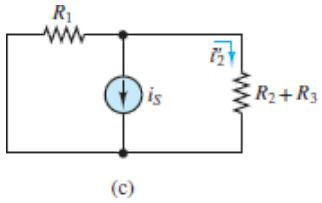


Figure 2.21(c) Circuit with voltage source set to zero

Comments: Superposition is not always a very efficient tool. Beginners may find it preferable to rely on more systematic methods, such as the node voltage method, to solve circuits. However, the importance of the principle of superposition is the insight it lends to the behavior of a circuit.

EXAMPLE 2.8 Principle of Superposition

Problem

Determine the voltage v_R across resistor R in the circuit of Figure 2.22(a).

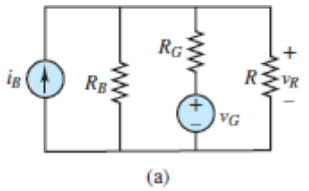


Figure 2.22(a) Circuit used to demonstrate the principle of superposition

Solution

Known Quantities: The values of the sources and resistors in the circuit of Figure 2.22(a) are $i_B = 12$ A; $v_G = 12$ V; $R_B = 1\Omega$; $R_G = 0.3\Omega$; $R = 0.23\Omega$.

Find: v_R.

Analysis: Refer to Figure 2.22(a) and the steps in the Focus on Problem Solving box "Superposition."

1. The objective is to find voltage v_R.
2. There are two independent sources in the circuit, so there will be two components of v_R.

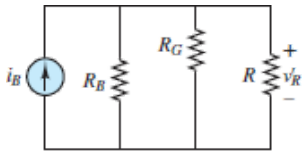
$$v_R = v_R' + v_R''$$

3. *Part 1:* Turn off the voltage source and replace it with a short-circuit. Redraw the circuit as shown in Figure 2.22(b), find the equivalent resistance of all three resistors in parallel, and apply Ohm's law to find v_R'.

$$R_{eq} = (R_B \parallel R_G \parallel R) = \frac{1}{1/R_B + 1/R_G + 1/R} = \frac{1}{1/1 + 1/0.3 + 1/0.23}$$

$$= \frac{(0.3)(0.23)}{(0.3)(0.23) + 0.23 + 0.3} = \frac{0.069}{0.599}\ \Omega$$

$$v_R' = i_B R_{eq} = (12\text{ A})\frac{0.069}{0.599} = 1.38\text{ V}$$



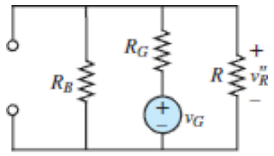
(b)

Figure 2.22(b) Circuit obtained by suppressing the voltage source

Part 2: Turn off the current source and replace it with an open-circuit. Redraw the circuit, as shown in Figure 2.22(c), and apply KCL at the upper node:

$$-\frac{v''_R}{R_B} - \frac{v''_R - v_G}{R_G} - \frac{v''_R}{R} = -v''_R\left[\frac{1}{R_B} + \frac{1}{R_G} + \frac{1}{R}\right] + \frac{v_G}{R_G} = 0$$

$$v''_R = \frac{v_G}{R_G}\frac{1}{1/R_B + 1/R_G + 1/R} = \frac{12}{0.3}\frac{1}{1/1 + 1/0.3 + 1/0.23} = 4.61 \text{ V}$$



(c)

Figure 2.22(c) Circuit obtained by suppressing the current source

This same result can be found by finding the equivalent resistance of R_B in parallel with R and applying voltage division.

$$R_{eq} = R_B \parallel R = \frac{R_B R}{R_B + R} = \frac{0.23}{1.23} \approx 0.187\ \Omega$$

$$v''_R = v_G\frac{R_{eq}}{R_{eq} + R_G} = (12\text{ V})\frac{0.23}{0.23 + (1.23)(0.3)} = 4.61 \text{ V}$$

4. Compute the voltage across R as the sum of the two component voltages:

$$v_R = v'_R + v''_R = 5.99 \text{ V}$$

Comments: The advantage offered by the principle of superposition in this problem is that it clearly reveals the contributions to v_R from each source. However, the work required to solve this problem using this method is extensive compared to other methods. For example, the voltage across R can easily be determined by applying KCL at the upper node.

CHECK YOUR UNDERSTANDING

Find the voltages v_a and v_b for the circuits of Example 1.14 by superposition.

CHECK YOUR UNDERSTANDING

Solve Example 1.17, using superposition.

CHECK YOUR UNDERSTANDING

Solve Example 1.19, using superposition.

2.5 THE SOURCE-LOAD PERSPECTIVE

An important analytic approach used throughout this book is to divide a circuit into two parts that are connected to each other at only two terminals. These two parts are known as the *source* and the *load,* as shown in Figure 2.23. In general, the load is the circuit element or segment of interest. By default, the source is everything not included in the load. Typically, the source provides energy and the load consumes it for some purpose. For example, consider a headlight attached to a car battery as shown in Figure 2.24. For the driver of the car, the headlight is the circuit element of interest since it enables the driver to see the road at night. From this perspective, the headlight is the load and the battery is the source, which is appealing because, in this example, power is transferred from the source (the battery) to the load (the headlight). (*Note:* It is not generally required nor necessarily true in all cases that power is transferred from source to load.)

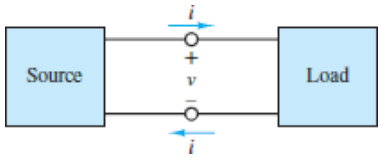


Figure 2.23 A circuit may be partitioned at two terminals. One part is the source; the other part is the load.

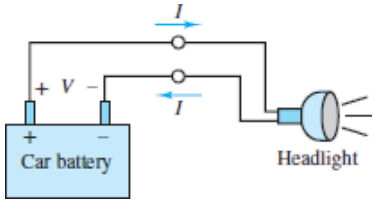


Figure 2.24 A headlight-battery system is a simple physical representation of a source-load perspective.

It is important not to confuse ideal independent and dependent voltage and current sources with the generalized source concept introduced here. Ideal sources, along with other circuit elements, are the constituents of a generalized source. In this book, ideal sources are referred to as either voltage or current sources, explicitly, to avoid confusion.

This source-load perspective leads to graphical solutions, which often lend insights to circuit behavior, and which can be essential in nonlinear problems Page 106involving diodes and transistors. Consider the two circuits shown in Figure 2.25(a) and (b). Each circuit is partitioned into a source and a load at terminals *A* and *B*. It is possible to analyze separately a source and its load and gain insight into the behavior of the complete circuit.

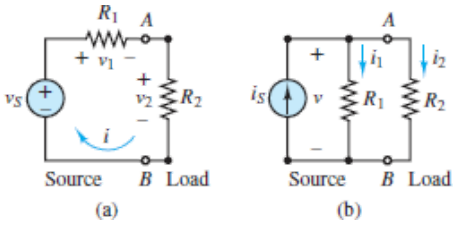


Figure 2.25 (a) A simple voltage divider partitioned into source and load. (b) A simple current divider partitioned into source and load.

The Source Networks

The source networks are shown in Figure 2.26(a) and (b) separate from their loads. These particular source networks are referred to in this book as *Thévenin* and *Norton* sources. KVL can be applied around a loop in Figure 2.26(a) and KCL can be applied at the upper node in Figure 2.26(b) to yield:

$$v_S = iR_1 + v \qquad i_S = i + \frac{v}{R_1} \tag{2.26}$$

These equations can be rearranged to yield:

$$i = \frac{v_S - v}{R_1} \qquad v = (i_S - i)R_1 \tag{2.27}$$

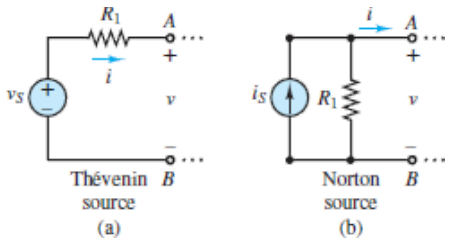


Figure 2.26 (a) Thévenin source, (b) Norton source.

Aside: As implied by the ellipses, these networks are parts of larger circuits and not stand-alone circuits themselves. Terminals *A* and *B* do not necessarily form an open-circuit, and so the current *i* is *not* assumed to be zero.

The *i*-*v* plots for each of these equations are shown in Figure 2.27(a) and (b).

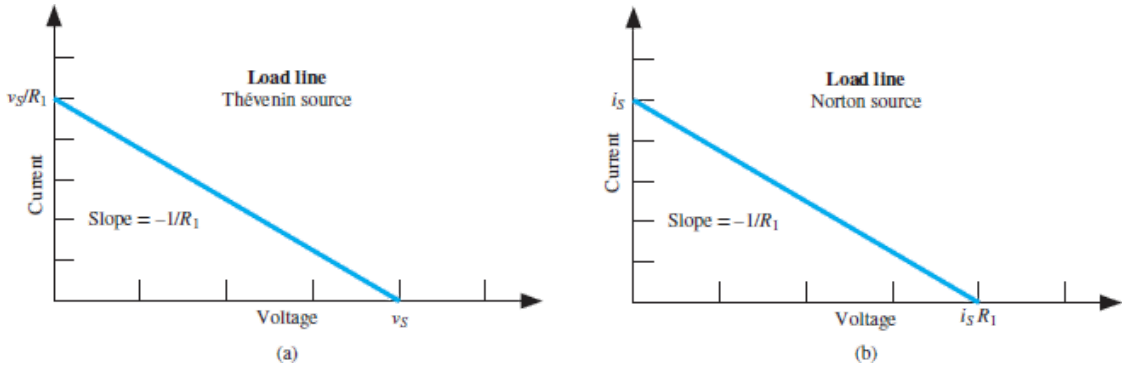


Figure 2.27 (a) The *i*-*v* plot for a Thévenin source; (b) the *i*-*v* plot for a Norton source

Notice that each plot is a straight line, known as the *load line,* with a (negative) slope of $-1/R_1$. Also, notice that when the terminals of the Thévenin Page 107and Norton source networks are connected by ideal wires (i.e., short-circuited) the resulting *short-circuit currents* i_{SC} through the wires are:

$$i_{SC} = \frac{v_S}{R_1} \qquad i_{SC} = i_S \tag{2.28}$$

Furthermore, when the terminals of the Thévenin and Norton sources are left open (i.e., open-circuited), the resulting *open-circuit voltages* v_{OC} across the terminals are:

$$v_{OC} = v_S \qquad v_{OC} = i_S R_1 \tag{2.29}$$

Interestingly, the short-circuit current and the open-circuit voltage are the two intercepts on the *i*-*v* plots and represent the source solution when the resistance

between terminals A and B is zero and infinity, respectively. All the other (i, v) points on the load line are solutions when the resistance between terminals A and B is nonzero but finite. As this resistance increases from zero the (i, v) solution moves along the load line from upper left to lower right.

The line representing all possible (i, v) solutions for the source network is known as the *load line*. The name implies that the solution for any resistive load from $0\Omega \rightarrow \infty$ is a point on the line.

It is important to note that if $v_S = i_S R_1$ the two load lines are identical and therefore the Thévenin and Norton source networks are *equivalent* from the perspective of the load. This result is a generalization of the concept of equivalent resistance.

The Load Network

The load for each of the two complete circuits can be represented by a simple resistor R_2, as shown in Figure 2.28. Its i-v relationship is simply Ohm's law.

$$v = iR_2 \qquad i = \frac{v}{R_2} \tag{2.30}$$

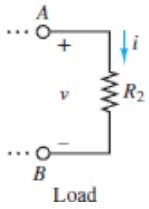


Figure 2.28 A generic load is represented by R_2.

The i-v plot is a simple straight line with a (positive) slope of $1/R_2$ and intercept at the origin. This line can be superimposed upon each of the source network plots. The result is shown in Figure 2.29(a) and (b).

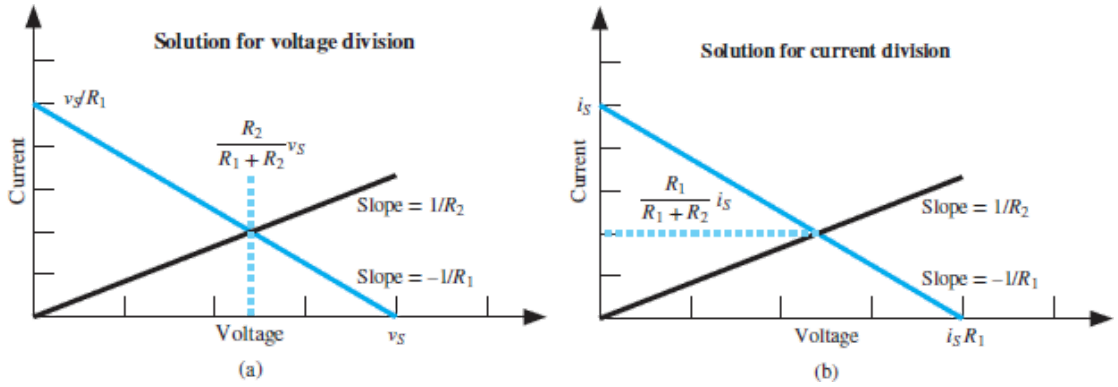


Figure 2.29 (a) For a Thévenin source, the intersection yields the voltage division result $v_2 = \frac{R_2}{R_1+R_2}v_S$. (b) For a Norton source, the intersection yields the current division result $i_2 = \frac{R_1}{R_1+R_2}i_s$.

Page 108

The intersection of the load line, which is comprised of the solutions for all possible resistive loads, with the line for a specific load yields the solution for that load. These plots allow one to envision the effect of either increasing or decreasing the load R_2 on the resulting voltage v across and current i through terminals A and B.

The algebraic solution for the intersection point is found by setting equal the source and load network solutions. (See equations 2.27 and 2.30.) The results are shown below on the left and right for the Thévenin and Norton sources, respectively, attached to a resistive load:

$$i = \frac{v_S - v}{R_1} \qquad v = (i_S - i)R_1 \tag{2.31}$$

$$= \frac{v}{R_2} \qquad = iR_2 \tag{2.32}$$

These equations can be rearranged to find:

$$\frac{v}{v_S} = \frac{R_2}{R_1+R_2} \qquad \frac{i}{i_S} = \frac{R_1}{R_1+R_2} \tag{2.33}$$

Voila! These are the voltage and current division expressions found earlier in this chapter for the same circuits as shown in Figure 2.25(a) and (b). This graphical method is particularly useful when the load is nonlinear, as is the case for a diode or transistor.

2.6 SOURCE TRANSFORMATIONS

It was stated in the previous section that the Thévenin and Norton sources shown in Figure 2.26 are equivalent if $v_S = i_S R_1$. In that case, the two load lines shown in Figure 2.27 are identical. This fact is the basis of an analytic tool known as *source transformations*.

Consider the Thévenin and Norton source networks shown in Figure 2.30, where the Thévenin source voltage is defined as v_T, the Norton source current is defined as i_N, and each source network has the same resistance $R_T = R_N$. Then, the two source networks are equivalent and can be interchanged if

$$v_T = i_N R_T = i_N R_N \tag{2.34}$$

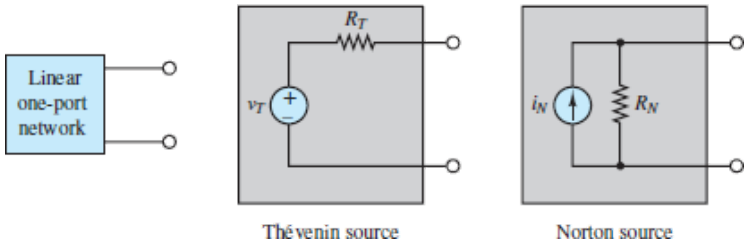


Figure 2.30 Simplified equivalent representations of a linear one-port network

Consider the upper network shown in Figure 2.31. The Thévenin source in the shaded box may be replaced by its equivalent Norton source as shown in the lower network. This replacement of a Thévenin source with a Norton source is a *source transformation*. Likewise, a Norton source may also be replaced by a Page 109Thévenin source. The computation of i_{SC} is now a straightforward application of current division.

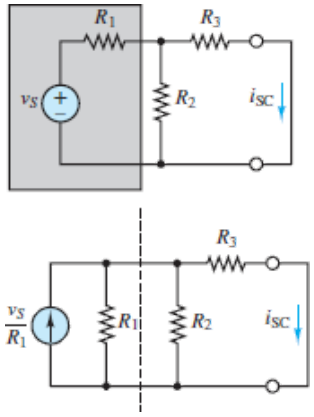


Figure 2.31 Result of source transformation

$$i_{SC} = \frac{1/R_3}{1/R_1 + 1/R_2 + 1/R_3} \frac{v_S}{R_1} \tag{2.35}$$

Source transformations can *simplify* a network, if employed correctly. To do so, identify the terminals of either a Thévenin or Norton source, as shown in Figure 2.32. Next, remove the source network between those terminals and reattach the equivalent source network to those same terminals. In general, a source transformation simplifies a circuit by creating either a series or parallel connection between two resistors that was not previously present in the circuit.

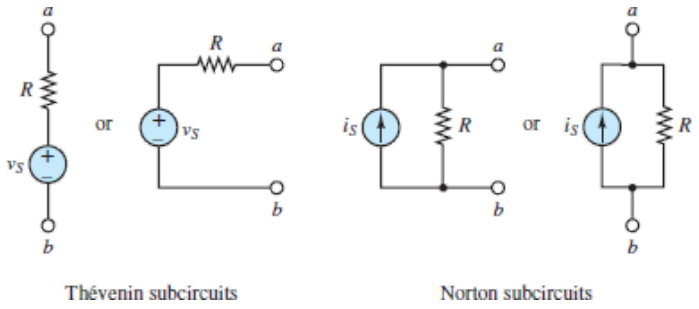


Figure 2.32 Networks amenable to source transformation

EXAMPLE 2.9 Source Transformations

Problem

Use source transformations to simplify the network shown in Figure 2.33 to a single Norton source seen by the load R_o.

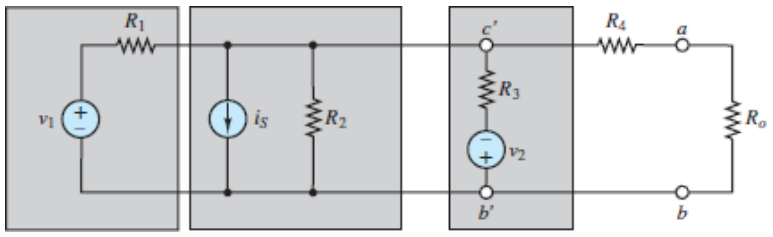


Figure 2.33

Solution

Known Quantities: Source voltages and current; resistor values.

Find: Thévenin equivalent resistance R_T; Norton current $i_N = i_{SC}$.

Schematics, Diagrams, Circuits, and Given Data: $v_1 = 50$ V; $i_S = 0.5$ A; $v_2 = 5$ V; $R_1 = 100\Omega$; $R_2 = 100\Omega$; $R_3 = 200\Omega$; $R_4 = 160\Omega$.

Assumptions: Assume the reference node is at the bottom of the circuit.

Page 110

Analysis: Highlight key terminals in the circuit to emphasize the Thévenin and Norton sources present in the circuit, as shown in Figures 2.33 and 2.34. The Thévenin source consisting of v_1 and R_1, which appears between terminals c'' and b'', can be replaced with a Norton source consisting of a current source v_1/R_1 in parallel with R_1. Similarly, the Thévenin source between terminals c' and b' can be replaced with a Norton source consisting of a current source v_2/R_3 in parallel with R_3. Both of these transformations are shown in Figure 2.35. Notice that the direction of the current source v_2/R_3 is in accord with the polarity of the voltage source v_2. The order of elements in parallel can be interchanged without changing the behavior of the overall circuit, as is shown in Figure 2.36(a) with numerical values inserted for each element.

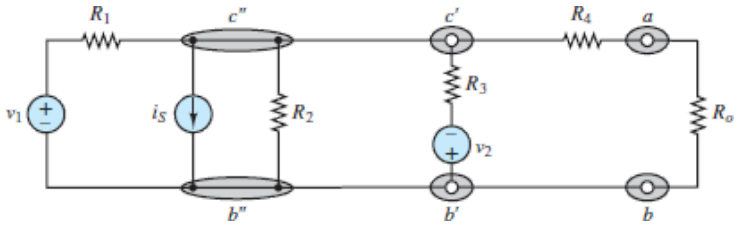


Figure 2.34

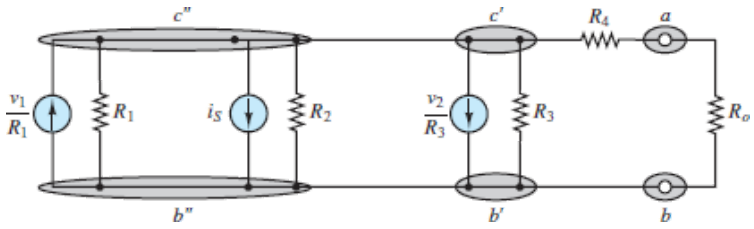


Figure 2.35

In Figure 2.36(a), the three current sources in parallel contribute a total current of 25 mA *leaving* the upper left node and so can be replaced with a single 25 mA current source directed downward as shown in Figure 2.36(b). Likewise, resistors R_1,

R_2 and R_3 are in parallel and have a 40Ω equivalent resistance between the upper left node and the bottom node.

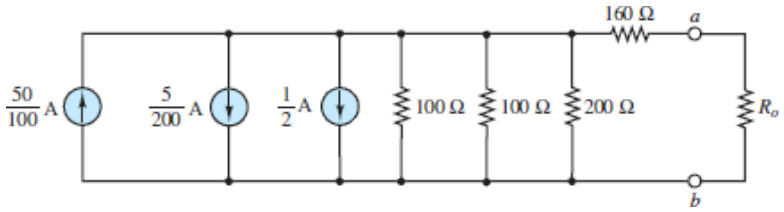


Figure 2.36(a) Transformed, but not yet simplified, circuit

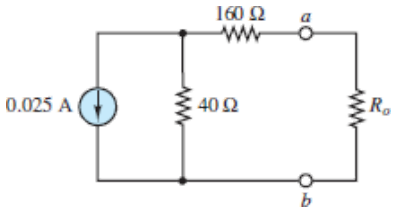


Figure 2.36(b) Simplified circuit

Page 111

The Norton source on the left can be transformed into a Thévenin source such that the 40-Ω resistor is in series with the 160-Ω resistor and the current source is transformed into a 0.025A×40Ω = 1-V voltage source. Finally, combine the two series resistors to form a single equivalent R_N = 200 Ω resistor and transform the resulting Thévenin source into an equivalent Norton source with a i_N = 5 mA current source directed downward, as shown in Figure 2.37.

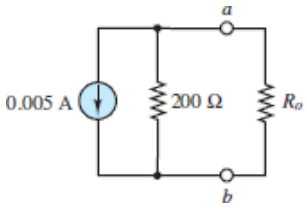


Figure 2.37

2.7 THÉVENIN AND NORTON EQUIVALENT NETWORKS

Recall that a network that has two, and only two, terminals at which it can attach to other networks, as in Figure 2.38, is a **one-port network**. Such a network is characterized by the relationship between the current i through and voltage v across its terminals for various loads (e.g., open-circuit, short-circuit). The key concepts are

- The impact of a one-port source on a one-port load is completely represented by the *i-v characteristic* of the source.
- Two one-port networks are *electrically equivalent* if their *i-v* characteristics are equivalent.
- Equivalent networks are those for which the voltage across and current through their terminals are the same *for any load.*

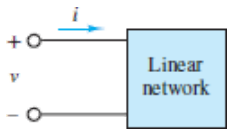


Figure 2.38 One-port network

This concept of *equivalence* was introduced earlier for a network of resistors. The central idea was that an entire network of resistors can be replaced by a single equivalent resistance without impacting whatever else is attached to the original network. Here, the concept of equivalence is generalized to networks that include resistors, ideal sources, and other linear circuit elements. Recall the following statement from earlier in this chapter.

Any network composed of linear elements only is itself linear. Common linear elements are ideal sources, resistors, capacitors, and inductors.

The essence of this section is captured in the statement of two very important theorems about linear networks.

Thévenin's Theorem

When viewed from its terminals, any linear one-port network may be represented by an equivalent circuit consisting of an ideal voltage source v_T in series with an equivalent resistance R_T.

Norton's Theorem

When viewed from its terminals, any linear one-port network may be represented by an equivalent circuit consisting of an ideal current source i_N in parallel with an equivalent resistance R_N.

Page 112

The Thévenin and Norton equivalent resistances (R_T and R_N) are the same for any particular linear one-port network.

Any one-port linear network, no matter how complicated, can always be represented by either of two simple equivalent networks, and the transformations leading to these equivalent representations are easily managed, with a little practice. In this section, techniques are presented for computing these equivalent networks, which reveal some simple—yet general—results for linear networks, and are useful for analyzing basic nonlinear circuits.

FOCUS ON PROBLEM SOLVING

THÉVENIN AND NORTON THEOREMS

Any one-port linear network can be simplified to either of two equivalent netv
forms. They are:

- A *Thévenin source*, comprised of an independent voltage source v_T in s with a resistor R_T, as shown in Figure 2.39.
- A *Norton source*, comprised of an independent current source i_N in parallel a resistor R_N, as shown in Figure 2.40.

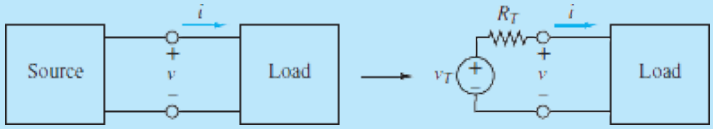


Figure 2.39 Illustration of Thévenin's theorem

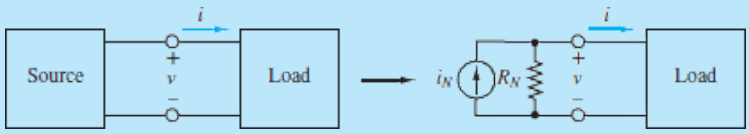


Figure 2.40 Illustration of Norton's theorem

Moreover, since these equivalent network forms are equivalent to the original li network, the forms themselves are equivalent. For example, a Thévenin source be interchanged with its equivalent Norton source, which is known as a *so transformation*.

The equivalent network of any specific one-port linear network is comprise specific values for v_T and R_T, or i_N and R_N, which are known as the:

- Thévenin voltage v_T and Thévenin equivalent resistance R_T.
- Norton current i_N and the Norton equivalent resistance R_N.

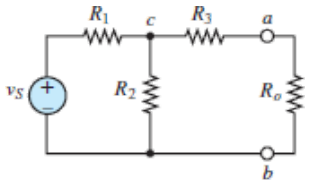
In addition, for any specific linear one-port network $R_T = R_N$ and $v_T = i_N R_T$. values of v_T and i_N are the open-circuit voltage v_{OC} across and the short-ci current i_{SC} through, respectively, the source network terminals.

Page 113

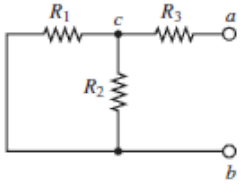
Computation of R_T or R_N: Networks Without Dependent Sources

The first step to calculate the Thévenin (or Norton) equivalent resistance of a one-port linear source network with no dependent sources is to identify the two terminals

(e.g., *a* and *b*) of the source network. Sometimes just the one-port source network is given in a problem, in which case the network terminals should be readily apparent. Other times a complete circuit is given such that it is necessary to define and/or identify the load and, by default, the source network. In Figure 2.41, the resistor R_o is chosen as the load such that terminals *a* and *b* define the one-port (two terminal) connection between the load and the source network.



Complete circuit



Circuit with load removed for computation of R_T. The voltage source is replaced by a short circuit.

Figure 2.41 Computation of Thévenin resistance

The second step is to remove the load and set all independent sources in the source network to zero; that is, replace all independent voltage sources with short-circuits and all independent current sources with open-circuits. The source network of Figure 2.41 is shown with the voltage source replaced by a short-circuit.

Finally, apply series and parallel equivalent resistance substitutions to find the effective equivalent resistance "seen" by the load R_o across terminals *a* and *b*. For example, in the circuit of Figure 2.42, R_1 and R_2 are in parallel since they are connected between the same two nodes, *b* and *c*. The total resistance between terminals *a* and *b* is simply:

$$R_T = R_3 + R_1 \parallel R_2 \tag{2.36}$$

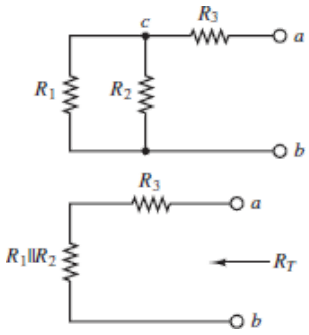


Figure 2.42 Equivalent resistance seen by the load

When series and parallel equivalent resistance substitutions are not sufficient, find R_T by once again turning off (i.e., setting to zero) all independent sources in the one-port network and attaching an arbitrary independent voltage source v to the network terminals. Next, the current i through those terminals can be computed. Then, R_T is simply:

$$R_T = \frac{v}{i} \tag{2.37}$$

For example, assume that the independent sources in a one-port network have been turned off and the resulting network is that shown in the top portion of Figure 2.43. That resistor network, seen by terminals a and b, cannot be simplified by series and parallel equivalent resistance substitutions. However, the Thévenin equivalent resistance to the left of terminals a and b can be computed by applying an arbitrary independent voltage source v as shown in the bottom portion of Figure 2.43 and applying Equation 2.37. This method is described by a simple four-step algorithm.

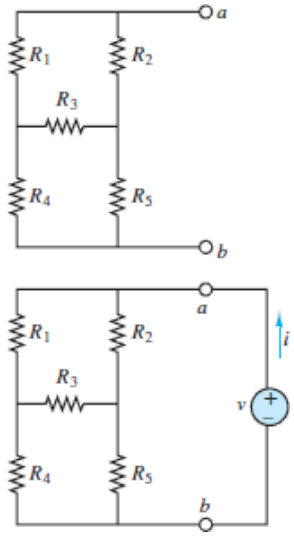


Figure 2.43 A general method of determining the Thévenin resistance

Step 1: *Set all independent voltage and current sources in the one-port network to zero. Replace them with short- and open-circuits, respectively.*

Step 2: *Attach an arbitrary independent voltage source* v_S *across the one-port network terminals.*

Step 3: *Compute the current* i_S *through the voltage source.*

Step 4: *Compute* $R_T = v_S/i_S$.

Computation of R_T or R_N: Networks With Dependent Sources

There are two methods available to calculate the Thévenin equivalent resistance R_T when a dependent source is present in a linear one-port network. The first method is the same as that described by the four-step algorithm at the end of the previous section. That method will always work.

Page 114

A second method can be applied when the network also contains at least one independent source. Without turning off any of the network's independent sources compute the open-circuit voltage v_{oc} across the network terminals and then the short-circuit current i_{sc} through the same terminals. Then, the Thévenin equivalent resistance R_T of the network is

$$R_T = \frac{v_{oc}}{i_{sc}} \tag{2.38}$$

For either of these two methods to be valid the following rule must be obeyed.

> Each dependent source and its associated dependent variable must be collocated in either the source network or the load when applying Thévenin's or Norton's theorem.

For any particular one-port network the Thévenin and Norton equivalent resistances are always equivalent to each other.

$$R_T = R_N \tag{2.39}$$

As a result, often only the R_T notation is used.

FOCUS ON PROBLEM SOLVING

THÉVENIN RESISTANCE

Use the following steps to compute the Thévenin equivalent resistance across terminals of a linear linear one-port network.

1. Identify the one-port network and label its terminals *a* and *b*.
2. Two methods exist for one-port networks without dependent sources. Both b with the same first step.
   (a) Turn off all independent voltage and current sources in the network replace them with short- and open-circuits, respectively.
   (b) When possible, use series and parallel equivalent resistance substitution simplify the network and eventually find R_T.
   (c) When series and parallel equivalent resistance substitutions are not suffic attach an arbitrary voltage source v_s to the terminals and compute the resu current i_s through those terminals. The Thévenin equivalent resistance is v_s/i_s.
3. Two methods also exist for one-port networks with dependent sources. The method is to follow steps 2(a) and 2(c). The second method is
   (a) Leave all independent sources in the network turned on.
   (b) Compute the open-circuit voltage v_{oc} across the network terminals.
   (c) Compute the short-circuit current i_{sc} through the network terminals.
   (d) Compute $R_T = v_{oc}/i_{sc}$.

When a dependent source is present in a one-port network, its associated dependent variable must also be part of that network.

EXAMPLE 2.10 Computing R_T for a Network Without a Dependent Source

Problem

Compute the Thévenin equivalent resistance seen by the load R_o in Figure 2.44.

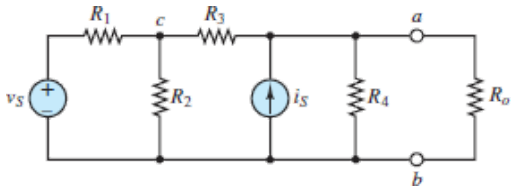


Figure 2.44

Solution

Known Quantities: Resistor values.

Find: The Thévenin equivalent resistance R_T.

Schematics, Diagrams, Circuits, and Given Data: $v_S = 5$ V; $R_1 = 2\Omega$; $R_2 = 2\Omega$; $R_3 = 1\Omega$; $i_S = 1$ A; $R_4 = 2\Omega$.

Analysis: Refer to Figure 2.44 and the steps in the Focus on Problem Solving box "Thévenin Resistance."

1. The source network is everything to the left of terminals a and b.
2. Turn off the voltage and current sources and replace them with short- and open-circuits, respectively. The result is shown in Figure 2.45.
3. There are three nodes remaining in the source network and no dependent source. Clearly, R_1 and R_2 are in parallel since they sit between nodes c and b. Their parallel equivalent resistance is in series with R_3. Thus, there are two parallel resistances from $a \rightarrow b$: $R_3 + (R_1 \mathbin{/\!/} R_2)$ and R_4. Finally, the equivalent resistance from $a \rightarrow b$ is

$$R_T = [(R_1 \parallel R_2) + R_3] \parallel R_4$$
$$= [(2 \parallel 2) + 1] \parallel 2 = 1\ \Omega$$

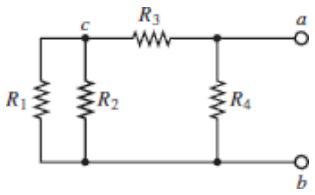


Figure 2.45

Comments: The network in this example is drawn in a fairly uncomplicated manner. However, sometimes a network may be drawn in a more confusing manner. In any case, it is easy to correctly calculate the equivalent resistance between the network terminals by seeing the network as a collection of nodes, between which sit various elements.

Page 116

EXAMPLE 2.11 Computing R_T for a Network With a Dependent Source

Problem

Compute the Thévenin equivalent resistance seen by the load R_o in Figure 2.46.

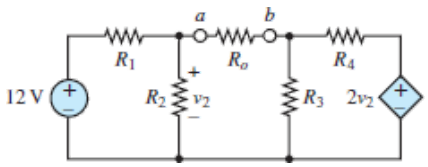


Figure 2.46

Solution

Known Quantities: Source and resistor values.

Find: The Thévenin equivalent resistance R_T seen by the load R_o.

Schematics, Diagrams, Circuits, and Given Data: R_1 = 24 kΩ; R_2 = 8 kΩ; R_3 = 9 kΩ; R_4 = 18 kΩ.

Analysis: Refer to Figure 2.46 and the steps in the Focus on Problem Solving box "Thévenin Resistance."

1. The source network is everything between terminals *a* and *b*.
2. Turn off the independent voltage source in Figure 2.46 and replace it with a short-circuit. As a result, R_1 and R_2 are in parallel and can be replaced by a single equivalent resistance.
3. The source network contains a dependent source. Attach an arbitrary independent voltage source v_S across terminals *a* and *b* and label its current i_S as shown in Figure 2.47. Included in this figure are two mesh currents i_1 and i_2 that can be used to solve the circuit using mesh analysis.

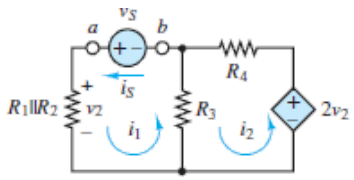


Figure 2.47

The counterclockwise circulation was chosen so that $i_1 = i_S$. Apply KVL around each mesh:

$$v_S - (R_1 \parallel R_2)i_1 - R_3(i_1 - i_2) = 0 \qquad \text{mesh 1}$$
$$2v_2 - R_4 i_2 - R_3(i_2 - i_1) = 0 \qquad \text{mesh 2}$$

Note that $v_2 = i_1(R_1 \mathbin{/\!/} R_2)$ such that the equations can be rewritten as:

$$v_S - (R_1 \parallel R_2)i_1 - R_3(i_1 - i_2) = 0 \qquad \text{mesh 1}$$
$$2(R_1 \parallel R_2)i_1 - R_4 i_2 - R_3(i_2 - i_1) = 0 \qquad \text{mesh 2}$$

Page 117

Collect coefficients of i_1 and i_2 and substitute values for the resistors.

$$15\,i_1 - 9\,i_2 = v_S \qquad \text{mesh 1}$$
$$21\,i_1 - 27\,i_2 = 0 \qquad \text{mesh 2}$$

Divide both sides of the mesh 2 equation by 3 and subtract the result from the mesh 1 equation.

$$8\,i_1 = v_S \qquad \text{or} \qquad \frac{v_S}{i_1} = \frac{v_S}{i_S} = R_T = 8\,\text{k}\Omega$$

Comments: This result can be computed by an alternate method where the 12-V source is left on. First, remove the load and compute the open-circuit voltage v_{oc} across terminals *a* and *b*. Second, connect terminals *a* and *b* with a wire and compute the short-circuit current i_{SC} through that wire. Last, compute R_T from its definition:

$$R_T \equiv \frac{v_{OC}}{i_{SC}}$$

Try it. Does it work?

Computing the Thévenin Voltage

This section describes the computation of the Thévenin voltage v_T for an arbitrary linear one-port network that may contain sources, both independent and dependent, and linear resistors. Thévenin's theorem states that any linear one-port network may be simplified to an equivalent network consisting of an independent voltage source v_T in series with a resistor R_T, as shown in Figure 2.48. When the network terminals are open the current i is necessarily zero such that the voltage drop across R_T is also zero. In this open-circuit case, KVL requires

$$v_T = iR_T + v_{OC} = v_{OC} \tag{2.40}$$

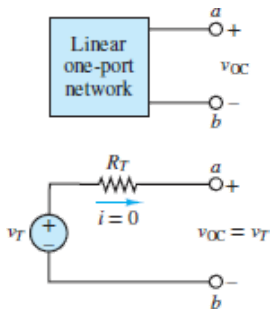


Figure 2.48 Equivalence of open-circuit and Thévenin voltage

LO

The Thévenin voltage v_T is equal to the **open-circuit voltage** v_{oc} across the terminals of a linear one-port network.

The following simple algorithm can be used to solve for the Thévenin voltage.

FOCUS ON PROBLEM SOLVING

THÉVENIN VOLTAGE

Follow these steps to compute the Thévenin voltage for a linear one-port network

1. Identify the network and label its terminals (e.g., a and b).
2. Define the open-circuit voltage v_{oc} across those terminals.
3. Apply any preferred method (e.g., the node voltage method) to solve for v_{oc}.
   - For a network without an independent source, the open-circuit voltage v simply zero, even when a dependent source is present.
4. The Thévenin voltage v_T of the network is, by definition, v_{oc}.

Page 118

Computation of the open-circuit voltage is best illustrated by examples. For instance, assume that everything to the left of terminals a and b in Figure 2.49 is considered the one-port source network attached to a load R_o. The Thévenin equivalent resistance from $a \rightarrow c \rightarrow b$ of the source network is $R_T = R_3 + R_1 \,//\, R_2$.

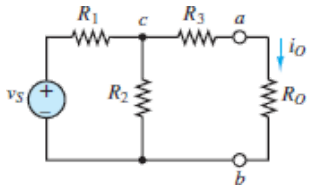


Figure 2.49

To compute v_{oc}, remove the load R_o, as shown in Figure 2.50, and observe that the current through R_3 must then be zero. Thus, R_1 and R_2 are in a *virtual* series connection and, as illustrated in Figure 2.51, v_{oc} is equal to the voltage across R_2, which can be found by voltage division in the *virtual* series loop $v_S \rightarrow R_1 \rightarrow R_2 \rightarrow v_S$.

$$v_T = v_{OC} = v_S \frac{R_2}{R_1 + R_2}$$

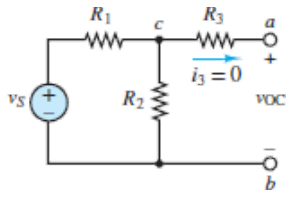


Figure 2.50

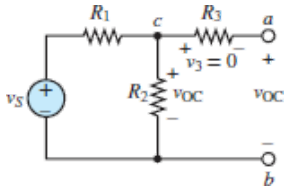


Figure 2.51

Now, consider, side by side, the original circuit and the circuit with the source network replaced by its Thévenin equivalent, as shown in Figure 2.52. The current i_o through the load R_o must be the same in both circuits.

$$i_o = v_T \cdot \frac{1}{R_T + R_o} = v_S \frac{R_2}{R_1 + R_2} \cdot \frac{1}{(R_3 + R_1 \parallel R_2) + R_o} \tag{2.41}$$

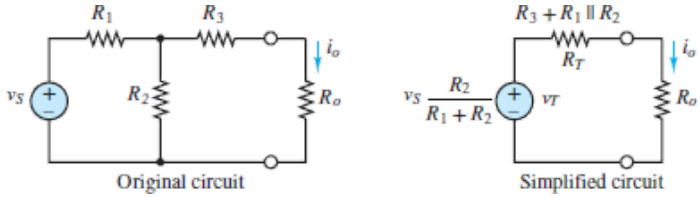


Figure 2.52 Two circuits with equivalent source networks for the load R_o

Notice that the latter portion of this expression is rather complicated. However, if you focus on the source network alone, it is often possible, with some practice, to readily determine R_T and v_T, by observation, and then apply Ohm's law or voltage division to the simplified circuit to find the current through or voltage across R_o. Practice! Practice!! Practice!!!

It is possible for v_T to be zero. In such a case, R_T may still be nonzero even though R_T is defined by $v_T = i_N R_T$. The implication is that when v_T is zero, i_N may also be zero, and vice versa, allowing finite, nonzero values for R_T. In this case the Thévenin equivalent of the source network is a simple resistor R_T. There are two other exceptional cases:

1. When v_T and R_T are both zero, i_N can be any value. Such a source network is trivial, being equivalent to a short-circuit.
2. When i_N is zero and R_T is infinitely large, v_T can be any value. Such a source network is also trivial, being equivalent to an open-circuit.

Page 119

EXAMPLE 2.12 Computing v_T for a Network With One Independent Source

Problem

Compute the open-circuit voltage v_{oc} in the network shown in Figure 2.53.

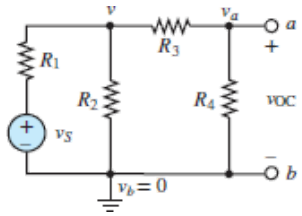


Figure 2.53

Solution

Known Quantities: Source voltage, resistor values.

Find: Open-circuit voltage v_{oc}.

Schematics, Diagrams, Circuits, and Given Data: $v_S = 12$ V; $R_1 = 1\Omega$;$R_2 = 10\Omega$; $R_3 = 10\Omega$; $R_4 = 20\Omega$.

Analysis: Refer to Figure 2.53 and the steps in the Focus on Problem Solving box "Thévenin Voltage."

1. In this problem, the one-port network is everything to the left of terminals *a* and *b*.
2. The open-circuit voltage v_{oc} is across terminals *a* and *b*, as shown in the figure.
3. There are four nodes in the network. Node *b* is selected as the reference with a voltage $v_b = 0$. Another node is fixed at v_S by the voltage source. For the other two nodes, the node voltage method will yield two KCL equations in the two

unknown node voltages, v and v_a. Apply KCL to obtain the following two equations:

$$\frac{v_S - v}{R_1} - \frac{v - 0}{R_2} - \frac{v - v_a}{R_3} = 0 \qquad \text{node } v$$

$$\frac{v - v_a}{R_3} - \frac{v_a - 0}{R_4} = 0 \qquad \text{node } v_a$$

Collect terms to find:

$$\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) v \qquad - \frac{1}{R_3} v_a = \frac{v_S}{R_1} \qquad \text{node } v$$

$$\frac{1}{R_3} v - \left(\frac{1}{R_3} + \frac{1}{R_4}\right) v_a = 0 \qquad \text{node } v_a$$

Substitute numerical values and write the equations in matrix form as:

$$\begin{bmatrix} 1.2 & -0.1 \\ 0.1 & -0.15 \end{bmatrix} \begin{bmatrix} v \\ v_a \end{bmatrix} = \begin{bmatrix} 12 \\ 0 \end{bmatrix}$$

Solving yields $v = 10.6$ V and $v_a = 7.1$ V. Thus, $v_{OC} = v_a - 0 = 7.1$ V.

Comments: A common mistake in problems like this one is to assume that R_4 is the load (even though there is no mention of a load) and not part of the one-port source network. The fact that the voltage drop across R_4 is given as the open-circuit voltage v_{oc} suggests that the entire network to the left of terminals a and b is to be treated as the source network.

Page 120

EXAMPLE 2.13 Computing v_T and R_T for a Network With Two Independent Sources

Problem

Find the Thévenin equivalent of the source network and use it to compute the load current i in the circuit of Figure 2.54.

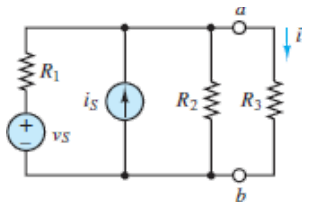


Figure 2.54

Solution

Known Quantities: Source and resistor values.

Find: v_T and R_T for the source network and the load current i.

Schematics, Diagrams, Circuits, and Given Data: $v_S = 24$ V; $i_S = 3$ A; $R_1 = 4\Omega$; $R_2 = 12\Omega$; $R_3 = 6\Omega$.

Analysis: Refer to Figure 2.54 and the steps in the Focus on Problem Solving boxes "Thévenin Resistance" and "Thévenin Voltage."

1. R_3 is the load. Everything else is the one-port source network.
2. Remove the load R_3 and solve for R_T and v_T of the source network, and use them to find the load current i.
   - Find R_T: Set both voltage and current sources to zero and replace them with short- and open-circuits, respectively, as shown in Figure 2.55. The resulting equivalent resistance between terminals a and b is simply $R_T = R_1 \,/\!/\, R_2 = 4 \,/\!/\, 12 = 3\Omega$.
   - Find v_T: The circuit shown in Figure 2.56 has only three nodes. Node b is selected as the reference with a voltage $v_b = 0$. Of the remaining two nodes, one is fixed at a voltage v_S by the voltage source. Thus, only a single KCL equation at node v_a is needed for a solution:

$$\frac{v_S - v_a}{R_1} + i_S - \frac{v_a}{R_2} = 0 \quad \text{or} \quad v_a = (v_S + i_S R_1)\frac{R_2}{R_1 + R_2}$$

     Substitute numerical values to find: $v_a - v_b = v_{OC} = 27$V. The Thévenin voltage v_T is the open-circuit voltage v_{OC} across terminals a and b. (Note that the principle of superposition also readily yields a solution.)
   - Find i: Construct the Thévenin equivalent of the source network and reattach the load R_3, as shown in Figure 2.57. The load current is easily computed using Ohm's law.

$$i = \frac{27}{3+6} = 3 \text{ A}$$

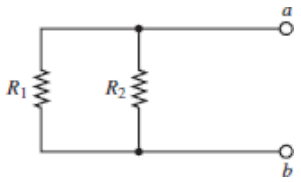


Figure 2.55

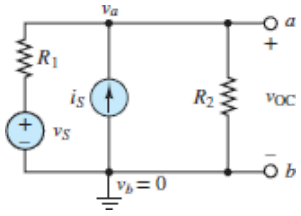


Figure 2.56

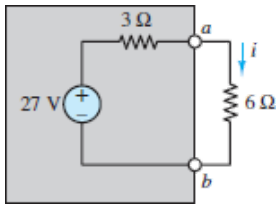


Figure 2.57 Simplified circuit

Comments: Equivalent circuit analysis has several key advantages. By reducing any complicated linear source network to a simple structure, one can more quickly determine:

- The voltage across and current through *any* load.
- The maximum possible load current v_T/R_T (for loads approaching short-circuits).
- The maximum possible load voltage v_T (for loads approaching open-circuits).
- The value of the load that gives maximum power transfer to the load (see Section 2.8).

EXAMPLE 2.14 Computing v_T for a Network With a Dependent Source

Problem

Find the Thévenin voltage v_T of the one-port source network seen by the load R_o in Figure 2.58.

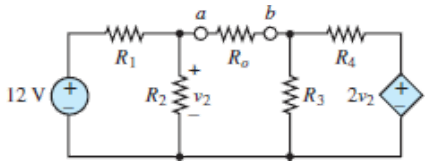


Figure 2.58

Solution

Known Quantities: Source and resistor values.

Find: v_T for the source network.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 24$ kΩ; $R_2 = 8$ kΩ; $R_3 = 9$ kΩ; $R_4 = 18$ kΩ.

Analysis: Refer to Figure 2.58 and the steps in the Focus on Problem Solving box "Thévenin Voltage." This circuit is identical to the one from Example 2.11, where the Thévenin equivalent resistance R_T seen by R_o was found to be 8 kΩ. In this example, the Thévenin voltage v_T seen by R_o is found.

1. The source network is everything except the load R_o. Remove the load from the source network as shown in Figure 2.59.
2. Define the open-circuit voltage v_{OC} as in Figure 2.59.
3. The resulting circuit has two series loops sharing one common node c. Define the voltage v_3 across R_3. Then KVL around the middle portion of the circuit yields:

$$v_2 = v_{OC} + v_3$$

Voltage division can be applied to the series loop on the left to solve for v_2.

$$v_2 = 12\text{V}\frac{R_2}{R_1 + R_2} = 12\text{V}\frac{8}{24 + 8} = 3\text{V}$$

Voltage division can also be applied to the series loop on the right to find v_3 in terms of v_2.

$$v_3 = 2v_2\frac{R_3}{R_3 + R_4} = 6\,\text{V}\frac{9}{9 + 18} = 2\,\text{V}$$

Plug these values into the KVL equation to find:

Page 122

4. The Thévenin voltage is $v_T = v_{OC} = 1$ V.

$$v_{OC} = v_2 - v_3 = 1\,\text{V}$$

Figure 2.59

Computing the Norton Current

The Norton current i_N is equal to the **short-circuit current** i_{SC} through the source network terminals.

Consider an arbitrary linear one-port network and its Norton equivalent, each attached to a short-circuit, as shown in Figure 2.60. The voltage across R_N is zero so there is no current through it. Thus, the current i_{SC} through the short-circuit is exactly the Norton current i_N.

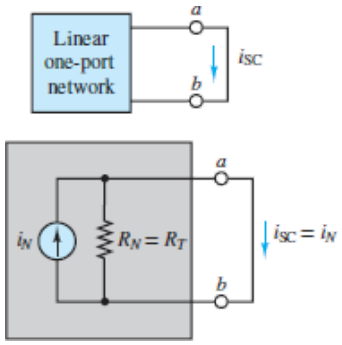


Figure 2.60 Illustration of Norton equivalent circuit

FOCUS ON PROBLEM SOLVING

NORTON CURRENT

Follow these steps to compute the Norton current for a linear one-port network.

1. Identify the one-port network and label its terminals (e.g., *a* and *b*).
2. Define the short-circuit current i_{SC} through those terminals.
3. Apply any preferred method (e.g., the mesh current method) to solve for i_{SC}.
   - For source networks without an independent source the short-circuit cu i_{SC} is simply zero, even when a dependent source is present.
4. The Norton current i_N of the network is, by definition, i_{SC}.

This simple observation suggests the basic method for finding the Norton current for any arbitrary linear one-port network. Attach a short-circuit wire to its terminals to determine the Norton current through the wire.

Consider the circuit of Figure 2.61, shown with a short-circuit attached to a one-port source network (i.e., in place of the load). The short-circuit current i_{SC} can be found easily using the mesh current method.

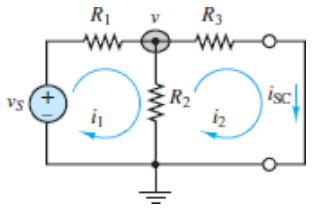


Figure 2.61 Computation of Norton current

In terms of the mesh currents i_1 and i_2, the KVL mesh equations are

$$\begin{aligned} v_S - R_1 i_1 - R_2(i_1 - i_2) = 0 \quad & \text{mesh 1} \\ -R_2(i_2 - i_1) - R_3 i_2 = 0 \quad & \text{mesh 2} \end{aligned}$$

Collect terms to find:

$$(R_1+R_2)i_1 - R_2 i_2 = v_S \quad \text{mesh 1}$$
$$-R_2 i_1 + (R_2+R_3)i_2 = 0 \quad \text{mesh 2}$$

Page 123

Multiply the mesh 2 equation by $(R_1 + R2)/R_2$ and add the result to the mesh 1 equation to find:

$$\left[\frac{(R_1+R_2)(R_2+R_3)}{R_2} - R_2\right] i_2 = v_S$$

Finally, multiply both sides of the equation by R_2 to obtain:

$$i_{SC} = i_2 = \frac{v_S R_2}{(R_1+R_2)(R_2+R_3) - R_2^2} = \frac{v_S R_2}{R_1R_2 + R_1R_3 + R_2R_3}$$

Alternatively, the node voltage method requires only one KCL node equation.

$$\frac{v_S - v}{R_1} = \frac{v}{R_2} + \frac{v}{R_3}$$

Multiply both sides of the equation by $R_1R_2R_3$ and collect terms to find:

$$v_S R_2 R_3 = v(R_1R_2 + R_1R_3 + R_2R_3)$$

or

$$v = v_S \frac{R_2R_3}{R_1R_2 + R_1R_3 + R_2R_3}$$

Finally, the short-current current is:

$$i_{SC} = \frac{v - 0}{R_3} = \frac{v_S R_2}{R_1R_2 + R_1R_3 + R_2R_3}$$

Of course, the results are the same for both methods. Great! Thus, the Norton current i_N is:

$$i_N = i_{SC} = \frac{v_S R_2}{R_1R_2 + R_1R_3 + R_2R_3}$$

Why solve for i_{SC} twice, using two separate methods? When time allows, it is always a good idea to validate your results!

EXAMPLE 2.15 Computing i_N for a Network With Two Independent Sources

Problem

Determine the Norton current i_N and the Norton equivalent for the network in Figure 2.62.

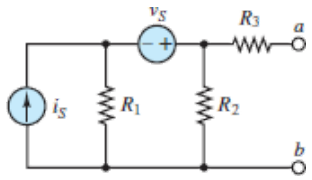


Figure 2.62

Solution

Known Quantities: Voltage source v_S and current source i_S; resistor values.

Find: Norton current $i_N = i_{SC}$; Equivalent resistance R_T.

Schematics, Diagrams, Circuits, and Given Data: $v_S = 6$ V; $i_S = 2$ A; $R_1 = 6\Omega$; $R_2 = 3\Omega$; $R_3 = 2\Omega$.

Assumptions: Assume the reference node is at the bottom of the circuit.

Page 124

Analysis: Refer to Figure 2.62 and the steps in the Focus on Problem Solving box "Norton Current."

- Find i_N: In Figure 2.63, the source network terminals a and b are defined and a short-circuit wire is attached to them. The mesh current method would work very well in this problem (do you see why?), but the node voltage method will also work, and this circuit provides a good opportunity to practice the use of a "supernode."

  (a) There are three nodes in this circuit. The reference node is labeled in the figure.

  (b) There are two nonreference nodes labeled v_1 and v_2.

  (c) The two-node voltage variables are related to each other by the voltage source v_S.

(d) Apply KCL at the boundaries of the supernode shown in the figure to find:

$$i_S - \frac{v_1 - 0}{R_1} - \frac{v_2 - 0}{R_2} - \frac{v_2 - 0}{R_3} = 0 \qquad \text{supernode}$$

$$v_2 - v_1 = v_S \qquad \text{constraint equation}$$

(e) v_2 is the primary objective since $v_2 - 0 = i_{SC}R_3$. Use the constraint equation to substitute for v_1 in the supernode equation.

$$i_S = \frac{v_2 - v_S}{R_1} + v_2\frac{R_2 + R_3}{R_2R_3}$$

$$i_S + \frac{v_S}{R_1} = v_2\left[\frac{1}{R_1} + \frac{R_2 + R_3}{R_2R_3}\right]$$

Form the common denominator $R_1R_2R_3$ for the bracketed term and find:

$$v_2 = \left(i_S + \frac{v_S}{R_1}\right)\left[\frac{R_1R_2R_3}{R_1R_2 + R_1R_3 + R_2R_3}\right]$$

$$= \left(2 + \frac{6}{6}\right)\left[\frac{6 \cdot 3 \cdot 2}{6 \cdot 3 + 6 \cdot 2 + 3 \cdot 2}\right]$$

$$= (2 + 1)\left[\frac{36}{36}\right] = 3\,\text{V}$$

(f) Finally, the short-circuit current is given by:

$$i_{SC} = \frac{v_2}{R_3} = \frac{3}{2} = 1.5\,\text{A} = i_N$$

- Find R_T: To compute the Thévenin equivalent resistance, set the independent voltage and current sources to zero and replace them with short- and open-circuits, respectively. The resulting resistor network is shown in Figure 2.64. It is easy to see that $R_T = R_1 \;//\; R_2 + R_3 = 6 \;//\; 3 + 2 = 4\Omega$.

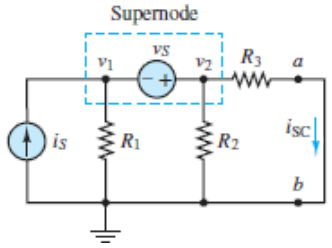


Figure 2.63

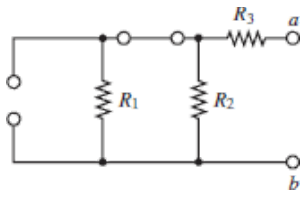


Figure 2.64

The Norton equivalent of the original one-port network is shown in Figure 2.65. Notice the polarity of the current source, which is dictated by the polarity defined for the short-circuit current i_{SC}.

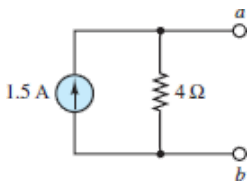


Figure 2.65 Norton equivalent network

Comments: Superposition is a reasonable alternative method for solving for i_{SC}. Take another look at Figure 2.63 and note that current division will quickly yield the component of i_{SC} due to the current source i_S. Also note that voltage division will quickly yield the component v_2 due to the voltage source v_S. Ohm's law can then be applied to find i_{SC}.

EXAMPLE 2.16 Computing i_N for a Network With a Dependent Source

Problem

Find the Norton current i_N of the one-port source network seen by the load R_o in Figure 2.66.

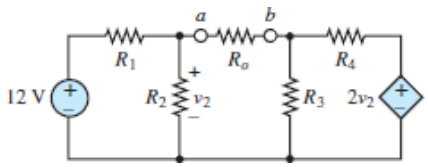


Figure 2.66

Solution

Known Quantities: Voltage source and resistor values.

Find: i_N for the source network.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 24\ \text{k}\Omega$; $R_2 = 8\ \text{k}\Omega$; $R_3 = 9\ \text{k}\Omega$; $R_4 = 18\ \text{k}\Omega$.

Assumptions: Assume the reference node is at the bottom of the circuit.

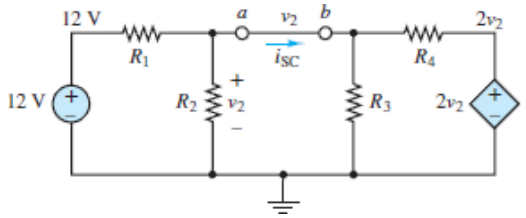


Figure 2.67

Analysis: Refer to Figure 2.66 and the steps in the Focus on Problem Solving box "Norton Current." This circuit is identical to the one from Example 2.14, where the Thévenin voltage seen by R_o was found to be 1 V. In this example, the Norton current i_N seen by R_o is found.

1. The one-port source network is everything except the load R_o. Remove the load from the source network and replace it with a short-circuit (i.e., a wire), as shown in Figure 2.67.
2. Define the short-circuit current i_{SC} as in Figure 2.67.
3. The resulting circuit has three nonreference nodes; however, the voltage of one is known while the voltages of the other two are both determined by v_2. Thus, there is only one unknown node voltage v_2, which can be found by applying KCL at node v_2.

$$\frac{12 - v_2}{R_1} - \frac{v_2 - 0}{R_2} - \frac{v_2 - 0}{R_3} - \frac{v_2 - 2v_2}{R_4} = 0$$

Plug in values for the resistors and multiply both sides of the equation by the common denominator to get:

$$3(12 - v_2) - 9v_2 - 8v_2 - 4(-v_2) = 0$$

or

$$16v_2 = 36 \qquad \text{which yields} \qquad v_2 = \frac{9}{4} = 2.25\,\text{V}$$

Page 126

4. To find i_{SC} apply KCL at the wire junction directly above R_2.

$$\frac{12 - v_2}{R_1} - \frac{v_2 - 0}{R_2} - i_{SC} = 0$$

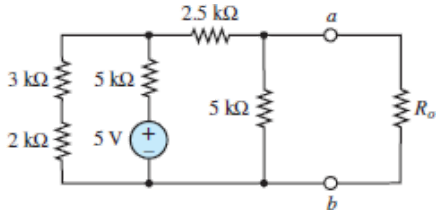
Plug in for v_2 to find

$$i_{SC} = \frac{9.75}{24} - \frac{2.25}{8} = \frac{3}{24} = \frac{1}{8} = 0.125\,\text{mA} = i_N$$

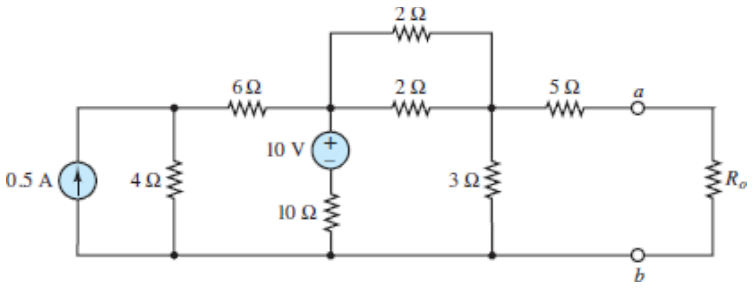
Comments: Note that the circuit in this example is identical to the one used in Examples 2.11 and 2.14. In these three example problems, the Thévenin equivalent resistance R_T, the Thévenin voltage v_T, and the Norton current i_N were found to be 8 kΩ, 1 V, and 0.125 mA, respectively, for the same linear one-port network. Although all three values were found by independent means, the result is that $v_T = i_N \cdot R_T$. Check it out! Amazing!!

CHECK YOUR UNDERSTANDING

Find the Thévenin equivalent resistance seen by the load resistor R_o.



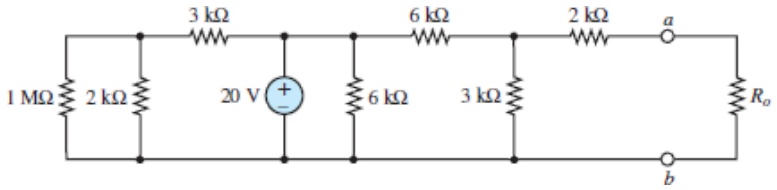
Find the Thévenin equivalent resistance seen by the load resistor R_o.



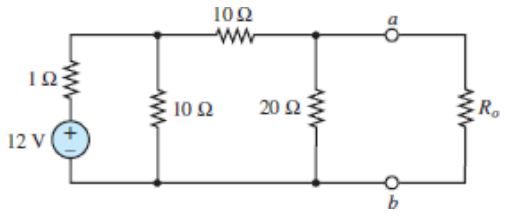
Answer: $R_T = 2.5\ \text{k}\Omega;\ R_T = 7\Omega$

CHECK YOUR UNDERSTANDING

For each circuit below, find the Thévenin equivalent resistance seen by the load resistor R_o.



Page 127



Answer: $R_T = 4.23\ \text{k}\Omega;\ R_T = 7.06\Omega$

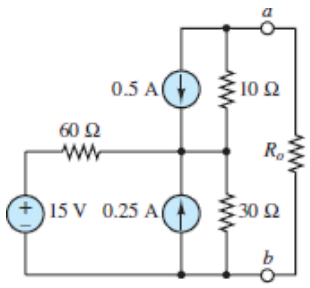
CHECK YOUR UNDERSTANDING

Find the open-circuit voltage v_{OC} for the circuit of Figure 2.53 if $R_1 = 5\Omega$.

Answer: 4.8 V

CHECK YOUR UNDERSTANDING

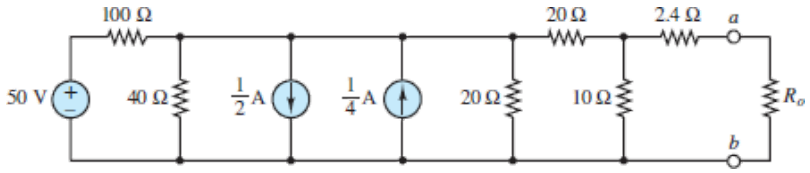
Find the Thévenin equivalent network seen by the load resistor R_o.



Answer: $R_T = 30\Omega$; $v_{OC} = v_T = 5$ V

CHECK YOUR UNDERSTANDING

Find the Thévenin equivalent network seen by the load resistor R_o. Source transformations are very useful in this problem.



Answer: $R_T = 10\Omega$; $v_{OC} = v_T = 0.704$ V

Experimental Determination of Thévenin and Norton Equivalents

Thévenin and Norton equivalent networks are often used as linear models of practical devices, such as batteries, power supplies, voltmeters, and ammeters, over a limited range of operation. While it is usually not possible nor feasible, because of the internal complexity of the devices, to determine those models analytically, simple experimental methods can be used instead. In practice, it is very useful to measure, for example, the equivalent internal (Thévenin) resistance of an instrument, so as to understand its operating limits and power requirements. Essentially, the linear model of a device is completely determined by its Thévenin (open-circuit) voltage v_T and its Norton (short-circuit) current i_N. The equivalent internal (Thévenin) resistance R_T is

$$R_T = \frac{v_T}{i_N} \tag{2.42}$$

Figure 2.68 illustrates the measurement of the short-circuit current and the open-circuit voltage. In practice, it is generally not a good idea to measure the short-circuit current directly with an ammeter since the input resistance of Page 129an ammeter is, by design, typically quite small. If these direct measurements are made the finite meter resistances r_A and r_V must be accounted for in the computation of the short-circuit current i_{SC} and the open-circuit voltage v_{OC}, respectively.

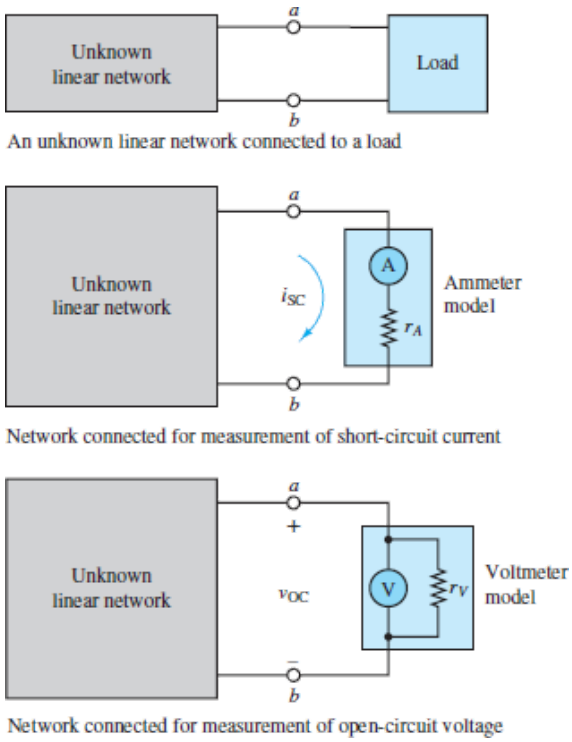


Figure 2.68 Measurement of open-circuit voltage and short-circuit current

Current and voltage division can be applied, respectively, to obtain the following expressions for the Norton current i_N and the Thévenin voltage v_T,

$$i_N = i_{SC}\left(1 + \frac{r_A}{R_T}\right) \tag{2.43}$$

$$v_T = v_{OC}\left(1 + \frac{R_T}{r_V}\right)$$

where R_T is the Thévenin equivalent resistance across terminals a and b of the unknown linear network. For an ideal ammeter, the internal resistance r_A is zero (a short-circuit). For an ideal voltmeter, the internal resistance r_V is infinite (an open-circuit). The two expressions in Equation 2.43 determine the "true" Thévenin and Norton equivalent networks using imperfect measurements of the short-circuit current and the open-circuit voltage, provided that the internal meter resistances are known. In practice, when the equivalent resistance seen by a voltmeter is much smaller than r_V, the measured v_{OC} will closely approximate the "true" v_{OC}. Likewise, when the equivalent resistance seen by an ammeter is much larger than r_A, the measured i_{SC} will closely approximate the "true" i_{SC}.

It is often not advisable to measure i_{SC} directly with an ammeter since its magnitude is not known. An ammeter is designed to approximate a short-circuit when inserted in a network, such that a large current may result and destroy an over-current protection fuse and perhaps damage the ammeter itself.

An alternative to measuring i_{SC} directly is to collect data along the *load line* of the unknown linear network and extrapolate i_{SC} from that data. Figure 2.27 depicts a typical load line associated with a linear network. Experimental load line data can be acquired by inserting resistive loads between the device terminals. The first load should be an open-circuit to determine directly the open-circuit voltage. The second load should be very large and followed by successively smaller loads. The load voltage can be measured by a voltmeter and the load current deduced by applying Ohm's law to the resistive load. For an ideal linear device, these data points will trace a straight line from the intersection with the voltage axis (v_{OC}) to the intersection with the current axis, which is the short-circuit current i_{SC}. In practice, experimental errors should be accounted for by using the load line data to compute a "best fit" trendline.

Page 130

FOCUS ON MEASUREMENTS

Experimental Determination of Thévenin Equivalent Networks

Problem:

Determine the Thévenin equivalent of an unknown source network from measurements of open-circuit voltage and short-circuit current.

Solution:

Known Quantities — Short-circuit current i_{SC}, open-circuit voltage v_{OC}, ammeter internal resistance r_A, and voltmeter internal resistance r_V.

Find — Equivalent resistance R_T; Thévenin voltage $vT = v_{OC}$.

Schematics, Diagrams, Circuits, and Given Data — Measured v_{OC} = 6.5 V; measured i_{SC} = 3.25 mA; r_A = 25Ω; r_V = 10 MΩ.

Assumptions — The unknown network is linear containing ideal sources and resistors only. The short-circuit current was able to be measured directly using an ammeter without damaging the instrument or its fuse.

Analysis — The unknown circuit shown in Figure 2.69 is replaced by its Thévenin equivalent and is connected to an ammeter to measure the short-circuit current and to a voltmeter to measure the open-circuit voltage. Ohm's law can be applied to the current measurement to find:

$$i_{SC} = \frac{v_T}{R_T + r_A}$$

Voltage division can be applied to the voltage measurement to find:

$$v_{OC} = \frac{r_V}{R_T + r_V} v_T$$

These expressions can be solved for v_T to yield:

$$\begin{aligned} v_T &= i_{SC}(R_T + r_A) \\ &= v_{OC}\left(1 + \frac{R_T}{r_V}\right) \end{aligned}$$

Or

$$i_{SC} R_T \left(1 + \frac{r_A}{R_T}\right) = v_{OC}\left(1 + \frac{R_T}{r_V}\right)$$

Since r_V is typically on the order of 10^6 times larger than r_A, one or both of the fractions in the previous expression will be negligible for a given R_T. Under the assumption that $R_T \ll r_V$ the above expression is approximated by:

$$i_{SC} R_T \left(1 + \frac{r_A}{R_T}\right) = v_{OC}$$

Under the assumption that $R_T \gg r_A$ the above expression is instead approximated by:

$$i_{SC} R_T = v_{OC} \left(1 + \frac{R_T}{r_V}\right)$$

If both assumptions are true, the Thévenin equivalent resistance is approximated by:

$$i_{SC} R_T = v_{OC}$$

which is the calculation that many inexperienced users make for every measurement, regardless of the relative values of R_T, r_A, and r_V. Of course, R_T is not known a priori so it is important to consider whether either or both of the limiting assumptions is reasonable.

Page 131

Consider the example measurement data listed above. The measured values of the short-circuit current and open-circuit voltage are:

$$i_{SC} = 3.25\,\text{mA} \qquad \text{and} \qquad v_{OC} = v_T = 6.5\,\text{V}$$

If both limiting assumptions are made, then the Thévenin equivalent resistance R_T between terminals *a* and *b* of the unknown network is approximately:

$$R_T \approx \frac{v_{OC}}{i_{SC}} = 2.0\,\text{k}\Omega$$

This value is 80 times larger than r_A but 5000 times smaller than r_V. Thus, one might expect that the impact of r_A is more significant than the impact of r_V for this particular network.

If only $R_T \ll r_V$ is assumed, then using the appropriate expression above yields:

$$R_T \approx \frac{v_{OC}}{i_{SC}} - r_A = 2.0\,\text{k}\Omega - 25\,\Omega = 1975\,\Omega$$

which is a 1.25% change from 2.0kΩ. If only $R_T \gg r_A$ is assumed, then using the appropriate expression above yields:

$$R_T \approx \frac{v_{OC}}{i_{SC}} \frac{r_V}{r_V - \frac{v_{OC}}{i_{SC}}} = (2.0\,\text{k}\Omega)\frac{10^7}{10^7 - 2.0\,\text{k}\Omega} = 2000.4\,\Omega$$

which is a negligibly small 0.02% change from 2.0 kΩ. If neither limiting assumption is made, then R_T is:

$$R_T = \frac{v_{OC} - i_{SC} r_A}{i_{SC} - \frac{i_{SC}}{r_V}} = 1975.4\,\Omega$$

As expected it is important in this example to include the impact of r_A when calculating the "true" value of R_T. The impact of r_V on the calculation is negligible.

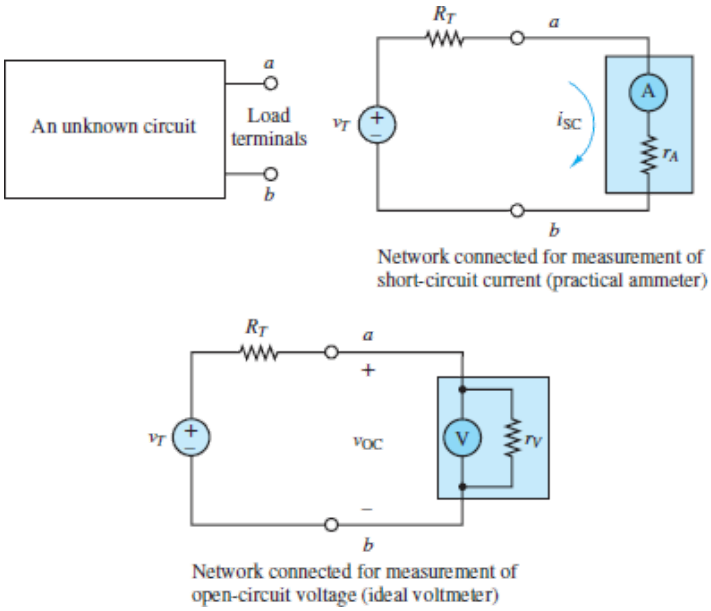


Figure 2.69

Page 132

2.8 MAXIMUM POWER TRANSFER

The reduction of any linear resistive circuit to its Thévenin or Norton equivalent form is a very convenient conceptualization, as far as the computation of load-related quantities is concerned. One such computation is that of the power absorbed by the load. The Thévenin and Norton models imply that some of the power generated by

the source will necessarily be dissipated by the internal circuits within the source. Given this unavoidable power loss, a logical question to ask is, How much power can be transferred to the load from the source under the most ideal conditions? Or, alternatively, what is the value of the load resistance that will absorb maximum power from the source? The answer to these questions is contained in the **maximum power transfer theorem**.

The model employed in the discussion of power transfer is illustrated in Figure 2.70, where a one-port linear network is represented by means of its Thévenin equivalent network. The power absorbed by the load P_o is

$$P_o = i_o^2 R_o \tag{2.44}$$

and the load current is

$$i_o = \frac{v_T}{R_o + R_T} \tag{2.45}$$

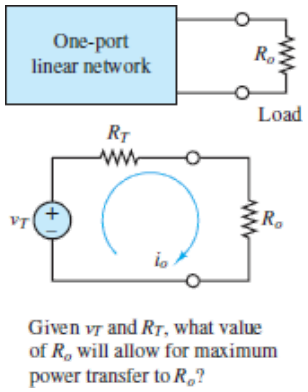


Figure 2.70 Power transfer between source and load

Combining the two expressions, the load power can be computed as

$$P_o = \frac{v_T^2}{(R_o + R_T)^2} R_o \tag{2.46}$$

The expression for P_o can be differentiated with respect to R_o and set to zero to find the value of R_o that gives the maximum power absorbed by the load. (Here, v_T and R_T are assumed constant.)

$$\frac{dP_o}{dR_o} = 0 \tag{2.47}$$

Plug in for P_o and solve to obtain:

$$\frac{dP_o}{dR_o} = \frac{v_T^2(R_o + R_T)^2 - 2v_T^2 R_o(R_o + R_T)}{(R_o + R_T)^4} \tag{2.48}$$

Thus, at the maximum value of P_o the following expression must be satisfied.

$$(R_o + R_T)^2 - 2R_o(R_o + R_T) = 0 \tag{2.49}$$

The solution of this equation is

LO

$$\boxed{R_o = R_T} \tag{2.50}$$

Thus, to transfer maximum power to a load, the load resistance must **match**the Thévenin equivalent resistance. Figure 2.71 depicts a plot of the load power divided by v_T^2 versus the ratio of R_o to R_T. Note that load power is maximized when $R_o = R_T$.

Page 133

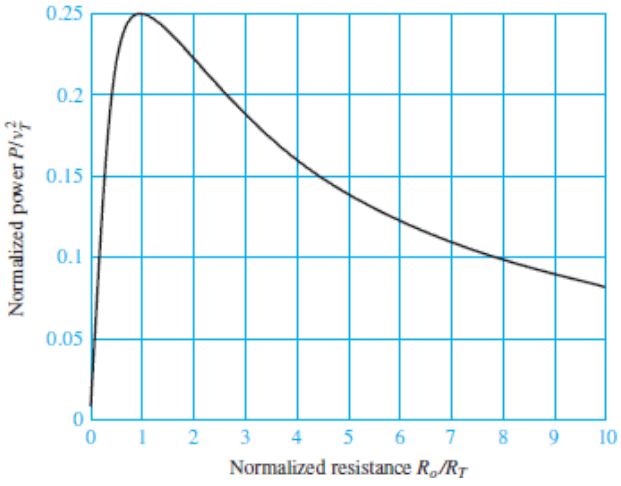


Figure 2.71 Graphical representation of maximum power transfer

An phenomenon related to power transfer is **source loading**, which is illustrated in Figure 2.72. When a one-port linear network is connected to a load, the voltage across the load will be somewhat lower than the *open-circuit voltage* (the Thévenin voltage) of the source. The extent of the decrease in voltage depends on the amount of current drawn by the load. With reference to Figure 2.72, the voltage decrease is equal to iR_T, and therefore the load voltage will be

$$v_o = v_T - i_o R_T \tag{2.51}$$

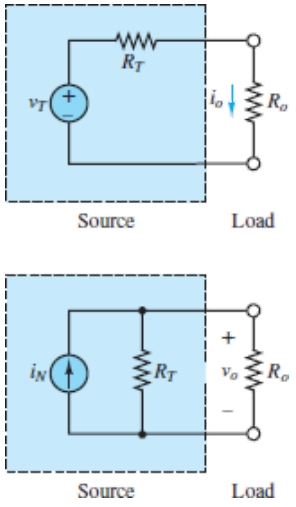


Figure 2.72 Loading

It should be apparent that it is desirable to have as small an internal resistance as possible in a practical voltage source.

In the case of a current source, the current through the load will be somewhat lower than the *short-circuit current* (the Norton current) of the source:

$$i_o = i_N - \frac{v_o}{R_T} \tag{2.52}$$

It is therefore desirable to have a very large internal resistance in a practical current source. Refer to the discussion of practical sources later in this chapter to see that they are often represented by Thévenin and Norton equivalent networks.

EXAMPLE 2.17 Maximum Power Transfer

Problem

Use the maximum power transfer theorem to determine the increase in power delivered to a loudspeaker resulting from matching the speaker resistance to the amplifier output resistance R_T, as depicted in the simplified model of Figure 2.73.

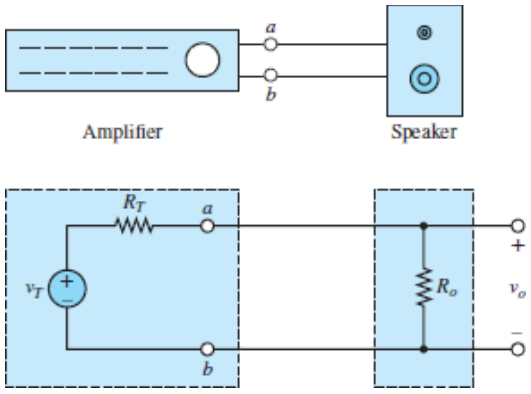


Figure 2.73 A simplified model of an audio system

Solution

Known Quantities: Source equivalent resistance R_T; unmatched speaker load resistance R_U; matched loudspeaker load resistance R_M.

Find: Difference between power delivered to loudspeaker with unmatched and matched loads, and corresponding percentage increase.

Schematics, Diagrams, Circuits, and Given Data: $R_T = 8\Omega$; $R_U = 16\Omega$; $R_M = 8\Omega$.

Assumptions: The amplifier can be modeled as a one-port linear network.

Analysis: Consider connecting (unwittingly) an 8-Ω amplifier to a 16-Ω speaker. The power delivered to the speaker can be computed using voltage division as follows:

$$v_U = \frac{R_U}{R_U + R_T} v_T = \frac{2}{3} v_T$$

and the load power is then computed to be

$$P_U = \frac{v_U^2}{R_U} = \frac{4}{9}\frac{v_T^2}{R_U} = 0.0278\, v_T^2$$

Repeat the calculation for the case of a matched 8-Ω speaker resistance R_M. The new load voltage v_M and the corresponding load power P_M are calculated as follows:

$$v_M = \frac{1}{2} v_T$$

and

$$P_M = \frac{v_M^2}{R_M} = \frac{1}{4}\frac{v_T^2}{R_M} = 0.03125\, v_T^2$$

The increase in load power is therefore

$$\Delta P = \frac{0.03125 - 0.0278}{0.0278} \times 100 = 12.5\%$$

Comments: In practice, an audio amplifier and a speaker are not well represented by the simple resistive models used in this example. Circuits that are appropriate to model amplifiers and loudspeakers are presented in later chapters.

Page 135

CHECK YOUR UNDERSTANDING

A practical voltage source has an internal resistance of 1.2 Ω and generates a 30-V output under open-circuit conditions. What is the smallest load resistance that can be connected to the source if the decrease in load voltage is to be no more than 2 percent with respect to the source open-circuit voltage?

A practical current source has an internal resistance of 12 kΩ and generates a 200-mA output under short-circuit conditions. What percentage drop in load current will be experienced (with respect to the short-circuit condition) if a 200-Ω load is connected to the current source?

Answer: 58.8 Ω; 1.64%

2.9 PRACTICAL VOLTAGE AND CURRENT SOURCES

Ideal independent sources are defined such that their prescribed output, a voltage or current, is completely independent of other factors. An ideal independent voltage source maintains a prescribed voltage across its terminals independent of the current through those terminals; likewise, an ideal independent current source maintains a prescribed current through its terminals independent of the voltage across those terminals. Neither of these ideal sources account for the effective *internal resistance* of practical voltage and current sources, which makes the output of a practical source dependent on the load that is *seen* by the source.

Consider, for example, a conventional car battery rated at 12 V, 450 ampere-hours (A-h). The latter rating implies that there is a limit (albeit a large one) to the amount of current the battery can deliver to a load and that, to some extent, the voltage output of the battery is dependent on the current drawn from it. This dependency can be observed as a drop in battery voltage when starting an automobile. Fortunately, a detailed understanding and analysis of the battery's physics are not necessary to model its behavior. Instead, the concept of internal resistance allows practical sources to be approximated by either of two different yet simple and effective models.

A practical voltage source can be approximated by a *Thévenin* model, which is composed of an ideal voltage source v_S in series with an internal resistance r_S. In practice, r_S is designed to be small compared to a typical equivalent resistance seen by the source.

A practical current source can be approximated by a *Norton* model, which is composed of an ideal current source i_S in parallel with an internal resistance r_S. In practice, r_S is designed to be large compared to a typical equivalent resistance seen by the source.

Page 136

The shaded portion of Figure 2.74 depicts the so-called *Thévenin* model. The source output current i_o depends upon the ideal voltage source v_S, the internal resistance r_S, and the load R_o. The maximum current is found in the limit that the load $R_o \to 0$ (i.e., a short-circuit load). The equivalent resistance "seen" by the ideal voltage source is $r_S + R_o$. The source output current i_o is given by Ohm's law.

$$i_o = \frac{v_S}{r_S + R_o} \quad \text{such that} \quad i_{o\,\max} = \frac{v_S}{r_S}$$

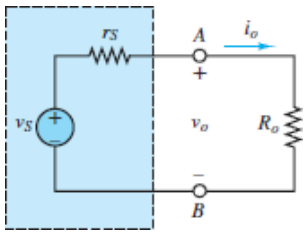


Figure 2.74 The Thévenin model of a real voltage source

The load voltage v_o can be found by direct application of voltage division.

$$\frac{v_o}{v_S} = \frac{R_o}{r_S + R_o}$$

In practice, the internal resistance r_S of a real voltage source is designed to be small compared to a typical load resistance R_o. In such cases, the load voltage v_o is approximately equal to the ideal source voltage v_S and the current requirements of a broad range of loads may be satisfied. Often, the effective internal resistance of a real voltage source is listed in its technical specifications. In cases where R_o is comparable to or smaller than r_S, the load voltage v_o will be significantly less than v_S. This result is known as a *loading effect*, such as when an automotive battery is required to start its engine.

The shaded portion of Figure 2.75 depicts the so-called *Norton* model. With this model, the source output voltage v_o depends upon the ideal current source i_o, the internal resistance r_S, and the load R_o. The maximum voltage is found in the limit that the load $R_o \rightarrow \infty$ (i.e., an open-circuit load). The equivalent resistance "seen" by the ideal current source is $r_S \;/\!/\; R_o$. The source output voltage v_o is given by Ohm's law.

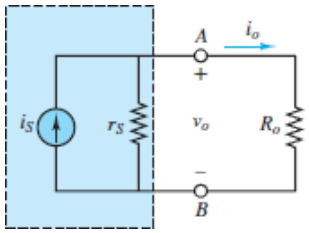


Figure 2.75 The Norton model of a real current source

$$v_o = i_S \frac{r_S R_o}{r_S + R_o} \quad \text{such that} \quad v_{o\,\max} = i_S r_S$$

The load current can be found by direct application of current division.

$$\frac{i_o}{i_S} = \frac{r_S \| R_o}{R_o} = \frac{r_S}{r_S + R_o}$$

In practice, the internal resistance r_S of a real current source is designed to be large compared to a typical load resistance R_o. In such cases, the load current i_o is approximately equal to the ideal source current i_S and the voltage requirements of a broad range of loads may be satisfied. Often, the effective internal resistance of a real

current source is listed in its technical specifications. In cases where R_o is comparable to or larger than r_S, the load current i_o will be significantly less than i_S. This result is also known as a *loading effect*.

2.10 MEASUREMENT DEVICES

In practice, the most commonly required measurements are of resistance, current, voltage, and power. An *ideal* measurement device would have no effect upon the quantity being measured. Of course, when a real measurement device is attached to a network, the network itself is changed (it now includes the measurement device) and it is quite possible that the quantity being measured is changed from what it was before the device was attached. At first glance, this problem may seem like a Page 137classic catch-22 scenario. That is, a quantity needs to be measured, so a measurement device must be used; but when the measurement device is used, the quantity is no longer what is was. To restore the quantity to its original state, the measurement device must be removed, but then … and so on and so on, around and around.

Luckily, if the characteristics of the measurement device are known, it is often possible to estimate the qualitative and quantitative impacts of a device on the measured quantity. In this section, simple models of real measurement devices are introduced that allow reasonable estimates of both.

The Ohmmeter

An **ohmmeter** measures the equivalent resistance *across* two nodes. In particular, an ohmmeter measures the resistance across an element when connected in *parallel* with it. Figure 2.76 depicts an ohmmeter connected across a resistor. One important rule needs to be remembered when using an ohmmeter:

> When using an ohmmeter to measure the resistance between two terminals of a circuit element it is important that the element and ohmmeter not be in parallel with any other resistive pathway.

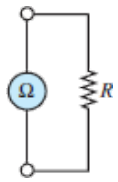


Figure 2.76 An ideal ohmmeter connected across a resistor

If the element is not disconnected from its network, the ohmmeter will measure the effective resistance of the element in parallel with the rest of its network. A common mistake made by inexperienced users of an ohmmeter is to attempt to measure the value of a discrete resistor by using fingers to clamp each end of the resistor to the ohmmeter probes. The user of the ohmmeter may be completely unaware that the measurement represents the equivalent resistance of the discrete resistor in parallel with the resistance from one hand to the other.

The Ammeter

An **ammeter** measures the current *through* an element when connected in *series* with it. Figure 2.77(a) shows an *ideal* ammeter inserted into a simple series loop to measure its current. An ideal ammeter has zero resistance and therefore is able to measure the current without alteration due to the presence of the ammeter. A more realistic model of an actual ammeter has an internal resistance in series with an ideal ammeter, as shown in Figure 2.77(b). To obtain an accurate measurement the internal resistance of the ammeter must be significantly *smaller* than the total equivalent Page 138resistance of the branch to which it is attached in series. For example, the internal resistance r_m of the ammeter must be significantly smaller than $R_1 + R_2$ in the series loop of Figure 2.77(a). In practice, it is necessary to observe two rules when using an ammeter:

1. When using an ammeter to measure the current *through* an element, the ammeter must be in series with the element.
2. When using an ammeter, its internal resistance should be much *smaller* than the total equivalent resistance in series with the ammeter.

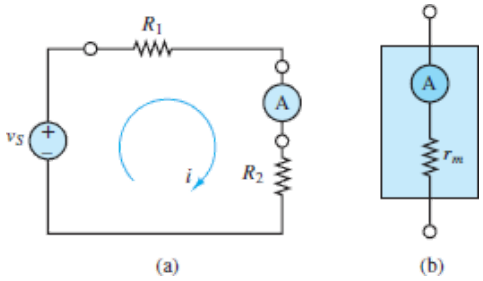


Figure 2.77 (a) An ideal ammeter in series with R_1 and R_2. (b) A practical model for an actual ammeter. r_m is the meter's internal resistance.

The Voltmeter

A **voltmeter** measures the voltage *across* an element when connected in *parallel* with it. Figure 2.78(a) shows an *ideal* voltmeter attached across resistor R_2, which is otherwise in a simple series loop. An ideal voltmeter has infinite resistance and therefore is able to measure the voltage without alteration due to the presence of the voltmeter. A more realistic model of an actual voltmeter has an internal resistance in parallel with an ideal voltmeter, as shown in Figure 2.78(b). To obtain an accurate measurement the internal resistance of the voltmeter must be significantly *larger* than the total equivalent resistance between the two nodes to which it is attached in parallel. For example, the internal resistance r_m of the voltmeter must be significantly larger than R_2 in the series loop of Figure 2.78(a). In practice, it is necessary to observe two rules when using a voltmeter:

1. When using a voltmeter to measure the voltage *across* an element, the voltmeter must be in parallel with the element.
2. When using a voltmeter, its internal resistance should be much *larger* than the total equivalent resistance in parallel with the voltmeter.

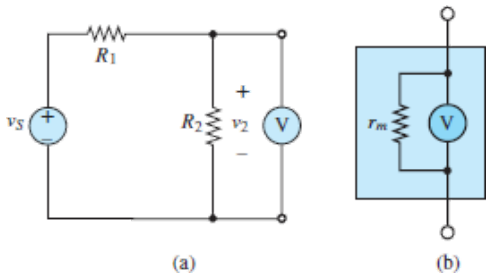


Figure 2.78 (a) An ideal voltmeter in parallel with R_2. (b) A practical model for an actual voltmeter. r_m is the meter's internal resistance.

The Wattmeter

A **wattmeter** is a three-terminal device [see Figure 2.79(a)] that measures the power dissipated by a circuit element. A wattmeter is essentially a combination of an Page 139ammeter and a voltmeter, as shown in Figure 2.79(b). Thus, it should be no surprise that an actual wattmeter is modeled with internal resistances at its terminals similar to those found in the practical ammeter and voltmeter models. A wattmeter simultaneously measures the current through and the voltage across an element and computes the product of these two quantities to determine the power dissipated.

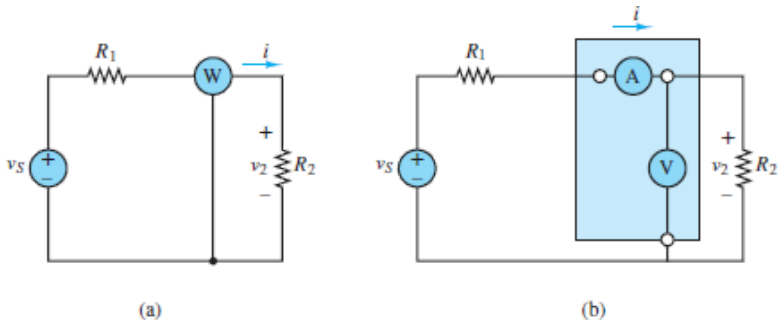


Figure 2.79 (a) An ideal wattmeter in series and parallel with R_2. (b) A model of an ideal wattmeter as a combination of an ideal ammeter and an ideal voltmeter. A practical model would replace the ideal meters with their own practical models.

EXAMPLE 2.18 Impact of a Practical Voltmeter

Problem

Use the tabulated data below to determine the effective internal resistance of the voltmeter shown in Figure 2.78(a), where the voltmeter is modeled as shown in Figure 2.78(b).

Solution

Known Quantities: $v_S = 5.0$ V; various values of $R_1 = R_2$; voltmeter data.

Find: The effective internal resistance r_m of the voltmeter.

Schematics, Diagrams, Circuits, and Given Data: Figure 2.78(a) and (b) and Table 2.1.

Table 2.1 Voltmeter data for determining internal resistance

$R_1 = R_2$	v_2 (V)
10 kΩ	2.49
470 kΩ	2.44
1 MΩ	2.38
4.7 MΩ	2.02
10 MΩ	1.67

Analysis: Substitute the practical model of a voltmeter shown in Figure 2.78(b) for the ideal voltmeter shown in Figure 2.78(a). Notice that the internal resistance r_m of the voltmeter is in parallel with R_2. Their parallel equivalent resistance is:

$$R_{EQ} = r_m \| R_2 = \frac{r_m R_2}{r_m + R_2}$$

The voltage across R_2 and the voltmeter can be found directly by voltage division:

$$\frac{v_2}{v_S} = \frac{R_{EQ}}{R_1 + R_{EQ}} \tag{2.53}$$

$$= \frac{r_m R_2}{R_1(r_m + R_2) + r_m R_2} \tag{2.54}$$

Divide the numerator and denominator by R_1 and gather coefficients of r_m.

$$= \frac{r_m(R_2/R_1)}{r_m(1 + R_2/R_1) + R_2} \tag{2.55}$$

Page 140

Multiply both sides by the denominator on the right and gather coefficients of r_m to find:

$$r_m = \frac{(v_2/v_S)R_2}{R_2/R_1 - (v_2/v_S)(1 + R_2/R_1)} \tag{2.56}$$

When $R_2 = R_1$:

$$= \frac{v_2/v_S}{1 - 2v_2/v_S} R_2 \tag{2.57}$$

Notice that $r_m = R_2$ when:

$$\frac{v_2/v_S}{1 - 2v_2/v_S} = 1$$

Solve for v_2/v_S to find:

$$\frac{v_2}{v_S} = \frac{1}{3}$$

Since v_S = 5.0 V, the previous condition is satisfied when v_2 = 5.0/3 = 1.67V. This value for v_2 is found in Table 2.1 for $R_1 = R_2$ = 10MΩ. Thus, the internal resistance of the voltmeter is

$$r_m = 10\,\text{M}\Omega$$

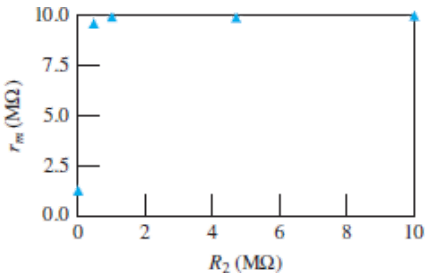
This value is typical of many handheld digital multimeters in voltmeter mode.

Comments: It is possible to acquire a separate estimate of r_m for each pair of values $R_1 = R_2$ and v_2 found in Table 2.1 by simply plugging in for v_2, v_S, R_1, and R_2, and solving for r_m. However, in practice, the calculated estimates for r_m will not be the same because they are particularly *sensitive* to experimental error when $R_2 \ll r_m$, as is the case for the first few pairs of data in the table. The least sensitivity to experimental error occurs when $R_2 = r_m$.

CHECK YOUR UNDERSTANDING

Find a separate estimate of r_m for each pair of values $R_1 = R_2$ and v_2 found in Table 2.1. Make a plot of r_m versus R_2.

Answer: Estimates for r_m are: 1.25M; 9.56M; 9.92M; 9.89M; 10.06M



2.11 NONLINEAR CIRCUIT ELEMENTS

The focus of this chapter so far has been on solving linear circuits. The example problems have had simple closed-form algebraic solutions that were readily interpreted with respect to their circuit diagrams. One reason for this result is that the loads have all been ideal resistors with a simple, linear i-v characteristic, namely Ohm's law. However, in practice, engineers are often faced with nonlinear i-v characteristics from elements such as diodes and transistors. This section explores two methods for analyzing nonlinear circuit elements.

Description of Nonlinear Elements

There are a number of useful cases in which a simple functional relationship exists between voltage and current in a nonlinear circuit element. Figure 2.80 depicts an element with an exponential *i*-*v* characteristic, described by the following equations:

$$\begin{aligned} i &= I_0(e^{\alpha v} - 1) \qquad v > 0 \\ i &= -I_0 \qquad\qquad\quad v \le 0 \end{aligned} \tag{2.58}$$

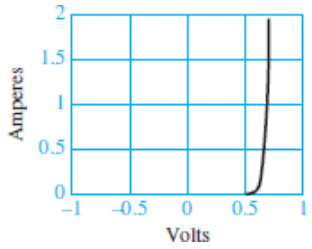


Figure 2.80 The *i*-*v* characteristic of an exponential resistor

In fact, this relationship approximates the nonlinear *i*-*v* characteristic of a semiconductor diode. The difficulty in dealing with circuits containing nonlinear elements is that it is not possible, in general, to obtain a closed-form analytic solution even for simple circuits.

One approach to analyzing a circuit containing a nonlinear element is to treat it as a load and to compute the Thévenin equivalent of the source network, as shown in Figure 2.81. Apply KVL to obtain the following equation:

$$v_T = R_T i_x + v_x \tag{2.59}$$

The second equation needed to solve for the unknown voltage v_x and the unknown current i_x is the *i*-*v* characteristic of the nonlinear load element. Assume the load is a semiconductor diode and consider, for the moment, only positive voltages. Since the reverse saturation current I_0 is typically very small, the circuit equations are well approximated by the following system:

$$\begin{aligned} i_x &= I_0 e^{\alpha v_x} \qquad v_x > 0 \\ v_T &= R_T i_x + v_x \end{aligned} \tag{2.60}$$

Although this system has two equations in two unknowns, one of the equations is nonlinear. By substituting the expression for i_x into the linear equation, the following expression is obtained:

$$v_T = R_T I_0 e^{\alpha v_x} + v_x \tag{2.61}$$

or

$$v_x = v_T - R_T I_0 e^{\alpha v_x} \tag{2.62}$$

This *transcendental* equation does not have a closed-form solution. How can v_x be found? One possibility is to generate a solution numerically, by guessing an initial value (for example, $vx = 0$) and iterating until a sufficiently precise solution is found. This method is explored further in the homework problems. Another method is based on a graphical analysis.

Page 142

Graphical (Load-Line) Analysis of Nonlinear Circuits

A nonlinear system of equations may be analyzed graphically. KVL can be applied to Figure 2.81 to write:

$$i_x = -\frac{1}{R_T} v_x + \frac{v_T}{R_T} \tag{2.63}$$

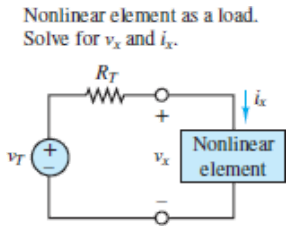


Figure 2.81 Representation of a nonlinear load in an otherwise linear circuit

This equation, which is referred to as the **load-line equation** and is valid for any load, linear or nonlinear, is a line in the (i_x, v_x) plane, with slope $-1/R_T$ and i_x intercept v_T/R_T. Its graphical representation is very useful and is shown in Figure 2.82.

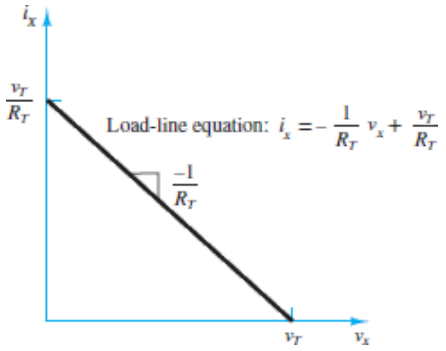


Figure 2.82 Load line

The other *i-v* characteristic is that of the nonlinear element. The intersection of the load line with the nonlinear *i-v* characteristic yields the solution, as depicted in Figure 2.83.

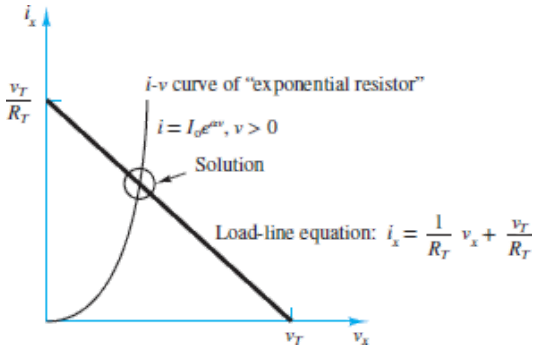


Figure 2.83 Graphical solution

It is important to emphasize that the methods introduced in this chapter for simplifying a linear source network can always be employed to solve circuits containing a single nonlinear load, as illustrated in Figure 2.84.

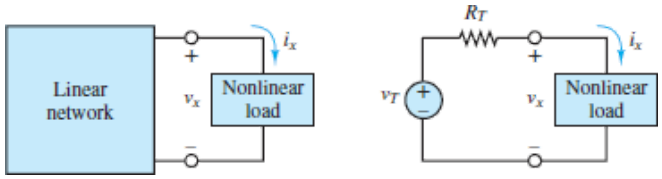


Figure 2.84 Transformation of a linear network to a Thévenin source

EXAMPLE 2.19 Nonlinear Load Power Dissipation

Problem

A linear generator is connected to a nonlinear load. Determine the power dissipated by the load.

Page 143

[

Solution

Known Quantities: Generator Thévenin equivalent circuit; load *i-v* characteristic and load line.

Find: Power dissipated by load P_x.

Schematics, Diagrams, Circuits, and Given Data: $R_T = 30\Omega$; $v_T = 15$ V; Figure 2.84; nonlinear load *i-v* characteristic.

Assumptions: None.

Analysis: Use the circuit model shown in Figure 2.84 to determine the voltage v_x and the current i_x, using graphical methods. The load-line equation for the circuit is given by:

$$i_x = -\frac{1}{R_T}v_x + \frac{v_T}{R_T} \quad \text{or} \quad i_x = -\frac{1}{30}v_x + \frac{15}{30}$$

This equation is a line in the $i_x v_x$ plane, with i_x intercept at 0.5 A and v_x intercept at 15 V. To determine the operating point of the circuit, superimpose the load line on the *i-v* characteristic of the device, as shown in Figure 2.85. The intersection of the load line with the device curve is the solution.

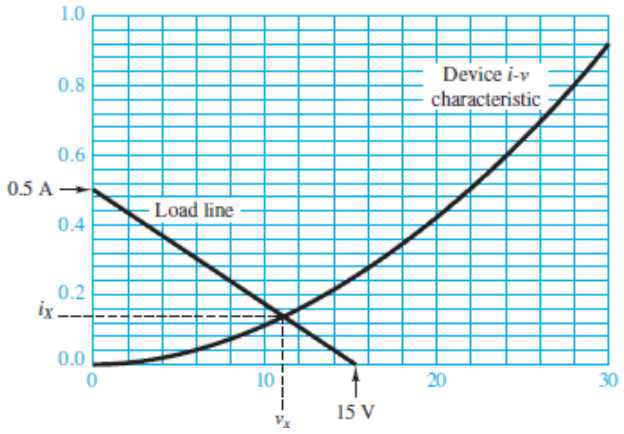


Figure 2.85 Graphical solution for circuit with nonlinear load

$$i_x = 0.14\,\text{A} \qquad v_x = 11\,\text{V}$$

Thus, the power dissipated by the nonlinear load is:

$$P_x = 0.14 \times 11 = 1.54\,\text{W}$$

The approach taken in this example is also, in essence, an experimental procedure. Many of the analytic methods presented in this chapter also apply to practical measurements.

EXAMPLE 2.20 Load-Line Analysis

Problem

A temperature sensor has a nonlinear i-v characteristic, shown in the figure on the left below. The load is connected to a linear network represented by a Thévenin source, as shown in Figure 2.84. Determine the current through the temperature sensor.

Page 144

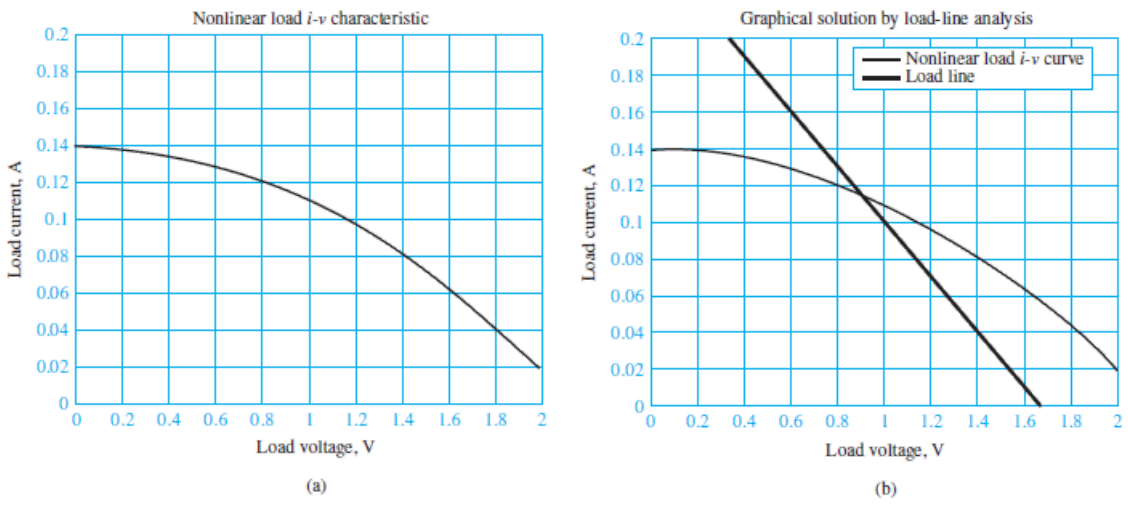
Solution

Known Quantities: $R_T = 6.67\,\Omega$; $V_T = 1.67$ V; $i_x = 0.14 - 0.03\,v_x^2$..

Find: i_x.

Analysis: The figure on the left depicts the device i-v characteristic. The figure on the right depicts a plot of both the device i-v characteristic and the load line obtained from

$$i_x = -\frac{1}{R_T}v_x + \frac{v_T}{R_T} = -0.15\,v_x + 0.25$$



The solution for v_x and i_x occurs at the intersection of the device and load-line characteristics: $i_x \approx 0.12$ A, $v_x \approx 0.9$ V..

CHECK YOUR UNDERSTANDING

Example 2.19 demonstrates a graphical solution method. Sometimes it is possible to determine the solution for a nonlinear load by analytical methods. Imagine that the same generator of Example 2.19 is now connected to a "square law" load, that is, one for which $v_x = \beta i_x^2$, with $\beta = 15.0$. Determine the load current i_x. [*Hint:* Assume that only positive solutions are possible, given the polarity of the generator.]

Answer: $i_x = 0.414$ A

CHECK YOUR UNDERSTANDING

Imagine that the same generator of Example 2.19 is connected to a nonlinear load i-v characteristic given by $i_x = 0.14 + 0.03 v_x^2$. Determine the load current i_x. [*Hint:* Assume that only positive solutions are possible, given the polarity of the generator.]

Answer: $i_x = 0.40$A

Page 145

Conclusion

This chapter introduced fundamental concepts needed to see meaningful structure and to successfully analyze electric circuits. The chapter also introduced methods that enable circuit problems to be simplified before finally being solved. The act of simplifying an electric circuit provides powerful insight into the essential nature of an electric circuit and how it could be modified to achieve a desired result. Such insight is a key aspect of any design methodology. Upon successful completion of this chapter, a student will have learned to:

1. Apply *voltage and current division* to calculate unknown voltages and currents in simple series, parallel, and series-parallel circuits. *Sections 2.1–2.2*.
2. Correctly redraw a resistive network, as necessary, and compute the equivalent resistance between two nodes. *Section 2.3*.

3. Apply the *principle of superposition* to linear circuits containing independent and dependent sources. *Section 2.4*.
4. Apply the source-load perspective to find graphical solutions to circuit problems. *Section 2.5*.
5. Apply source transformations to simplify and solve linear circuits containing independent and dependent sources. *Section 2.6*.
6. Determine *Thévenin and Norton equivalents* of linear one-port networks containing resistors and independent and dependent sources. *Section 2.7*.
7. Use equivalent circuit ideas to compute the maximum power transfer between a source and a load. *Section 2.8*.
8. Understand the impact of internal resistance in practical models of voltage and current sources as well as of voltmeters, ammeters, and wattmeters. *Sections 2.9–2.10*.
9. Use the concept of equivalent circuit to determine voltage, current, and power for nonlinear loads by using *load-line analysis* and analytical methods. *Section 2.11*.

The material covered in this chapter is essential to the development of more advanced techniques throughout the remainder of the book.

HOMEWORK PROBLEMS

Sections 2.1–2.2: Voltage and Current Division

2.1 Apply voltage division to the circuit of Figure P2.1. Assume that $v_S = 9\text{V}$, $R_1 = 8$ kΩ, $R_2 = R_3 = 10\text{k}\Omega$, $R_4 = 12$ kΩ. Find v_2.

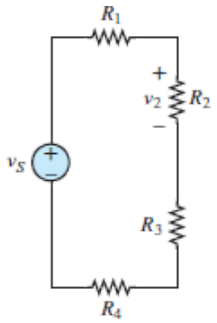


Figure P2.1

2.2 Refer to Figure P2.2, and assume that $v_S = 12$ V, $R_1 = 5\Omega$, $R_2 = 3\Omega$, $R_3 = 4\Omega$, and $R_4 = 5\Omega$. Apply voltage division to each resistive branch and KVL to find

the voltage v_{ab}.

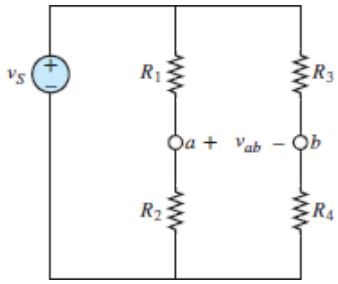


Figure P2.2

Page 146

2.3 Apply voltage division to each circuit in Figure P2.3 to find the value of R_o.

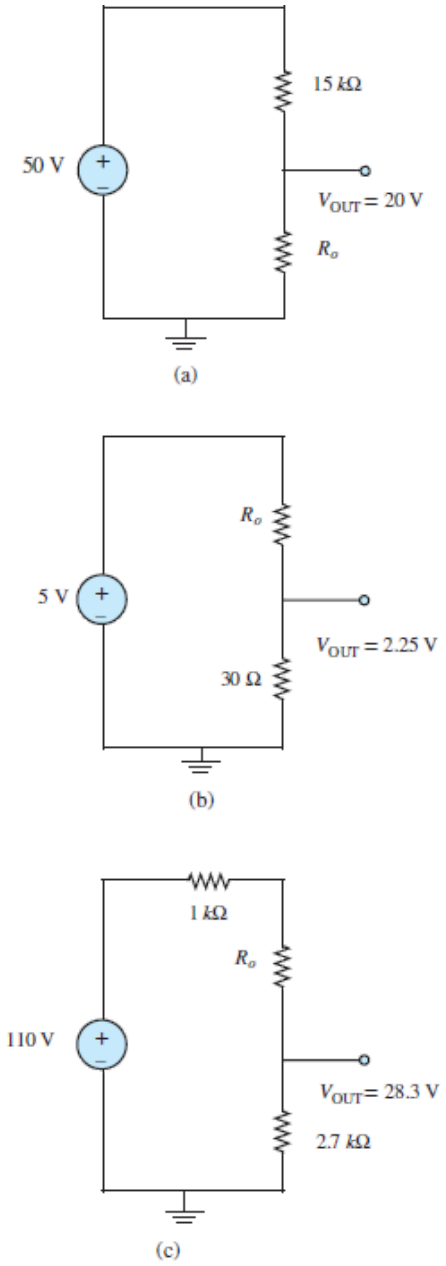


Figure P2.3

2.4 Apply the concepts of equivalent parallel resistance, voltage division, and current division to determine the current through each of the resistors R_4, R_5, and R_6 in Figure P2.4. $v_S = 10$ V, $R_1 = 20\Omega$, $R_2 = 40\Omega$, $R_3 = 10\Omega$, $R_4 = R_5 = R_6 = 15\Omega$.

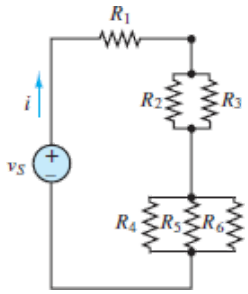


Figure P2.4

2.5 The voltage divider network of Figure P2.5 is designed to provide $v_{out} = v_S/>2$. However, in practice, the resistors may not be perfectly matched; that is, their tolerances are such that the resistances are unlikely to be identical. Apply voltage division to relate v_{out} to v_S and take the derivative of v_{out} to find an expression for dv_{out} in terms of the tolerances dR_1/R_1 and dR_2/R_2. Assume v_S = 10 V and nominal resistance values of $R_1 = R_2 = 5$ kΩ.

a. If the resistors have ±5 percent tolerance, find the expected range of possible output voltages.

b. Find the expected output voltage range for a tolerance of ±1 percent.

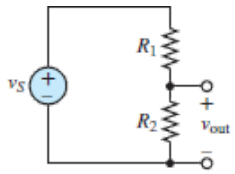


Figure P2.5

2.6 Apply voltage division to the circuit in Figure P2.6 to find an expression for the voltage across the variable resistor R. Use that expression to determine and plot the power absorbed by R, ranging from 0 to 30 Ω. Plot the power absorption as a function of R. Assume that $v_S = 15$ V, $R_S = 10$Ω.

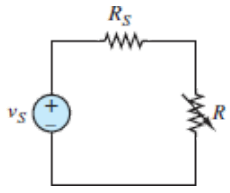


Figure P2.6

2.7 Apply voltage division to the circuit shown in Figure P2.7 to determine the terminal voltage v_o of the voltage source and the power absorbed by R_o.

$$v_S = 12\,\text{V} \quad R_S = 5\,\text{k}\Omega \quad R_o = 7\,\text{k}\Omega$$

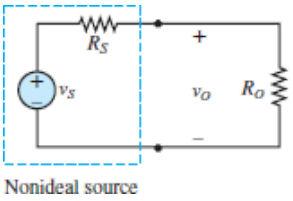


Figure P2.7

2.8 With no load R_o attached to the terminals of the nonideal source in Figure P2.7, the voltage drop v_o is 50.8 V. When a R_o = 10Ω load is attached, that voltage drop is 49 V. Apply voltage division to find an expression for v_o in terms of v_S, R_S and R_o.

a. Determine v_S and R_S for this nonideal source.

b. What voltage would be measured at the terminals in the presence of a 15-Ω load resistor?

c. How much current could be drawn from the nonideal source under short-circuit conditions?

2.9 Apply voltage division and KVL to determine the voltage v_o across terminals A and B in Figure P2.9.

$$v_S = 12\,\text{V}$$
$$R_1 = 11\,\text{k}\Omega \qquad R_3 = 6.8\,\text{k}\Omega$$
$$R_2 = 220\,\text{k}\Omega \qquad R_4 = 0.22\,\text{M}\Omega$$

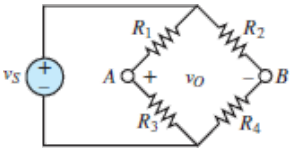


Figure P2.9

2.10 Refer to Figure P2.10 and assume v_S = 15V, R_1 = 12Ω, R_2 = 5Ω, R_3 = 8Ω, R_4 = 2Ω, R_5 = 4Ω, R_6 = 2Ω, and R_7 = 1Ω. Apply voltage division to find:

a. The voltage v_{ac} across nodes a and c.

b. The voltage v_{bd} across nodes b and d.

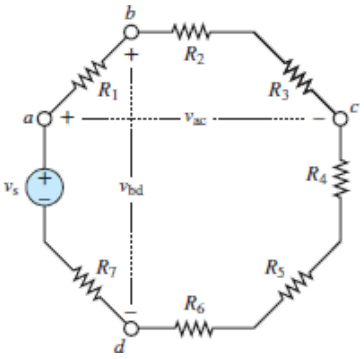


Figure P2.10

2.11 The circuit of Figure P2.11 is used to measure the internal resistance r_B of a battery.

a. A fresh battery is being tested, and it is found that the voltage v_{out} is 2.28 V with the switch open and 2.27 V with the switch closed. Apply voltage division to find the internal resistance of the battery.

b. The same battery is tested one year later. v_{out} is found to be 2.28 V with the switch open but 0.31 V with the switch closed. Apply voltage division to find the one-year-old internal resistance of the battery.

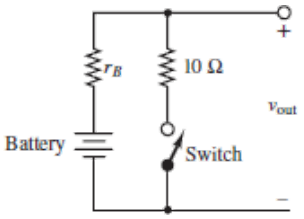


Figure P2.11

2.12 For the circuit shown in Figure P2.12, assume $i_S = 5$ A, $R_1 = 10\Omega$, $R_2 = 7\Omega$, $R_3 = 8\Omega$, $R_4 = 4\Omega$, and $R_5 = 2\Omega$. How many nodes are in the circuit? Use series and parallel equivalent resistance concepts to simply the network to the left of the current source into a single equivalent resistance. Apply current division to find the magnitude of the current through the branch containing R_4 and R_5.

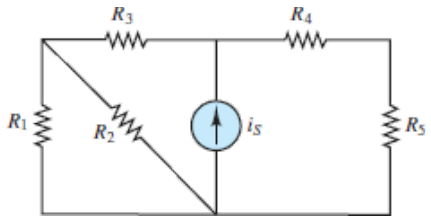


Figure P2.12

Page 148

2.13 Consider the practical ammeter, depicted in Figure P2.13, consisting of an ideal ammeter in series with a 1-kΩ resistor. (An ideal ammeter acts like a short-circuit.) The meter sees a full-scale deflection when the current through it is 30 μA. Depending on the setting of the rotary switch, the ammeter will read full scale when the current I equals 10 mA, 100 mA and 1 A. Apply current division to determine the appropriate values of R_1, R_2, and R_3.

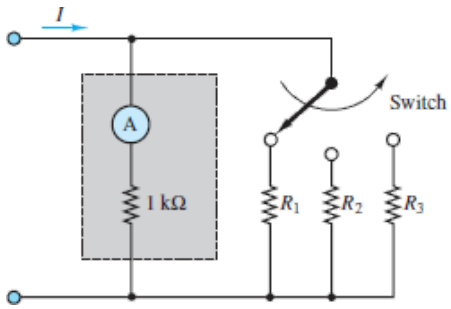


Figure P2.13

2.14 How many nodes are in the circuit shown in Figure P2.14? Use series and parallel equivalent resistance concepts to simplify the network to the right of node V_1 into a single equivalent resistance. Apply current division to find the magnitude of the current through the 3Ω resistor.

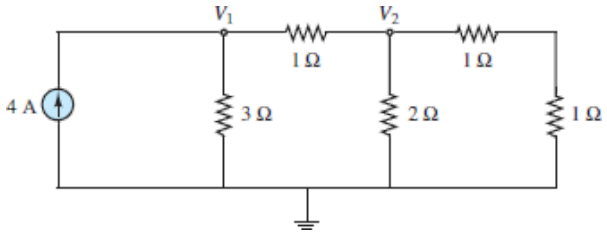


Figure P2.14

2.15 How many nodes are in the circuit shown in Figure P2.15? Use series and parallel equivalent resistance concepts to simplify the network to the right of the current source into a single equivalent resistance. Apply current division to find the magnitude of the current through R_1. Assume $R_1 = 10\Omega$, $R_2 = 9\Omega$, $R_3 = 4\Omega$, $R_4 = 4\Omega$, $i_S = 2$ A.

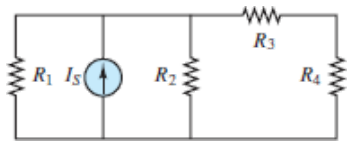


Figure P2.15

2.16 How many nodes are in the circuit shown in Figure P2.16? Apply current division to find the current through each resistive branch. Apply KVL and Ohm's law to find the magnitude of the voltage across nodes *a* and *b*. Assume $R_1 = 12\Omega$, $R_2 = 10\Omega$, $R_3 = 5\Omega$, $R_4 = 2\Omega$, $I_S = 3$ A.

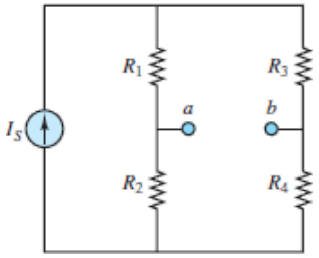


Figure P2.16

Section 2.3: Equivalent Resistance between Two Nodes

2.17 Find the equivalent resistance seen by the voltage source in Figure P2.17. Use that result and voltage division to find v_2.

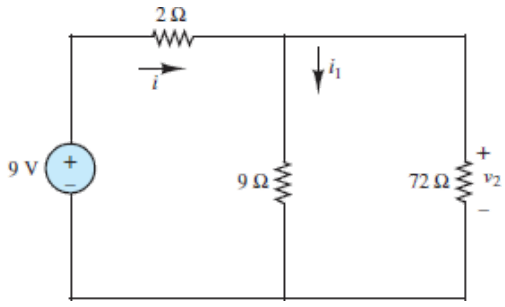


Figure P2.17

2.18 Find the equivalent resistance seen by the voltage source and the current i in the circuit of Figure P2.18.

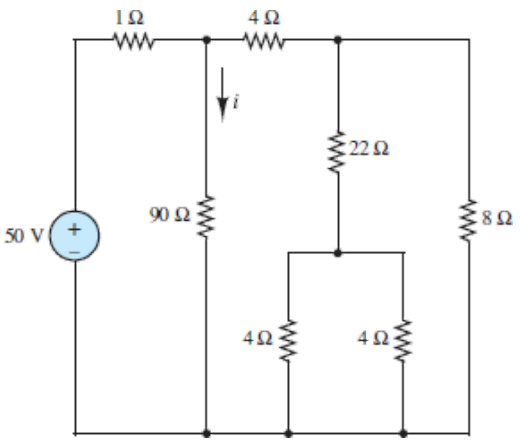


Figure P2.18

Page 149

2.19 In the circuit of Figure P2.19, the power absorbed by the 15-Ω resistor is 15 W. Find R.

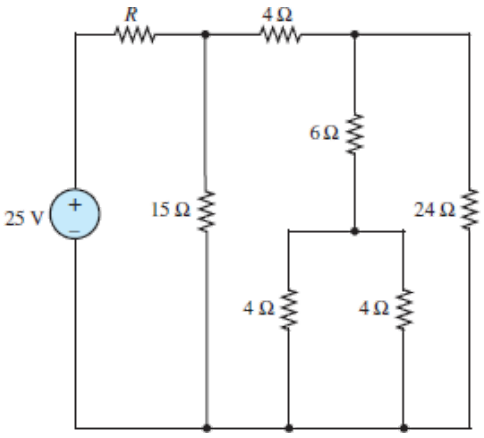


Figure P2.19

2.20 Find the equivalent resistance between terminals a and b in the circuit of Figure P2.20.

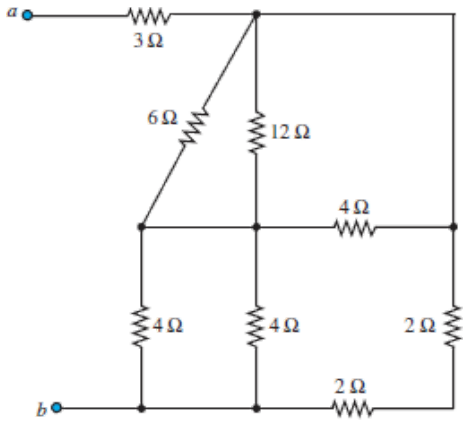


Figure P2.20

2.21 For the circuit shown in Figure P2.21, find the equivalent resistance seen by the voltage source. How much power is delivered by it?

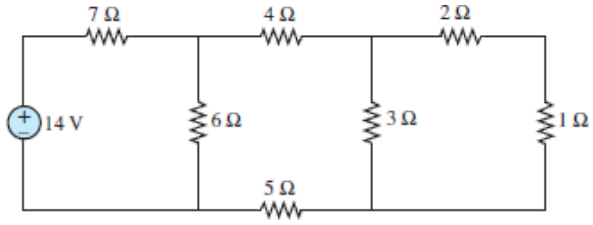


Figure P2.21

2.22 For the circuit shown in Figure P2.22, find the equivalent resistance seen by the current source. How many nodes are in the circuit? Assume $R_1 = 2\Omega$, $R_2 = 3\Omega$, $R_3 = 85\Omega$, $R_4 = 2\Omega$, and $R_5 = 4\Omega$.

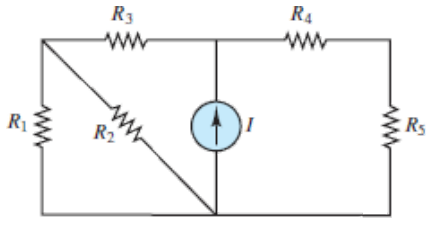


Figure P2.22

2.23 Refer to Figure P2.23. Assume $v_S = 20$ V, $R_1 = 10\Omega$, $R_2 = 5\Omega$, $R_3 = 8\Omega$, $R_4 = 2\Omega$, $R_5 = 4\Omega$, $R_6 = 2\Omega$, $R_7 = 1\Omega$, and $R_8 = 10\Omega$. How many nodes are in the circuit?

a. Determine the equivalent resistance seen by the voltage source v_S.

b. Apply voltage division to find the voltage across R_7 and R_8.

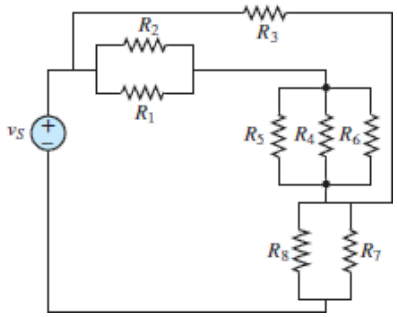


Figure P2.23

2.24 Find the equivalent resistance seen by the voltage source in Figure P2.24. How many nodes are in the circuit? Assume $R_1 = 12\Omega$, $R_2 = 5\Omega$, $R_3 = 8\Omega$, $R_4 = 2\Omega$, $R_5 = 4\Omega$, $R_6 = 2\Omega$, $R_7 = 1\Omega$, $R_8 = 10\Omega$, and $R_9 = 10\Omega$.

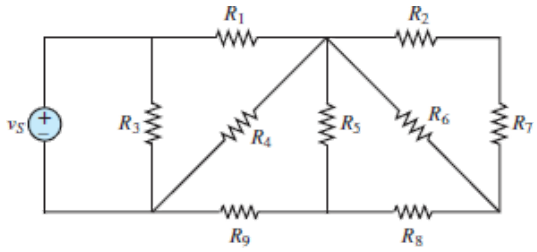


Figure P2.24

Page 150

2.25 For the circuit shown in Figure P2.25, assume $v_S = 10$ V, $R_1 = 9\Omega$, $R_2 = 4\Omega$, $R_3 = 4\Omega$, $R_4 = 5\Omega$, and $R_5 = 4\Omega$. Find:

a. The number of nodes in the circuit.

b. The equivalent resistance seen by the voltage source v_S.

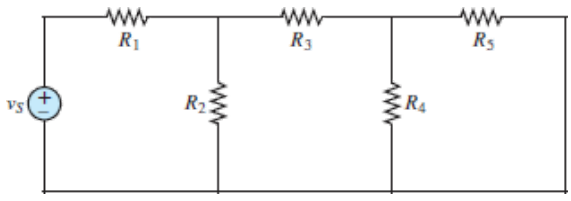


Figure P2.25

2.26 Determine the equivalent resistance of the infinite network of resistors in the circuit of Figure P2.26.

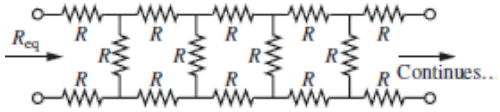


Figure P2.26

2.27 In the circuit of Figure P2.27, find the equivalent resistance between terminals *a* and *b* if terminals *c* and *d* are open and again if terminals *c* and *d* are shorted together. Also, find the equivalent resistance between terminals *c* and *d* if terminals *a* and *b* are open and again if terminals *a* and *b* are shorted together.

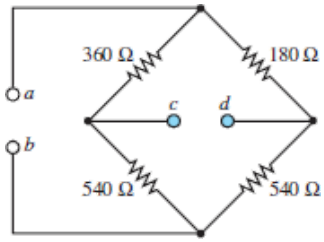


Figure P2.27

2.28 Refer to Figure P2.27 and determine the equivalent resistance between terminals *a* and *b* if terminal *c* is wired (shorted) to terminal *a* and terminal *d* is wired (shorted) to terminal *b*.

2.29 Apply the node voltage method to find the magnitude of the current through the voltage source. Use it and the definition of equivalent resistance between two nodes to find the equivalent resistance seen by the voltage source in Figure P2.29. How many nodes are in the circuit? Assume: $R_1 = 12\Omega$, $R_2 = 5\Omega$, $R_3 = 8\Omega$, $R_4 = 2\Omega$, and $R_5 = 4\Omega$.

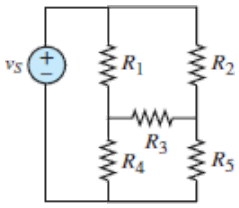


Figure P2.29

2.30 Refer to Figure P2.30 and assume $v_S = 15$ V, $R_1 = 12\Omega$, $R_2 = 5\Omega$, $R_3 = 8\Omega$, $R_4 = 2\Omega$, $R_5 = 4\Omega$, $R_6 = 2\Omega$, $R_7 = 1\Omega$, and $R_8 = R_9 = 10\Omega$. Find:

a. The number of nodes in the circuit.

b. The equivalent resistance seen by the voltage source v_S.

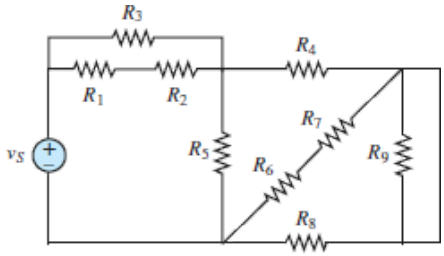


Figure P2.30

Section 2.4: The Principle of Superposition

2.31 Refer to Figure P2.31, and assume that $v_S = 7$ V, $i_S = 3$ A, $R_1 = 20\Omega$, $R_2 = 12\Omega$, and $R_3 = 10\Omega$. Apply the principle of superposition to find:

a. The component of i_1 due to v_S.

b. The component of i_2 due to i_S.

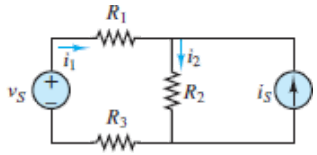


Figure P2.31

2.32 With reference to Figure P2.32, determine the current i through R_1 due only to the source V_{S2}.

$V_{S1} = 110$ V $V_{S2} = 90$ V
$R_1 = 560\,\Omega$ $R_2 = 3.5$ kΩ
$R_3 = 810\,\Omega$

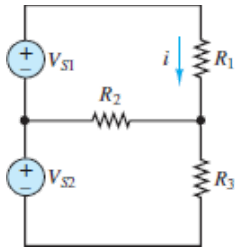


Figure P2.32

2.33 Refer to Figure P2.33 and use the principle of superposition to find the voltages at nodes A, B, and C. Assume V^1 = 12 V, V_2 = 10 V, R_1 = 2Ω, R_2 = 8Ω, R_3 = 12Ω, R_4 = 8Ω.

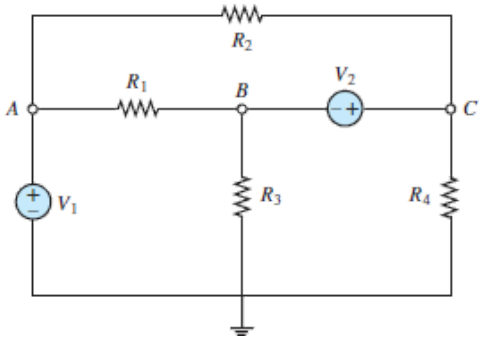


Figure P2.33

2.34 Use the principle of superposition to determine the voltage v across R_2 in Figure P2.34.

$$V_{S1} = V_{S2} = 12\,\text{V}$$
$$R_1 = R_2 = R_3 = 1\,\text{k}\Omega$$

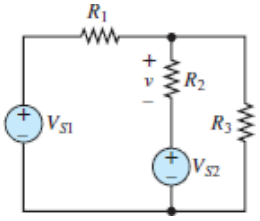


Figure P2.34

2.35 Refer to Figure P2.35 and use the principle of superposition to determine the component of the current i through R_3 that is due to V_{S2}.

$$V_{S1} = V_{S2} = 450\,\text{V}$$

$$R_1 = 7\,\Omega \qquad R_2 = 5\,\Omega$$

$$R_3 = 10\,\Omega \qquad R_4 = R_5 = 1\,\Omega$$

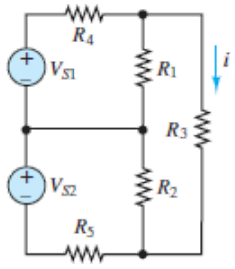


Figure P2.35

2.36 Refer to Figure P2.36 and use the principle of superposition to determine the current i through R_4 due to the current source i_S. Assume: $R_1 = 12\Omega$, $R_2 = 8\Omega$, $R_3 = 5\Omega$, $R_4 = 3\Omega$, $v_S = 3$ V, and $i_S = 2$ A.

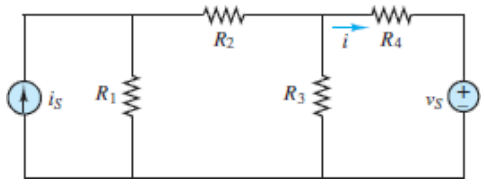


Figure P2.36

2.37 Refer to Figure P2.36 and use the principle of superposition to determine the current i through R_4 due to the voltage source v_S. Assume: $R_1 = 12\Omega$, $R_2 = 8\Omega$, $R_3 = 5\Omega$, $R_4 = 3\Omega$, $v_S = 3$ V, and $i_S = 2$ A.

2.38 Use the principle of superposition node to determine the voltages V_a and V_b in Figure P2.38. Let $R_1 = 10\Omega$, $R_2 = 4\Omega$, $R_3 = 6\Omega$, $R_4 = 6\Omega$, $V_1 = 2$ V, $V_2 = 4$ V, $I1 = 2$ A.

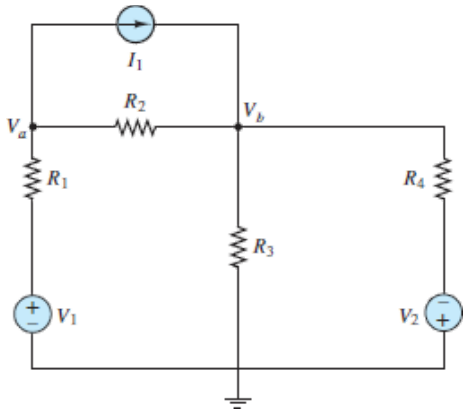


Figure P2.38

Page 152

2.39 Use the principle of superposition to determine the current i through R_3 in Figure P2.39. Let $R_1 = 10\Omega$, $R_2 = 4\Omega$, $R_3 = 2\Omega$, $R_4 = 2\Omega$, $R_5 = 2\Omega$, $VS = 10$ V, $i_S = 2$ A.

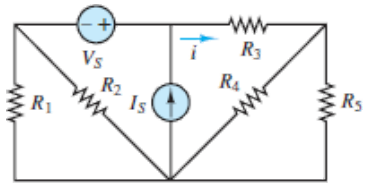


Figure P2.39

2.40 Figure P2.40 represents a temperature measurement system, where temperature T is linearly related to the voltage source V_{S2} by a transduction constant k. Use the principle of superposition to determine the components of V_{ab} due to V_{S1} and V_{S2} and then to determine the temperature.

$$V_{S2} = kT \qquad k = 10\,\text{V}/^\circ\text{C}$$
$$V_{S1} = 24\,\text{V} \qquad R_s = R_1 = 12\,\text{k}\Omega$$
$$R_2 = 3\,\text{k}\Omega \qquad R_3 = 10\,\text{k}\Omega$$
$$R_4 = 24\,\text{k}\Omega \qquad V_{ab} = -2.524\,\text{V}$$

In practice, the voltage across R_3 is used as the measure of temperature, which is introduced to the circuit through a temperature sensor modeled by the voltage source V_{S2} in series with R_s.

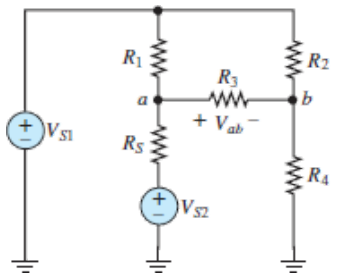


Figure P2.40

2.41 In Figure P2.41, use the principle of superposition to determine the components of the current through the voltage source v_S due to v_S and i_S, respectively. Use those results to determine the total current through v_S and the power supplied by it. Let $R_1 = 12\Omega$, $R_2 = 10\Omega$, $R_3 = 5\Omega$, $R_4 = 5\Omega$, $v_S = 10$ V, $i_S = 5$ A. (Note: Power is not a linear function of voltage or current and so power cannot be computed using the component currents separately.)

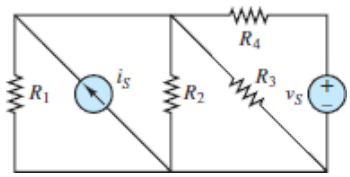


Figure P2.41

2.42 Use the principle of superposition to determine components of the current i_o through R_1 due to each independent source in Figure P2.42. Let $R_1 = 8\Omega$, $R_2 = 2\Omega$, $R_3 = 3\Omega$, $R_4 = 4\Omega$, $R_5 = 2\Omega$, $V_1 = 15$ V, $I1 = 2$ A, $I2 = 3$ A.

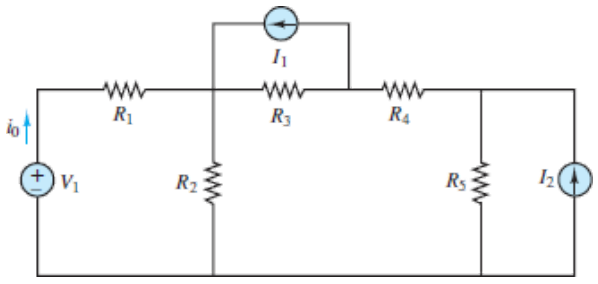


Figure P2.42

Section 2.6: Source Transformations

2.43 Apply two source transformations and current division in the circuit of Figure P2.43 to find I_2. Let $R_1 = 12\Omega$, $R_2 = 6\Omega$, $R_3 = 10\Omega$, $V_1 = 4$ V, $V_2 = 1$ V.

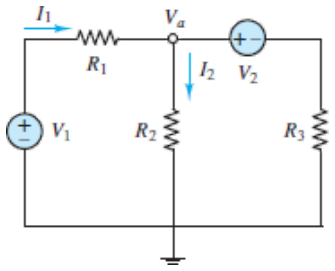


Figure P2.43

2.44 Apply source transformations to find the voltage V_o across R_0 for the circuit of Figure P2.44. Assume that $R_1 = 2\Omega$, $RV = R_2 = R_0 = 4\Omega$, $VS = 4V$, and $i_S = 0.5$ A.

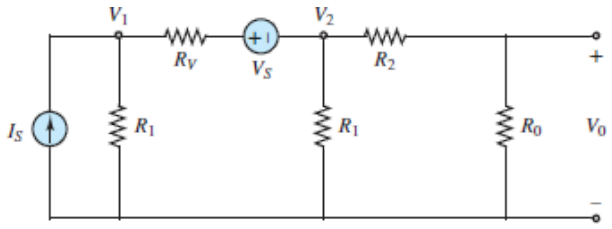


Figure P2.44

Page 153

2.45 Apply source transformations to find the mesh current I_3 for the circuit shown in Figure P2.45.

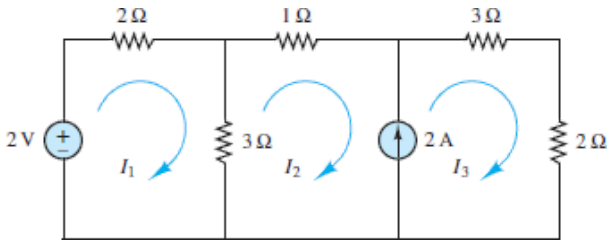


Figure P2.45

2.46 Apply source transformations to find the voltage V across the current source in Figure P2.46.

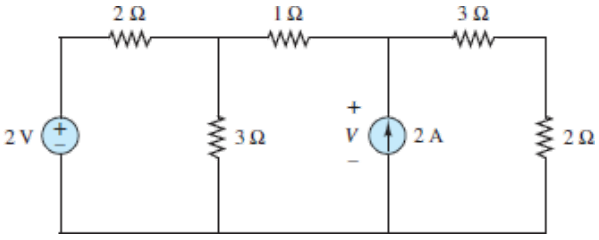


Figure P2.46

2.47 Apply a single source transformation and then voltage division to find the magnitude of the voltage across R_1 in Figure P2.47. Let $R_1 = 10\Omega$, $R_2 = 5\Omega$, $V_1 = 2$ V, $V_2 = 1$ V, $i_S = 2$ A.

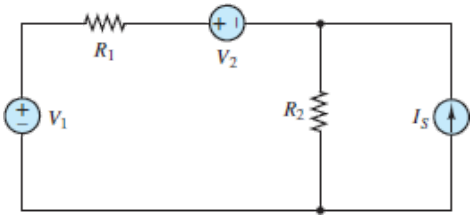


Figure P2.47

2.48 Transform each of the three Thévenin sources to Norton sources and apply current division to find the current through R_1 in Figure P2.48. Let $R_1 = 6\Omega$, $R_2 = 3\Omega$, $R_3 = 3\Omega$, $V_1 = 4$ V, $V_2 = 1$ V, $V3 = 2$ V.

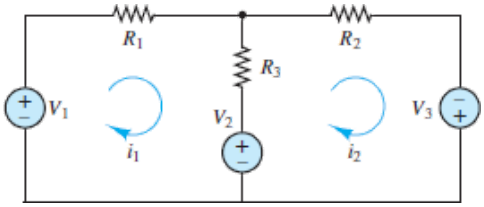


Figure P2.48

2.49 Simplify the circuit in Figure P2.49 by applying source transformations to the right half of the circuit. Solve for the node voltage v_1. (Note: Dependent sources can also be part of a source transformation as long as its reference variable is not obscured by the transformation.)

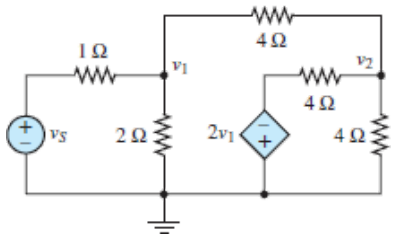


Figure P2.49

2.50 The circuit shown in Figure P2.50 is a simplified DC version of an AC three-phase wye-wye (Y-Y) electrical distribution system commonly used to supply industrial loads, particularly rotating machines.

$$V_{S1} = V_{S2} = V_{S3} = 170\,\text{V}$$
$$R_{w1} = R_{w2} = R_{w3} = 0.7\,\Omega$$
$$R_1 = 1.9\,\Omega \qquad R_2 = 2.3\,\Omega$$
$$R_3 = 11\,\Omega$$

a. Determine the number of nonreference nodes.

b. Determine the number of unknown node voltages.

c. Apply source transformations to find v_n'.

Notice that once v_n' is known the other unknown node voltages can be computed directly by voltage division.

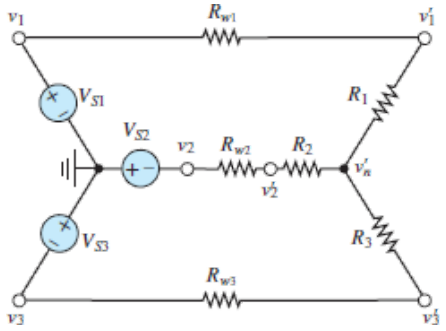


Figure P2.50

2.51 Apply source transformations to simplify the circuit in Figure P2.21. Solve for the magnitude of the current through the 1Ω resistor.

2.52 Apply source transformations to reduce the one-port network on the left side of Figure P2.82 to a Thévenin source. Apply voltage division to solve for the magnitude of the measured voltage when the voltmeter is attached.

Section 2.7: Thévenin and Norton Equivalent Networks

2.53 Find the Thévenin equivalent of the network seen by the 3-Ω resistor in Figure P2.53.

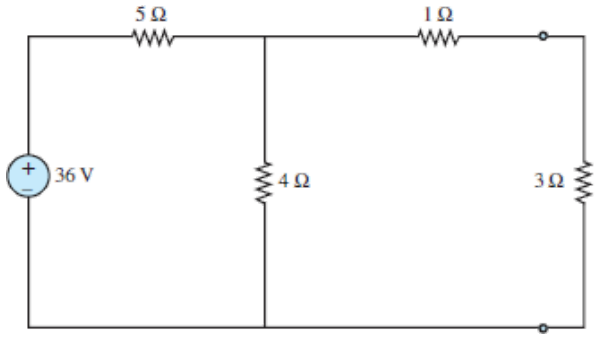


Figure P2.53

2.54 Find the Thévenin equivalent of the network seen by the 3-Ω resistor in Figure P2.54. Use it and voltage division to find the voltage v across the 3-Ω resistor.

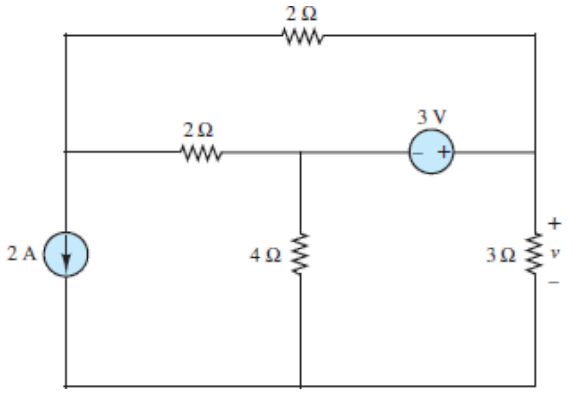


Figure P2.54

2.55 Find the Norton equivalent of the network seen by R_2 in Figure P2.55. Use it and current division to compute the current i through R_2. Assume $I_1 = 10$ A, $I_2 = 2$ A, $V_1 = 6$ V, $R_1 = 3\Omega$, and $R_2 = 4\Omega$.

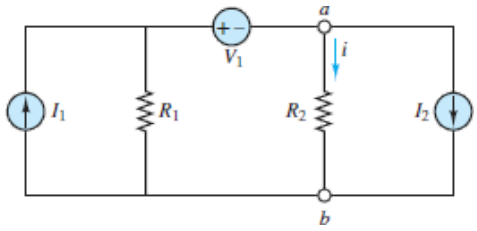


Figure P2.55

2.56 Find the Norton equivalent of the network between nodes *a* and *b* in Figure P2.56.

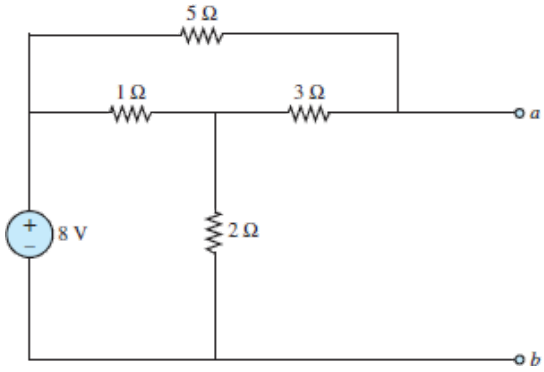


Figure P2.56

2.57 Find the Thévenin equivalent of the network seen by R in Figure P2.57, and use the result to compute the current i_R. Assume V_o = 10 V, I_o = 5 A, R_1 = 2Ω, R_2 = 2Ω, R_3 = 4Ω, and R = 3Ω.

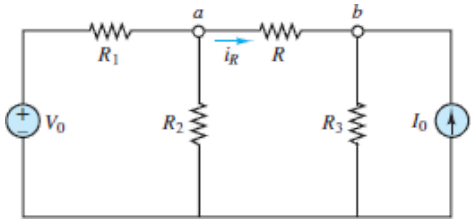


Figure P2.57

2.58 Find the Thévenin equivalent resistance seen by the load R_o in Figure P2.58.

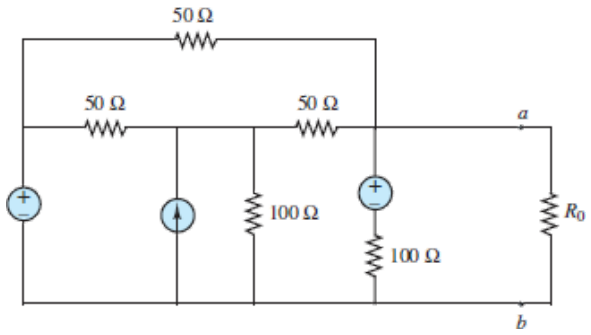


Figure P2.58

2.59 Find the Thévenin equivalent of the network seen by the load R_o in Figure P2.59.

Page 155

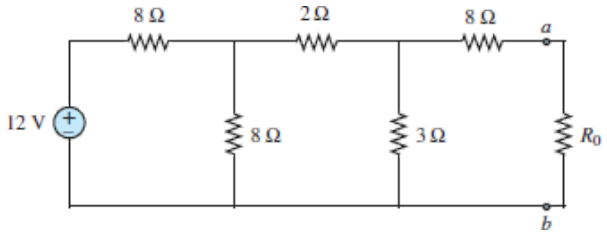


Figure P2.59

2.60 Find the Thévenin equivalent network seen by the load R_o in Figure P2.60, where $R_1 = 10\Omega$, $R_2 = 20\Omega$, $R_g = 0.1\Omega$, and $R_p = 1\Omega$.

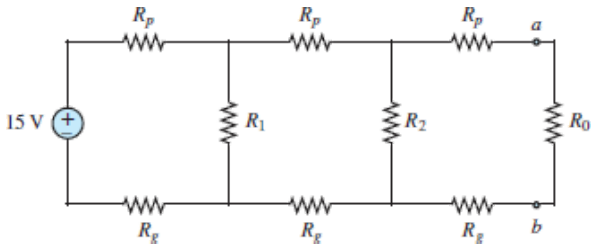


Figure P2.60

2.61 A Wheatstone bridge such as that shown in Figure P2.61 is used in numerous practical applications, such as determining the value of an unknown resistor R_X.

Find the Thévenin equivalent network seen by terminals a and b in terms of R, R_X, and V_S. Use it to find the value of R_x when $R = 1$ kΩ, $V_S = 12$ V, and $V_{ab} = 12$ mV.

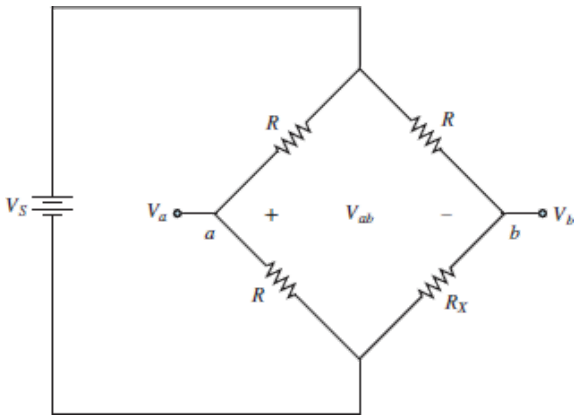


Figure P2.61

2.62 Thévenin's theorem can be useful when dealing with a Wheatstone bridge. For the circuit of Figure P2.62:

a. Express the Thévenin equivalent resistance seen by the load resistor R_o in terms of R_1, R_2, R_3, and R_X.

b. Determine the Thévenin equivalent network seen by R_o. Apply voltage division and use the result to compute the power dissipated by R_o. Assume $R_o = 500\Omega$, $V_S = 12$ V, $R_1 = R_2 = R_3 = 1$ kΩ, and $R_X = 996\Omega$.

c. When R_o is replaced by an open-circuit, the Thévenin equivalent network supplies no power. What is the net power supplied by the entire Wheatstone bridge circuit when R_o is replaced by an open-circuit? Are the results the same? What do you conclude?

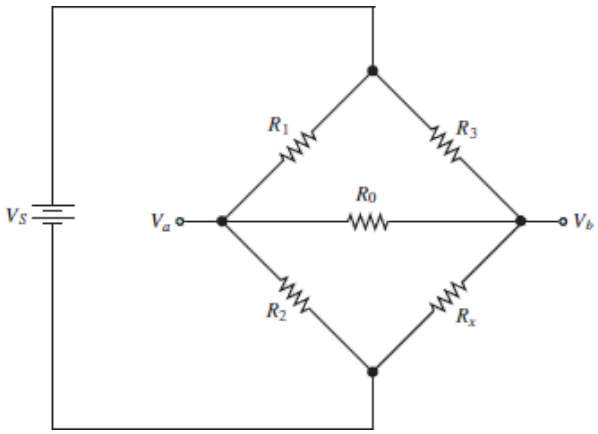


Figure P2.62

2.63 The circuit shown in Figure P2.63 is one form of a *differential amplifier*. Find an expression for the voltage drop v_{ba} from terminal b to terminal a in terms of v_1 and v_2 using Thévenin's or Norton's theorem. Notice that the figure implies zero current through sources v_1 and v_2.

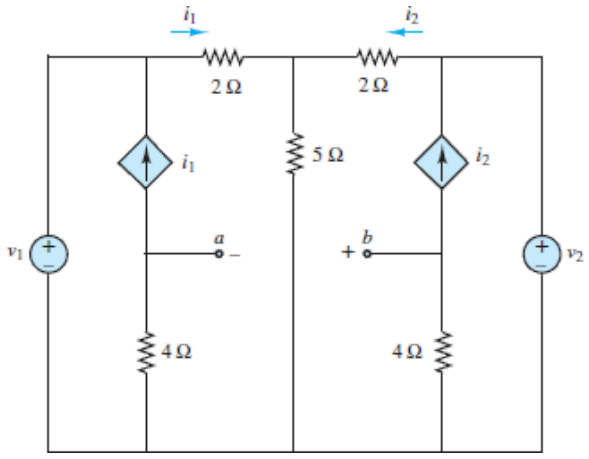


Figure P2.63

Page 156

2.64 Use source transformations to find the Thévenin equivalent network seen by resistor R_3 in the circuit of Figure P2.38. Assume $R_1 = 10\Omega$, $R_2 = 4\Omega$, $R_3 = R_4 = 6\Omega$, $V_1 = 2$ V, $V_2 = 4$ V and $I1 = 2$ A.

2.65 Find the Thévenin equivalent network seen by resistor R_4 in the circuit of Figure P2.33. Assume $R_1 = 2\Omega$, $R_2 = 8\Omega$, $R_3 = 12\Omega$, $R_4 = 8\Omega$, $V_1 = 12$ V and $V_2 = 10$ V.

2.66 Find the Thévenin equivalent network seen from node a to node b in Figure P2.66. Let $R_1 = 10\Omega$, $R_2 = 8\Omega$, $R_3 = 5\Omega$, $R_4 = 4\Omega$, $R_5 = 1\Omega$, $VS = 10$ V, $i_S = 2$ A.

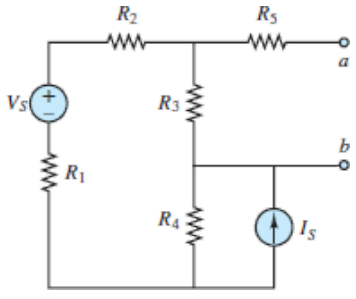


Figure P2.66

2.67 Find the Thévenin equivalent network seen by R_3 in Figure P2.40. Compute the Thévenin (open-circuit) voltage V_T in terms of the temperature T. Use that result to determine the temperature when R_3 is attached to that network.

2.68 Find the Norton equivalent network seen by R_5 in Figure P2.68. Use it and current division to compute the current through R_5. Assume $R_1 = 15\Omega$, $R_2 = 8\Omega$, $R_3 = 4\Omega$, $R_4 = 4\Omega$, $R_5 = 2\Omega$, $I1 = 2$ A, $I2 = 3$ A.

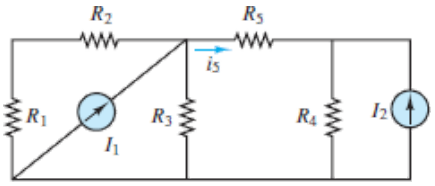


Figure P2.68

2.69 Find the Thévenin equivalent network seen by R in Figure P2.69. Use it and voltage division to compute the magnitude of the voltage across R. Assume:

$I_B = 12$ A $\quad R_B = 1\Omega$
$V_G = 12$ V $\quad R_G = 0.3\Omega$
$R = 0.23\Omega$

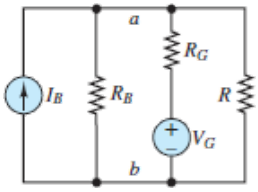


Figure P2.69

2.70 Find the Norton equivalent network between terminals a and b in Figure P2.70. Let $R_1 = 6\Omega$, $R_2 = 3\Omega$, $R_3 = 2\Omega$, $R_4 = 2\Omega$, $Vs = 10$ V, $i_S = 3$ A.

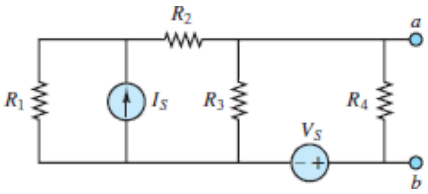


Figure P2.70

2.71 Find the Norton equivalent network seen by R_4 in Figure P2.71. Use it and current division to determine the current through R_4. Assume $R_1 = 8\Omega$, $R_2 = 5\Omega$, $R_3 = 4\Omega$, $R_4 = 3\Omega$, $Vo = 10$ V, and $Io = 2$ A.

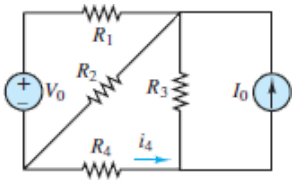


Figure P2.71

Section 2.8: Maximum Power Transfer

2.72 The Thévenin equivalent network seen by a load R_o is depicted in Figure P2.72. Assume $V_T = 10$ V, $R_T = 2\Omega$, and that the value of R_o is such that maximum power is transferred to it. Determine:

a. The value of R_o.

b. The power P_o dissipated by R_o.

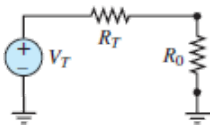


Figure P2.72

Page 157

2.73 The Thévenin equivalent network seen by a load R_o is depicted in Figure P2.72. Assume $V_T = 25$ V, $R_T = 100\Omega$, and that the value of R_o is such that maximum power is transferred to it. Determine:

a. The value of R_o.

b. The power P_o dissipated by R_o.

Section 2.9: Practical Voltage and Current Sources

2.74 A practical voltage source is modeled in Figure P2.74 as an ideal source V_S in series with a resistance R_S. This model accounts for internal power losses found in a real voltage source. The following data characterizes the real (nonideal) source:

When $R \to \infty$ $\quad V_R = 20\ \text{V}$

When $R = 2.7\ \text{k}\Omega$ $\quad V_R = 18\ \text{V}$

Determine the internal resistance R_S and the ideal voltage V_S.

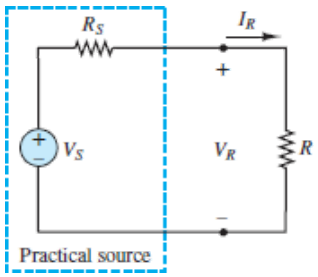


Figure P2.74

2.75 A practical voltage source is modeled in Figure P2.74 as an ideal source V_S in series with a resistance R_S. This model accounts for internal power losses found in a real voltage source. A load R is connected across the terminals of the model. Assume:

$$V_S = 12\ \text{V} \qquad R_S = 0.3\ \Omega$$

Plot the power dissipated in the load as a function of the load resistance. What can you conclude?

2.76 Consider NiMH hobbyist batteries depicted in Figure P2.76.

a. If $V_1 = 12.0$ V, $R_1 = 0.15\Omega$, and $R_o = 2.55\Omega$, find the load current I_o and the power dissipated by the load.

b. If battery 2 with $V_2 = 12$ V and $R_2 = 0.28\Omega$ is placed in parallel with battery 1, apply source transformations to determine whether the load current I_o will increase or decrease. Will the power dissipated by the load increase or decrease? By how much?

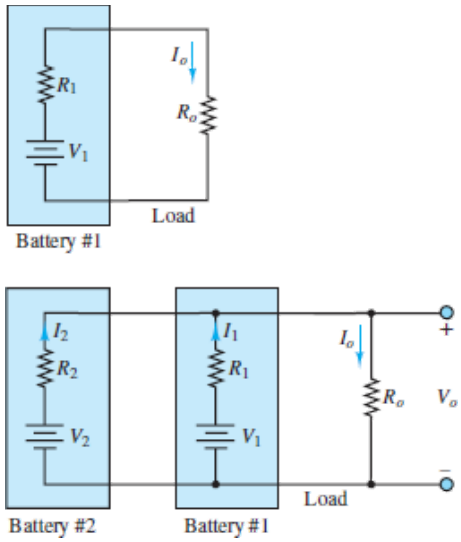


Figure P2.76

Section 2.10: Measurement Devices

2.77 A *thermistor* is a nonlinear device that changes its terminal resistance value as its surrounding temperature changes. The resistance and temperature generally have a relation in the form of:

$$R_{\text{th}}(T) = R_0 e^{-\beta(T-T_0)}$$

where R_{th} = resistance at temperature T, Ω
R_0 = resistance at temperature $T_0 = 298$ K, Ω
β = material constant, K^{-1}
T, T_0 = absolute temperature, K

a. If $R_0 = 300\Omega$ and $\beta = -0.01\text{K}-1$, plot $R_{\text{th}}(T)$ as a function of the surrounding temperature T for $350 \leq T \leq 750$.

b. If the thermistor is in parallel with a 250-Ω resistor, find the expression for the equivalent resistance and plot $R_{\text{th}}(T)$ on the same graph for part a.

2.78 A moving-coil meter movement has a meter resistance $r_M = 200\ \Omega$, and full-scale deflection is caused by a meter current $i_m = 10\mu\text{A}$. The meter is to be used to display pressure, as measured by a sensor, up to a maximum of 100 kPa. Models of the meter and pressure sensor are shown in Figure P2.78 Page 158along with the relationship between measured pressure and the sensor output v_o.

a. Devise a circuit that will produce the desired behavior of the meter, showing all appropriate connections between the terminals of the sensor

and the meter.

b. Determine the value of each component in the circuit.

c. What is the linear range, that is, the minimum and maximum pressure that can accurately be measured?

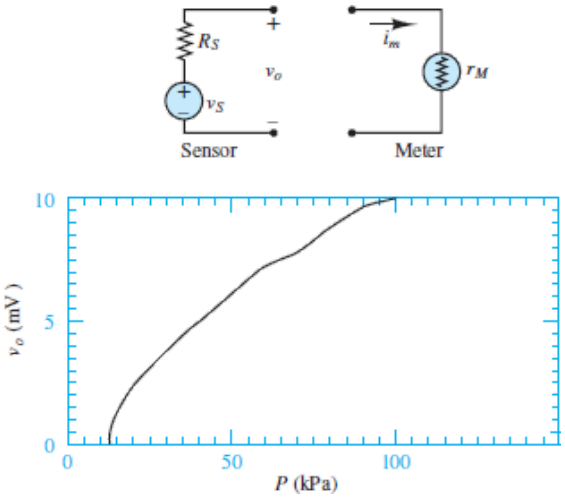


Figure P2.78

2.79 A circuit that measures the internal resistance of a practical ammeter is shown in Figure P2.79, where R_S = 50,000Ω, v_S = 12 V, and R_p is a variable resistor that can be adjusted at will.

a. Assume that $r_a \ll$ 50,000Ω. Estimate the current i.

b. If the meter displays a current of 150 μA when R_p = 15Ω, find the internal resistance of the meter r_a.

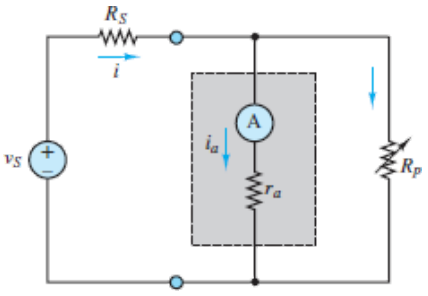


Figure P2.79

2.80 A practical voltmeter has an internal resistance r_m. What is the value of r_m if the meter reads 11.81 V when connected as shown in Figure P2.80? Assume V_S =

12 V and $R_S = 25$ kΩ.

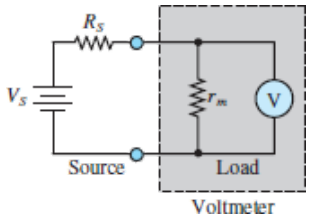


Figure P2.80

2.81 Using the circuit of Figure P2.80, find the voltage that the meter reads if $V_S = 24$ V and R_S has the following values: $R_S = 0.2r_m$, 0.4 r_m,0.6 r_m, 1.2 r_m, 4 r_m, 6 r_m, and 10 r_m. How large (or small) should the internal resistance of the meter be relative to R_S?

2.82 A voltmeter is used to determine the voltage across a resistive element in the circuit of Figure P2.82. The instrument is modeled by an ideal voltmeter in parallel with a 120-kΩ resistor, as shown. The meter is placed to measure the voltage across R_4. Assume $R_1 = 8$ kΩ, $R_2 = 22$ kΩ, $R_3 = 50$ kΩ, $R_S = 125$ kΩ, and $i_S = 120$ mA. Find the voltage across R_4 with and without the voltmeter in the circuit for the following values:

a. $R_4 = 100\Omega$

b. $R_4 = 1$ kΩ

c. $R_4 = 10$ kΩ

d. $R_4 = 100$ kΩ

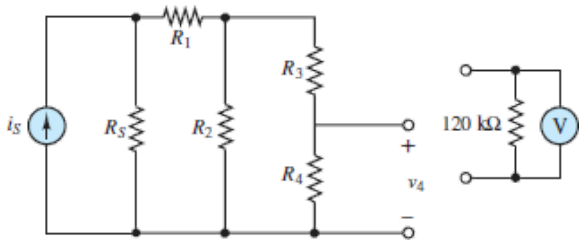


Figure P2.82

2.83 An ammeter is used as shown in Figure P2.83. The ammeter model consists of an ideal ammeter in series with a resistance. The ammeter model is placed in the branch as shown in the figure. Find the current through R_5 both with and

without the ammeter in the circuit for the following values, assuming that Page 159$R_S = 20\Omega$, $R_1 = 800\Omega$,$R_2 = 600\ \Omega$, $R_3 = 1.2\ \text{k}\Omega$, $R_4 = 150\Omega$, and $v_S = 24$ V.

a. $R_5 = 1\text{k}\Omega$

b. $R_5 = 100\Omega$

c. $R_5 = 10\Omega$

c. $R_5 = 1\Omega$

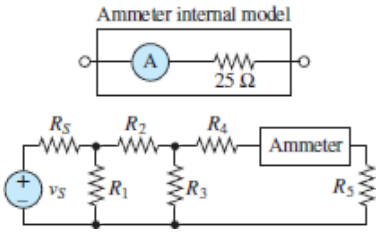


Figure P2.83

2.84 Figure P2.84 shows an *aluminum* cantilevered beam loaded by the force F. Strain gauges R_1, R_2, R_3, and R_4 are attached to the beam as shown in Figure P2.84 and connected into the circuit shown. The force causes a tension stress on the top of the beam that causes the length (and therefore the resistance) of R_1 and R_4 to increase and a compression stress on the bottom of the beam that causes the length (and therefore the resistance) of R_2 and R_3 to decrease. The result is a voltage of 50 mV at node B with respect to node A. Determine the force if

$$R_o = 1\ \text{k}\Omega \qquad v_S = 12\ \text{V} \qquad L = 0.3\ \text{m}$$
$$w = 25\ \text{mm} \qquad h = 100\ \text{mm} \qquad Y = 69\ \text{GN/m}^2$$

(See Focus on Measurements: The Wheatstone Bridge and Force Measurements.)

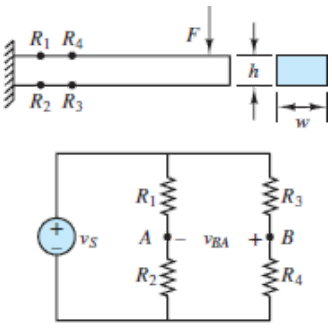


Figure P2.84

2.85 Refer to Figure P2.84 but assume that the cantilevered beam loaded by a force F is made of *steel*. Strain gauges R_1, R_2, R_3, and R_4 are attached to the beam and connected in the circuit shown. The force causes a tension stress on the top of the beam that causes the length (and therefore the resistance) of R_1 and R_4 to increase and a compression stress on the bottom of the beam that causes the length (and therefore the resistance) of R_2 and R_3 to decrease. The result is a voltage v_{BA} across nodes B and A. Determine this voltage if $F = 1.3$ MN and

$$R_o = 1\,\text{k}\Omega \qquad v_S = 24\,\text{V} \qquad L = 1.7\,\text{m}$$
$$w = 3\,\text{cm} \qquad h = 7\,\text{cm} \qquad Y = 200\,\text{GN/m}^2$$

(See Focus on Measurements: The Wheatstone Bridge and Force Measurements.)

Section 2.11: Nonlinear Circuit Elements

2.86 Apply nodal analysis to find two equations in terms of the node voltages v_1 and v_2 shown in Figure P2.86. The two nonlinear resistors R_a and R_b are characterized by:

$$i_a = 2v_a^3$$
$$i_b = v_b^3 + 10v_b$$

The resulting nonlinear (but not transcendental) equations cannot be solved by the methods used for simultaneous *linear* equations. While the equations can be solved analytically, consult your instructor before attempting to solve these equations.

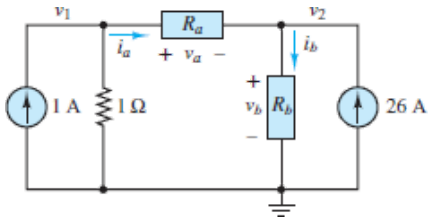


Figure P2.86

2.87 Many practical circuit elements are nonlinear; however, it is usually possible to linearize the *V-I* relationship near any specific point on the nonlinear *V-I* curve. Such a point is often referred to as an *operating point.* In other words, in the vicinity of an operating point $[V_0, I_0]$ the *V-I* relationship can be linearly approximated by:

$$I = mV + b \qquad \text{where} \quad m = \text{slope} \quad b = \text{intercept}$$

The inverse of the slope m at the operating point is defined as incremental resistance R_{inc}:

$$R_{inc} = \left.\frac{dV}{dI}\right|_{[V_0, I_0]} \approx \left.\frac{\Delta V}{\Delta I}\right|_{[V_0, I_0]}$$

Page 160

a. Refer to Figure P2.87 and find the operating point of the nonlinear element.

b. Find the incremental resistance of the nonlinear element at the operating point.

c. If V_T is increased to 20 V, what is the new operating point and the new incremental resistance?

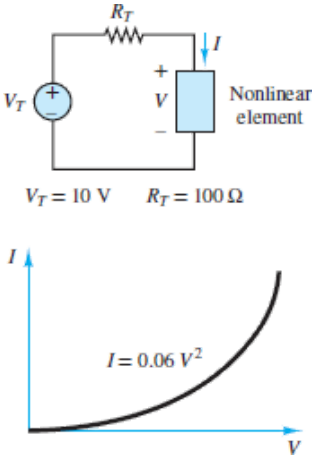


Figure P2.87

2.88 The device D in the circuit in Figure P2.88 is an induction motor with a nonlinear i-v characteristic. Determine the current through and the voltage across the motor.

$$V_S = 450\ \text{V} \qquad R = 9\ \Omega$$

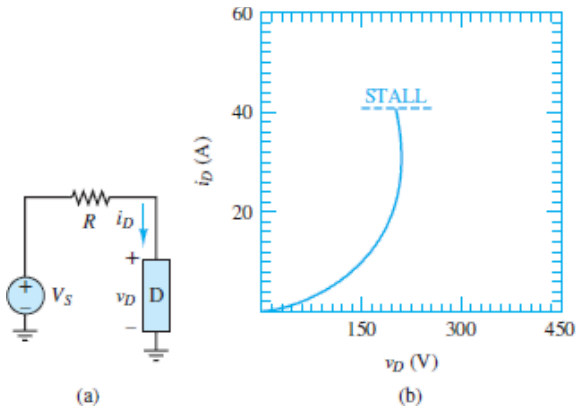


Figure P2.88

2.89 The nonlinear diode in Figure P2.89 has the *i-v* characteristic shown. Assume:

$$V_S = V_{TH} = 1.5\,\text{V} \qquad R = R_{eq} = 60\,\Omega$$

Determine the voltage across and the current through the diode.

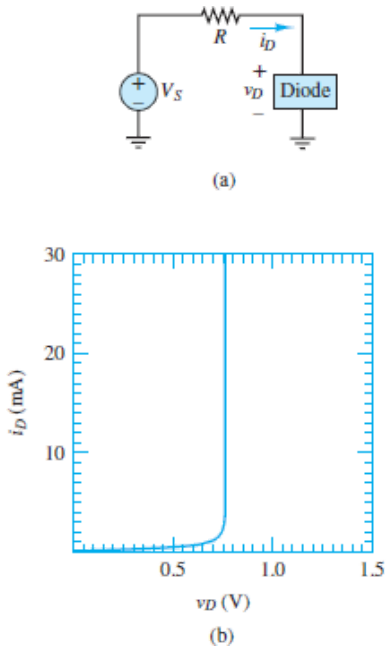


Figure P2.89

2.90 The resistance of the device *D* in Figure P2.90 is a nonlinear function of pressure *P*. The *i-v* characteristics of *D* are shown for various pressures. Assume:

$$V_S = 2.5\,\text{V} \qquad R = 125\,\Omega$$

a. Plot the DC load line.

b. Plot the voltage across D as a function of pressure.

c. Determine the current through D when $P = 30$ psig.

Page 161

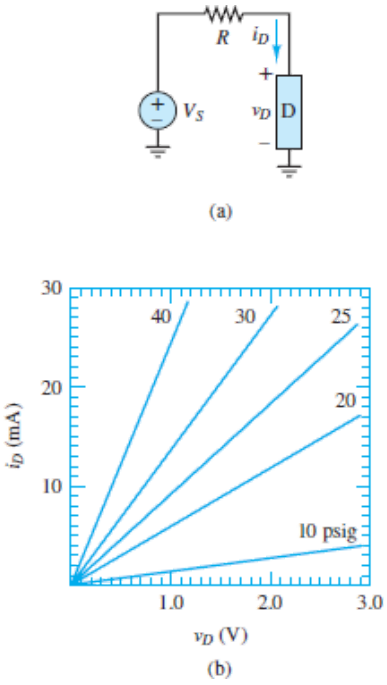


Figure P2.90

2.91 The nonlinear device D in Figure P2.91 has the following transcendental i-v characteristic:

$$i_D = I_0 e^{v_D/V_{thermal}}$$

where

$I_o = 10^{-10}$ A and $V_{thermal} = 25$ mV

Assume that $V_S = 2$ V and $R = 40\Omega$. Determine an expression for the DC load line. Then use an iterative technique to determine the voltage across and current through D.

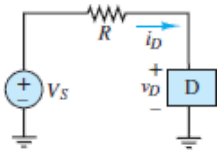


Figure P2.91

2.92 The resistance of the device D in Figure P2.90 is a nonlinear function of pressure P. The i-v characteristics of D are shown for various pressures. Assume:

$$V_S = 3.0\,\text{V} \qquad R = 100\,\Omega$$

Construct the DC load line and determine the current through D when P = 40 psig.

2.93 The so-called forward-bias i-v relationship for a silicon diode is:

$$i_D = I_{\text{SAT}}\left[e^{(v_D/V_{\text{thermal}})} - 1\right]$$

where I_{SAT} and V_{thermal} are known as the *saturation current* and *thermal voltage,* respectively. At room temperature (20°C):

$$I_{\text{SAT}} = 10^{-12}\,\text{A} \quad \text{and} \quad V_{\text{thermal}} = \frac{kT}{q} = 25.3\,\text{mV}$$

where k is Boltzmann's constant, T is absolute temperature in kelvins, and q is the charge of an electron.

Consider the circuit shown in Figure P2.93. KVL applied around the loop results in a transcendental equation for the loop current $i = i_D$. Such equations cannot be solved in terms of a closed-form expression i = …. Instead, graphical or iterative procedures must be used.

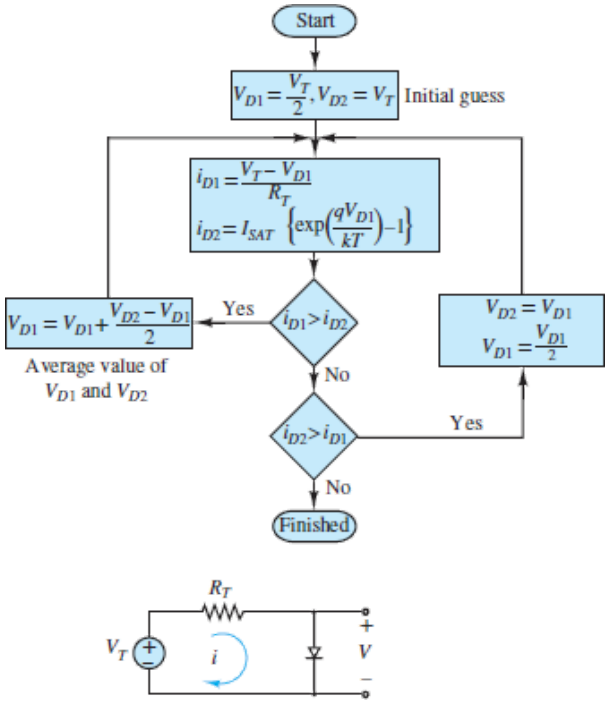


Figure P2.93

Page 162

a. Use graphical analysis to estimate the current through and the voltage across the diode. Assume $R_T = 22\Omega$ and $V_T = 12$ V.

b. Use the iterative algorithm depicted in the flowchart of Figure P2.93 to construct a computer program that solves for V and i. The algorithm relies upon the fact that $0 < V < VT$ to make an initial guess $V_{D1} = V_T/2$ for the voltage across the diode. The algorithm then determines whether the current i_{D1} through R_T is greater than, less than, or equal to the diode current i_{D2} for the guessed diode voltage. In the first case, a new guess for V_{D1} is set equal to the average value of V_{D1} and V_{D2}, which stores the most recent value of V_{D1} that resulted in $i_{D2} > i_{D1}$. The initial value $V_{D2} = V_T$ guarantees that $i_{D1} = 0$ and, thus, $i_{D2} > i_{D1}$ for the first pass through the iterative algorithm. The result is that V_{D1} and V_{D2} bracket the actual value of V. Each pass through the algorithm narrows the bracket until the difference $|i_{D1} - i_{D2}| < \varepsilon$, where ε is some sufficiently small error term.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for “Focus on measurements” weighing scales: Media Bakery.

C H A P T E R

3

AC NETWORK ANALYSIS

Chapter 3 introduces capacitors and inductors, which are energy storage elements, and methods for solving circuits that contain them. This chapter also introduces AC circuits, which contain time-dependent sinusoidal voltage and current sources, as opposed to DC circuits, which contain constant sources only. Solutions of AC circuits containing capacitors and/or inductors result in differential equations because the *i*-*v* relationships for capacitors and inductors involve time derivatives. Luckily, the method of phasor analysis can be used to convert differential equations into algebraic equations, which are much easier to solve. However, the price ("there is no such thing as a free lunch") of using phasor analysis is that the algebraic equations contain complex quantities, which must be added, subtracted, multiplied, and divided. (Most calculators can perform these operations.) More importantly, it is necessary to understand the meaning of and relationships among complex quantities. With some practice and patience even those students with no prior experience using complex numbers can become proficient with phasor analysis.

Sinusoids are an especially important class of signals for two reasons. First, nearly all residential and industrial electric power is generated, transmitted, and distributed as a sinusoidal waveform. All turbine-based power systems (e.g., coal-fired power stations, solar power arrays, hydroelectric dams, wind turbines) produce periodic rotating motion, which is represented mathematically by a sinusoid. Second,

all periodic waveforms (e.g., sawtooth, triangle, square) can be reconstructed as the sum of component sinusoidal waves (Fourier's theorem).

Page 164

Sinusoidal signals (voltages and currents) have three basic characteristics: amplitude (or magnitude), frequency, and phase (or phase angle). In this book, an AC circuit has one frequency shared by all voltages and currents that is equal to the frequency of an independent voltage or current source. As a result, it is not necessary to compute frequency in phasor analysis. On the other hand, while the amplitude and phase of voltages and currents in an AC circuit are constant, they are not uniform and are determined not only by an independent source but also by the elements present in the circuit. Consequently, AC circuit analysis is concerned with the computation of the amplitude and phase of one or more voltages and currents. (DC circuit analysis was only concerned with amplitude.) Phasor analysis is well suited for AC circuit analysis because a phasor bundles together amplitude and phase into a single quantity.

In phasor analysis, resistors, capacitors, and inductors are represented as impedance elements, which allows Ohm's law to be generalized. Kirchhoff's laws can also be generalized as phasor relationships. Consequently, AC circuits can be solved using the same DC methods (e.g., node voltage, mesh current, voltage division, current division, superposition, Thévenin's and Norton's theorem, and source transformations) discussed in Chapters 1 and 2. The only difference is that these relationships now involve phasors, that is, complex quantities.

The average and effective (root-mean-square) amplitude of a waveform are introduced in this chapter. An effective value represents the equivalent DC value required to supply or dissipate the same power as the AC waveform and thus provides a means of comparing different waveforms. This rather extensive chapter concludes with an introduction to single-phase AC power and the concepts of power factor, apparent, real and reactive power, power triangles and power factor correction.

In this chapter and throughout the book, angles are given in units of radians, unless indicated otherwise.

Learning Objectives

Students will learn to...

1. Compute current, voltage, and energy of capacitors and inductors. *Section 3.2*.

2. Calculate the average and effective (root-mean-square) value of an arbitrary periodic waveform. *Section 3.3*.
3. Convert time-domain sinusoidal voltages and currents to phasor notation, and vice versa; and represent circuits using impedances. *Sections 3.4* and *3.5*.
4. Apply DC circuit analysis methods to AC circuits in phasor form. *Section 3.6*.
5. Compute average AC power and the power factor of a complex load. *Section 3.7*.
6. Compute apparent, real and reactive power for complex loads and draw a power triangle. *Section 3.8*.
7. Compute the capacitance required to correct the power factor of a complex load *Section 3.9*.

3.1 CIRCUITS CONTAINING ENERGY STORAGE ELEMENTS

The resistive circuits studied in Chapters 1 and 2 had no dependence on time. The sources had constant (DC) values and the *i*-*v* relationship for resistors (Ohm's law) had no time dependence. As a result, all the equations obtained in those chapters Page 165were algebraic and the voltages and currents were all constants. If a sinusoidal source is present in a resistive circuit, the voltages and currents in the circuit will no longer be constant but instead will vary sinusoidally in time with the same frequency as the source. A circuit with a sinusoidal source is known as an **AC** circuit.

Purely resistive AC circuits offer no new challenges compared to DC circuits. However, when capacitors and/or inductors are introduced into an AC circuit, such as that shown in Figure 3.1, the resulting behavior is significantly more interesting and challenging. The reason is that capacitors and inductors require time to charge (store energy) and discharge (release energy). The result, in general, is that time delays are present in an AC circuit, as shown in Figure 3.2. These time delays are expressed as *angles* known as *phase shifts*. Consequently, in the solution of AC circuits it is necessary to keep track of two variables (amplitude and phase) for each voltage and current. By contrast, when solving DC circuits, it was only necessary to keep track of one variable (amplitude).

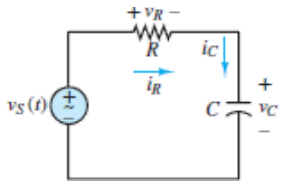


Figure 3.1 A simple RC series loop with a sinusoidal voltage source. The resulting voltages v_R and v_C and currents i_R and i_C are also sinusoidal and shifted in time with respect to the voltage source.

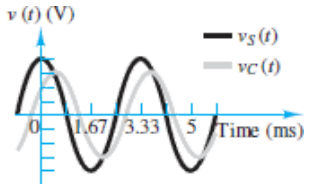


Figure 3.2 Waveforms for the AC circuit of Figure 3.1

To clarify this discussion, consider the simple series loop shown in Figure 3.1, which consists of a sinusoidal voltage source, a resistor, and a capacitor. Apply KVL around the loop to obtain:

$$v_S - v_R - v_C = 0 \qquad \text{or} \qquad v_R + v_C = v_S \tag{3.1}$$

The so-called *state variable* for this circuit is the voltage v_C across the capacitor. The state variables in a circuit are the voltages across capacitors and the currents through inductors.

The constitutive *i-v* relationships for the resistor and capacitor are

$$v_R = i_R R \qquad \text{and} \qquad i_C = C\frac{dv_C}{dt} \tag{3.2}$$

The resistor and capacitor currents are the same for this simple loop. Thus:

$$v_R = i_R R = i_C R = RC\frac{dv_C}{dt} \tag{3.3}$$

Plug this result into Equation 3.1 to obtain:

$$RC\frac{dv_C}{dt} + v_C = v_S \tag{3.4}$$

Divide both sides of Equation 3.4 by *RC* to find the standard form:

$$\frac{dv_C}{dt} + \frac{1}{RC} v_C = \frac{1}{RC} v_S \tag{3.5}$$

The result is a first-order, linear, ordinary differential equation. The solution for v_C has two parts: (1) a transient solution, and (2) a steady-state solution. The complete solution of the differential equation is the sum of these two parts.

The differential equation for v_R has a similar form.

$$\frac{dv_R}{dt} + \frac{1}{RC} v_R = \frac{dv_S}{dt} \tag{3.6}$$

Notice that the left-hand side of this equation is identical to that found in equation 3.5. Only the right-hand side is different. The constant *RC* has units of time and is known as a *time constant.*

Page 166

For more complicated circuits, the process is largely the same except that KVL and KCL may have to be applied multiple times and the circuit may contain multiple resistors, capacitors, and inductors. The result is multiple first- and perhaps second-order linear, ordinary differential equations. It is not difficult to imagine that for even modest circuits the procedure and results may become quite complicated and cumbersome.

To avoid these complications, an alternative approach is to dispense with time derivatives and solve for the steady-state (particular) and transient (homogeneous) solutions separately using the following two methods:

- *Steady-state solution.* To solve for the steady-state solution, Euler's formula is employed to represent sinusoids as complex exponentials and to eliminate the time derivatives in the constitutive *i-v* relations for capacitors and inductors. The result is algebraic equations with complex constants and variables. These equations can be solved using standard algebra techniques. The only complication is that the arithmetic involves complex numbers rather than real numbers.
- *Transient solution*. Whenever possible, Thévenin's and Norton's theorems are used to simplify complicated circuits and to focus on solving for the state variables. Solutions of the simplified first- and second-order circuits are found through well-established methods that require only a modest understanding of differential equations. These transient solutions are covered in chapter 4.

In a linear circuit with a sinusoidal source, all voltages and currents are sinusoids at the same frequency as the source. These voltages and currents are *scaled* versions of the source and may be shifted in time (i.e., phase shifted).

3.2 CAPACITORS AND INDUCTORS

The ideal resistor was introduced in Chapter 1. In addition to resistance, which always dissipates energy, an electric circuit may also exhibit capacitance and inductance, which act to store and release energy, in the same way that an expansion tank and flywheel, respectively, act in a mechanical system. These two distinct energy storage mechanisms are represented in electric circuits by two ideal circuit elements: the ideal capacitor and the ideal inductor, which approximate the behavior of actual discrete capacitors and inductors. They also approximate the bulk properties of capacitance and inductance that are present in any physical system. In practice, any element of an electric circuit will exhibit some resistance, some inductance, and some capacitance, that is, some ability to dissipate and store energy.

The energy of a capacitor is stored within the electric field between two conducting plates while the energy of an inductor is stored within the magnetic field of a conducting coil. Both elements can be charged (i.e., stored energy is increased) or discharged (i.e., stored energy is decreased). Ideal capacitors and inductors can store energy indefinitely; however, in practice, discrete capacitors and Page 167inductors exhibit “leakage,” which typically results in a gradual reduction in the stored energy over time.

All of the relationships for capacitors and inductors exhibit duality, which means that the capacitor relations mirror the inductor relations. Examples of duality are apparent in Table 3.1, where C is capacitance and L is inductance.

Table 3.1 Properties of capacitors and inductors

	Capacitors	Inductors
Differential i-v	$i = C\frac{dv}{dt}$	$v = L\frac{di}{dt}$
Integral i-v	$v_C(t) = \frac{1}{C}\int_{-\infty}^{t} i_C(\tau)\,d\tau$	$i_C(t) = \frac{1}{L}\int_{-\infty}^{t} v_L(\tau)\,d\tau$
DC equivalent	Open-circuit	Short-circuit
Two in series	$C_{eq} = \frac{C_1 C_2}{C_1 + C_2}$	$L_{eq} = L_1 + L_2$
Two in parallel	$C_{eq} = C_1 + C_2$	$L_{eq} = \frac{L_1 L_2}{L_1 + L_2}$
Stored energy	$W_C = \frac{1}{2}Cv_C^2$	$W_L = \frac{1}{2}Li_L^2$

MAKE THE CONNECTION

Hydraulic Analog of a Capacitor

If the walls of a vessel have some elasticity, energy is stored in the walls when the vessel is filled by a fluid or gas (e.g., an inflated balloon). The ratio of the mass of the fluid or gas to the *potential energy* stored in the walls per unit mass is the **fluid capacitance** of the vessel, a property similar to **electrical capacitance**. Figure 3.3 depicts a gas bag accumulator, such as an expansion tank attached to the hot water line in many residential homes. The two-chamber arrangement permits fluid to displace a membrane separating an incompressible fluid (e.g., water) from a compressible fluid (e.g., air). The analogy shown in Figure 3.3 assumes that the reference pressure p_0 and the reference voltage v_2 are both zero.

$$q_f = C_f \frac{d\Delta p}{dt} = C_f \frac{dp}{dt}$$
$$i = C\frac{d\Delta v}{dt} = C\frac{dv_1}{dt}$$

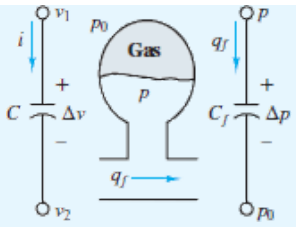


Figure 3.3 Analogy between electrical and fluid capacitance

The Ideal Capacitor

A capacitor is a device that can store energy due to a charge separation. In general, a capacitor (and thus, capacitance) is present when any two conducting surfaces are separated by a distance. A simple example is two parallel plates of shared cross-sectional area A separated by a distance d. The gap between the plates may be a vacuum or filled with some dielectric material, such as air, mica, or Teflon. The impact of the dielectric material on the capacitance is represented by the dielectric constant κ.[1] Figure 3.4 depicts a typical configuration and the circuit symbol for a capacitor.

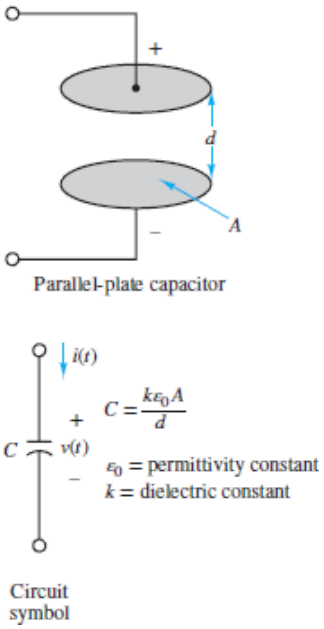


Figure 3.4 Structure of parallel-plate capacitor

The capacitance C of an *ideal* parallel-plate capacitor such as the one described above is:

$$C = \frac{\kappa \varepsilon_0 A}{d} \tag{3.7}$$

where $\varepsilon_0 = 8.85 \times 10^{-12}\, F/m$ is the permittivity constant of a vacuum.

The presence of a dielectric or vacuum between the conducting plates does not permit charge to pass directly from one plate to the other. However, although no charge can literally pass from one plate of an ideal capacitor directly through to the other, charge can exit one plate and enter the other through pathways in the circuit to which the capacitor is attached. The result is the equivalent effect of a current through the capacitor.

At all times the charge separation q_C is proportional to the applied voltage V_C,

$$q_C = CV_C \tag{3.8}$$

Page 168

where the capacitance C is a measure of the ability of the device to accumulate charge. The unit of capacitance is coulomb per volt, or farad (F). The farad is an impractically large unit for many common electronic applications; units of microfarads (1 μ F = 10^{-6} F) and picofarads (1 pF = 10^{-12} F) are more common in practice.

The current through a capacitor is defined as the time rate of change of its stored charge.

$$i_C(t) = \frac{dq_C(t)}{dt} \tag{3.9}$$

The *i-v* relationship for a capacitor is obtained by differentiating both sides of equation 3.8 and plugging the result into equation 3.9.

$$i_C(t) = C\frac{dv_C(t)}{dt} \qquad i\text{-}v \text{ relation for capacitor} \tag{3.10}$$

One immediate implication of equation 3.10 is that the current through a capacitor in a DC circuit is zero. Why? Since the voltage across a capacitor in a DC circuit must, by definition, be constant, the time derivative of the voltage must be zero. Thus, equation 3.10 requires the current through the capacitor to also be zero.

A capacitor in a DC circuit is equivalent to an open-circuit.

Equation 3.10 can be integrated to yield an expression for the change in voltage across a capacitor due to the current through it.

$$\Delta v_C = \frac{1}{C}\int^{t} i_C(\tau)\,d\tau \tag{3.11}$$

Equation 3.11 indicates that the change in voltage across a capacitor depends upon the accumulation of charge over time, which is represented as the time integral of the capacitor current. To calculate the voltage at a specific time it is necessary to know the voltage V_0 across the capacitor at some previous time t_0.

$$v_C(t) = V_0 + \frac{1}{C}\int_{t_0}^{t} i_C(\tau)\,d\tau \qquad t \geq t_0 \tag{3.12}$$

Equivalent Capacitance

Just as resistors in series or parallel can be represented by an equivalent resistance, so capacitors in series or parallel can be represented by an equivalent capacitance. For two capacitors in series and parallel, the equivalent capacitances are, respectively:

$$C_{eq} = \frac{C_1 C_2}{C_1 + C_2} \qquad \text{and} \qquad C_{eq} = C_1 + C_2 \tag{3.13}$$

Notice that the rule for the equivalent capacitance of two capacitors in series is the product divided by the sum, which is the same rule used for two resistors in parallel. Page 169Likewise, the equivalent capacitance of two capacitors in parallel is simply the sum of the two, which is the same rule used for two resistors in series. More general rules are illustrated in Figure 3.5.

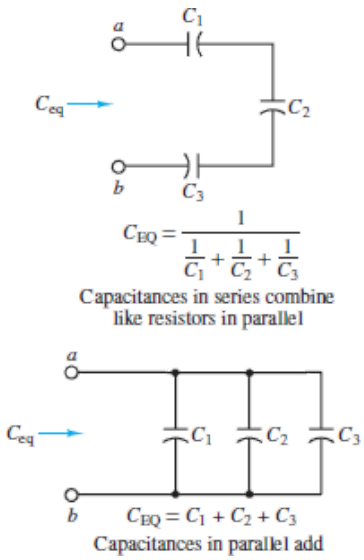


Figure 3.5 Equivalent capacitance in a circuit

LO

When calculating equivalent capacitance, capacitors in series combine like resistors in parallel and capacitors in parallel combine like resistors in series.

Discrete Capacitors

Actual capacitors are rarely constructed of two parallel plates separated by air because this configuration either yields very low values of capacitance or requires very large plate areas. To increase capacitance, physical capacitors are often made of tightly rolled sheets of metal film, with a dielectric (e.g., paper or Mylar) sandwiched in between. Table 3.2 illustrates typical values, materials, maximum voltage ratings, and useful frequency ranges for various types of capacitors.

Table 3.2 Capacitors

Material	Capacitance range	Maximum voltage (V)	Frequency range (Hz)
Mica	1 pF to 0.1 μF	100–600	10^3–10^{10}
Ceramic	10 pF to 1 μF	50–1,000	10^3–10^{10}
Mylar	0.001 μF to 10 μF	50–500	10^2–10^8
Paper	1,000 pF to 50 μF	100–10,000	10^2–10^8
Electrolytic	0.1 μF to 0.2 F	3–600	10–10^4

In practice, actual capacitors exhibit some leakage between the plates. Imperfect construction techniques invariably provide some capability for charge to pass from one plate to the other. This imperfection is often represented by an equivalent resistance in parallel with an ideal capacitor.

Energy Storage in Capacitors

The energy stored in a capacitor $W_C(t)$ is derived from the definition of energy as the time integral of power, which is the product of voltage and current:

$$P_C(t) = i_C(t)v_C(t) = C\frac{dv_C(t)}{dt}v_C(t) = \frac{d}{dt}\left[\frac{1}{2}Cv_C^2(t)\right] \tag{3.14}$$

The total energy stored in the capacitor is found by integrating the power.

$$W_C(t) = \int P_C(\tau)d\tau = \int \frac{1}{d\tau}\left[\frac{1}{2}Cv_C^2(\tau)\right]d\tau$$

LO

$$W_C = \frac{1}{2}Cv_C^2 \qquad \text{Energy stored in a capacitor} \tag{3.15}$$

Page 170

FOCUS ON MEASUREMENTS

Capacitive Displacement Transducer and Microphone

As shown in Figure 3.4, the capacitance of a flat parallel-plate capacitor is

$$C = \frac{\varepsilon A}{d} = \frac{\kappa \varepsilon_0 A}{d}$$

where ε is the **permittivity** of the dielectric material, κ is the dielectric constant, $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m is the permittivity of a vacuum, A is the area of each of the plates, and d is their separation. The dielectric constant for air is $\kappa_{\text{air}} \approx 1$. Thus, the capacitance of two flat parallel plates of area 1 m^2, separated by a 1-mm air gap is 8.854 nF, a very small value for such large plates. As a result, flat parallel-plate capacitors are impractical for use in most electronic devices. On the other hand, parallel-plate capacitors find application as motion transducers, that is, as devices that can measure the motion or displacement of an object. In a capacitive motion transducer, the plates are designed to allow relative motion when subjected to an external force. Using the capacitance value just derived for a parallel-plate capacitor, one can obtain the expression

$$C = \frac{8.854 \times 10^{-3} A}{x} \quad \text{pF}$$

where C is the capacitance in picofarads, A is the area of the plates in square millimeters, and x is the variable separation distance in millimeters. Note that the change in C due to a change in x is nonlinear, since $C \propto 1/x$. However, for small changes in x, the change in C is approximately linear.

The sensitivity S of the transducer is defined as the rate of change in capacitance C with respect to a change in separation distance x.

$$S = \frac{dC}{dx} = -\frac{8.854 \times 10^{-3} A}{x^2} \quad \frac{\text{pF}}{\text{mm}}$$

Thus, the sensitivity is itself a function of the separation distance, as shown in Figure 3.6. Note that as $x \rightarrow 0$, the slope of $C(x)$ increases and so the sensitivity S increases as well. Figure 3.6 depicts this behavior for a transducer with area equal to 10 mm^2. This type of capacitive displacement transducer is used in the popular condenser microphone, in which sound pressure waves act to deflect a thin metallic foil. The Page 171change in capacitance can then be converted to a change in voltage or current by means of a suitable circuit. An extension of this concept that permits measurement of differential pressures is shown in Figure 3.7. A three-terminal variable capacitor is made of two rigid surfaces and one thin, flexible plate, often made of steel, between them. Typically, the rigid surfaces are spherical depressions ground into glass disks and coated with a conducting material. Inlet orifices expose

the deflecting plate to the outside fluid or gas. When the pressure on both sides of the deflecting plate is the same, the capacitance between terminals b and d, denoted by C_{db}, will be equal to that between terminals b and c, denoted by C_{bc}. A pressure differential will cause the thin, flexible plate to deflect toward one of the rigid surfaces and away from the other. As a result, the two capacitances will change, with an increase on the side where the deflecting plate has come closer to the rigid surface and a corresponding decrease on the other side.

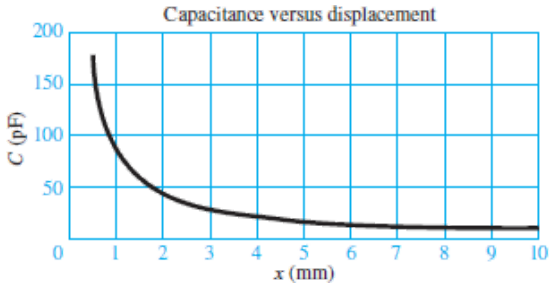


Figure 3.6 Response of a capacitive displacement transducer

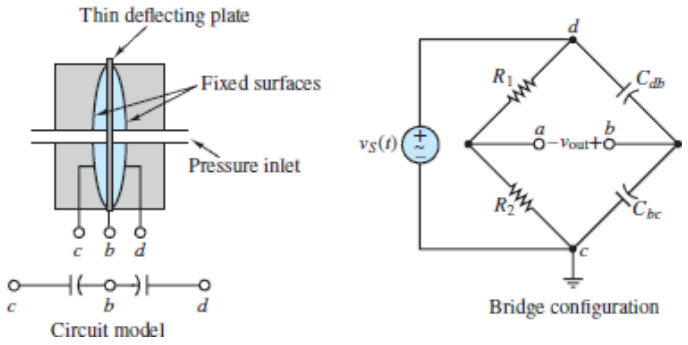


Figure 3.7 Capacitive pressure transducer and related bridge circuit

A Wheatstone bridge circuit, such as that shown in Figure 3.7, is ideally suited to set the output voltage v_{out} to zero when the differential pressure across the transducer is also zero.

LO

EXAMPLE 3.1 Charge Separation in Ultracapacitors

Problem

Ultracapacitors are finding application in a variety of fields, including as a replacement or supplement for batteries in hybrid-electric vehicles. These "supercapacitors" store energy electrostatically by polarizing an electrolytic solution. Although it is an electrochemical device (also known as an electrochemical double-layer capacitor), there are no chemical reactions involved in its energy storage mechanism. This mechanism is highly reversible, allowing the ultracapacitor to be charged and discharged hundreds of thousands of times. An ultracapacitor can be viewed as two nonreactive porous plates suspended within an electrolyte, with a voltage applied across the plates. The applied potential on the positive plate attracts the negative ions in the electrolyte while the potential on the negative plate attracts the positive ions. This effectively creates two layers of capacitive storage, one where the charges are separated at the positive plate and another at the negative plate.

Page 172

Recall that capacitors store energy in the form of separated electric charge. The greater the area for storing charge and the closer the separated charges, the greater the capacitance. A conventional capacitor gets its area from plates of a flat, conductive material. To achieve high capacitance, this material can be wound in great lengths, and sometimes a texture is imprinted on it to increase its surface area. A conventional capacitor separates its charged plates with a dielectric material, sometimes a plastic or paper film, or a ceramic. These dielectrics can be made only as thin as the available films or applied materials.

An ultracapacitor gets its area from a porous carbon-based electrode material, as shown in Figure 3.8. The porous structure of this material allows its surface area to approach 2,000 square meters per gram (m^2/g), much greater than can be accomplished using flat or textured films and plates. An ultracapacitor's charge separation distance is determined by the size of the ions in the electrolyte, which are attracted to the charged electrode. This charge separation [less than 10 angstroms (Å)] is much smaller than can be achieved using conventional dielectric materials. The combination of enormous surface area and extremely small charge separation gives the ultracapacitor its outstanding capacitance relative to conventional capacitors.

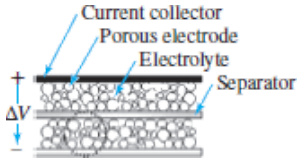


Figure 3.8 Ultracapacitor structure

Use the data provided to calculate the charge stored in an ultracapacitor and calculate how long it will take to discharge the capacitor at the maximum current rate.

Solution

Known Quantities: Technical specifications are as follows:

Capacitance	100 F	(−10%/+30%)
Series resistance	DC	15 mΩ(±25%)
	1 kHz	7 mΩ(±25%)
Voltage	Continuous	2.5 V; peak 2.7 V
Rated current	25 A	

Find: Charge separation at nominal voltage and time to complete discharge at maximum current rate.

Analysis: Based on the definition of charge storage in a capacitor, calculate

$$Q = CV = 100 \text{ F} \times 2.5 \text{ V} = 250 \text{ C}$$

To calculate how long it would take to discharge the ultracapacitor, approximate the current as:

$$i = \frac{dq}{dt} \approx \frac{\Delta q}{\Delta t}$$

Since the available charge is 250 C, the time to completely discharge the capacitor, assuming a constant 25-A discharge, is

$$\Delta t = \frac{\Delta q}{i} = \frac{250 \text{ C}}{25 \text{ A}} = 10 \text{ s}$$

To calculate the energy, use equation 3.15:

$$W_C = \frac{1}{2} C v_C^2 = \frac{1}{2}(100 \text{ F})(2.5 \text{ V})^2 = 312.5 \text{ J}$$

Page 173

EXAMPLE 3.2 Calculating Capacitor Current From Voltage

Problem

Calculate the current through a capacitor from knowledge of its terminal voltage.

Solution

Known Quantities: Capacitor terminal voltage for $t > 0$; capacitance value.

Find: Capacitor current $t > 0$.

Assumptions: None.

Schematics, Diagrams, Circuits, and Given Data: $v(t) = 5\left(1 - e^{-t/10^{-6}}\right)$ V; $t \geq 0$ s; $C = 0.1$ μF. The terminal voltage is plotted in Figure 3.9.

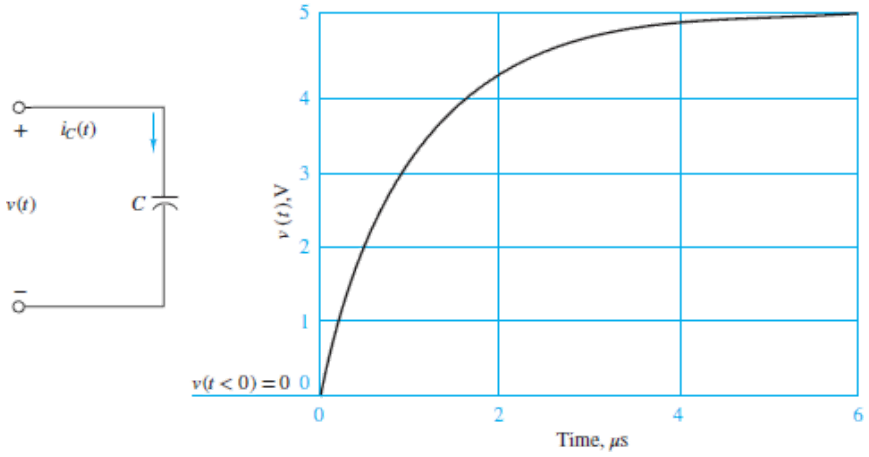


Figure 3.9

Assumptions: The capacitor is initially discharged: $v(t = 0) = 0$.

Analysis: Differentiate the voltage across the capacitor to find the current through it.

$$i_C(t) = C\frac{dv(t)}{dt} = 10^{-7}\frac{5}{10^{-6}}\left(e^{-t/10^{-6}}\right) = 0.5e^{-t/10^{-6}} \quad \text{A} \qquad t \geq 0$$

A plot of the capacitor current is shown in Figure 3.10. Note how the current jumps to 0.5 A just after $t = 0$. The ability of the current through a capacitor to change instantaneously is an important property.

Comments: As the voltage approaches the constant value 5 V, the charge stored in the capacitor approaches its maximum value and the current through the capacitor approaches zero. The total charge stored is $Q = 0.5 \times 10^{-6}$ C. This is a fairly small amount of charge, but it can produce a significant current for a brief period. For example, when fully charged the capacitor could provide 100 mA for 5 μs:

$$I = \frac{\Delta Q}{\Delta t} = \frac{0.5 \times 10^{-6}}{5 \times 10^{-6}} = 0.1\,\text{A}$$

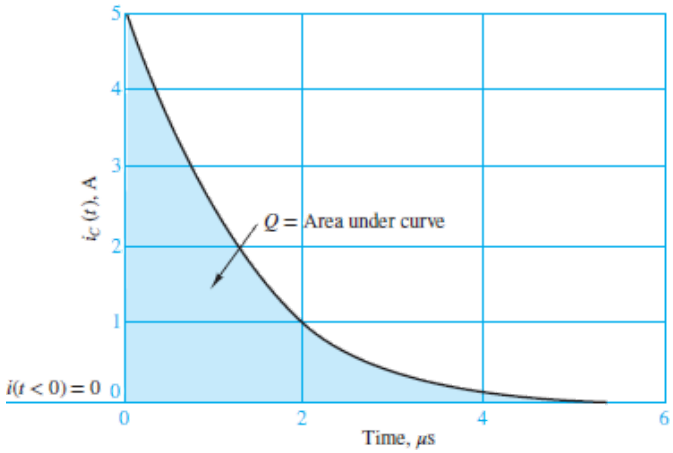


Figure 3.10

Page 174

There are many practical applications of the energy storage property of capacitors.

EXAMPLE 3.3 Calculating Capacitor Voltage From Current and an Initial Condition

Problem

Solve for the voltage across a capacitor from knowledge of its current and initial charge.

Solution

Known Quantities: Capacitor current; initial capacitor voltage; capacitance value.

Find: Capacitor voltage as a function of time.

Schematics, Diagrams, Circuits, and Given Data:

$$i_C(t) = \begin{cases} 0 & t < 0\,\text{s} \\ 10\,\text{mA} & 0 \le t \le 1\,\text{s} \\ 0 & t > 1\,\text{s} \end{cases}$$

$$v_C(t = 0) = 2\ \text{V} \qquad C = 1{,}000\ \mu\text{F}$$

The capacitor current is plotted in Figure 3.11(a).

Assumptions: The capacitor is initially charged such that $V_0 = v_C(t = 0) = 2$ V.

Analysis: The integral relationship between voltage and current for a capacitor can be used to find voltage when current is known.

$$\Delta v = \frac{1}{C}\int_0^t i_C(\tau)d\tau = 10t \qquad t \ge 0$$

$$v_C(t) = \begin{cases} 10t + 2\ \text{V} & 0 \le t \le 1\,\text{s} \\ 12\ \text{V} & t > 1\,\text{s} \end{cases}$$

Page 175

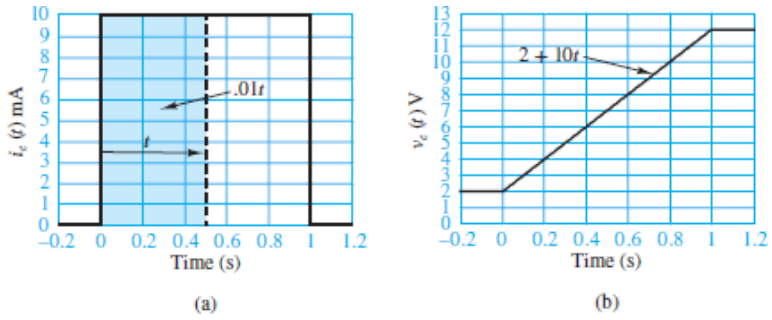


Figure 3.11

Comments: Once the current stops, at $t = 1$ s, the capacitor voltage remains constant because the charge remains constant. That is, $V = Q/C$ = constant = 12 V at $t = 1$ s. Remember, the final value of the capacitor voltage depends on two factors: (1) the initial value of the capacitor voltage and (2) the history of the capacitor current. Figure 3.11(a) and (b) depict the current through and voltage across the capacitor as functions of time.

CHECK YOUR UNDERSTANDING

Compare the energy stored in the ultracapacitor of Example 3.1 with a (similarly sized) electrolytic capacitor used in power electronics applications. Calculate the energy stored for a 2,000-μF electrolytic capacitor rated at 400 V.

Answer: 160 J

CHECK YOUR UNDERSTANDING

Compare the charge separation achieved in the ultracapacitor of Example 3.1 with a (similarly sized) electrolytic capacitor used in power electronics applications, by calculating the charge separation for a 2,000-μF electrolytic capacitor rated at 400 V.

Answer: 0.8 C

CHECK YOUR UNDERSTANDING

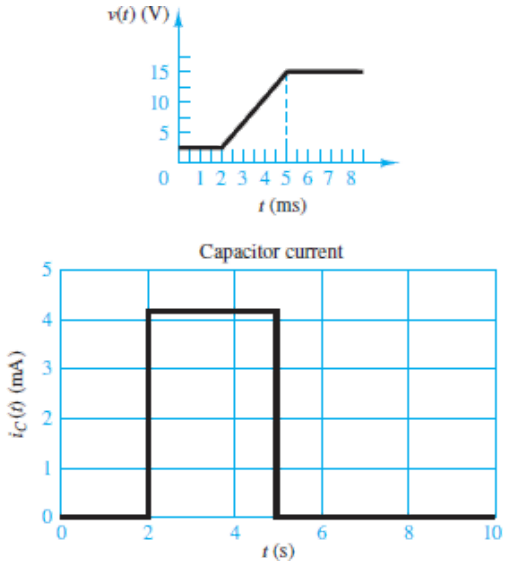
Find the maximum current through the capacitor of Example 3.3 if the capacitor voltage is described by $v_C(t) = 5t + 3$ V for $0 \leq t \leq 5$ s.

Answer: 5 mA

Page 176

CHECK YOUR UNDERSTANDING

The voltage waveform shown below appears across a 1,000-μF capacitor. Plot the capacitor current $i_C(t)$.



The Ideal Inductor

An inductor is an element that can store energy in a magnetic field within and around a conducting coil. In general, inductance is present whenever a conducting wire forms a loop. A simple example is a solenoid, which is a narrow and tightly wound coil of length ℓ, cross-sectional area A, and N turns. Inductors are typically made by winding wire around a **core**, which can be an insulator or a ferromagnetic material, as shown in Figure 3.12. A current through the coil establishes a magnetic field through and around the core. In an ideal inductor, the resistance of the wire is zero.

The inductance L is defined by the following ratio:

$$L \equiv \frac{N\Phi}{i_L}$$

where Φ is the magnetic flux through the inductor core and i_L is the current through the inductor coil. The inductance of an ideal solenoid is:

$$L = \frac{\mu N^2 A}{\ell}$$

where μ is the permeability of the core. Another inductor found in many applications is the toroid, which is also depicted in Figure 3.12.

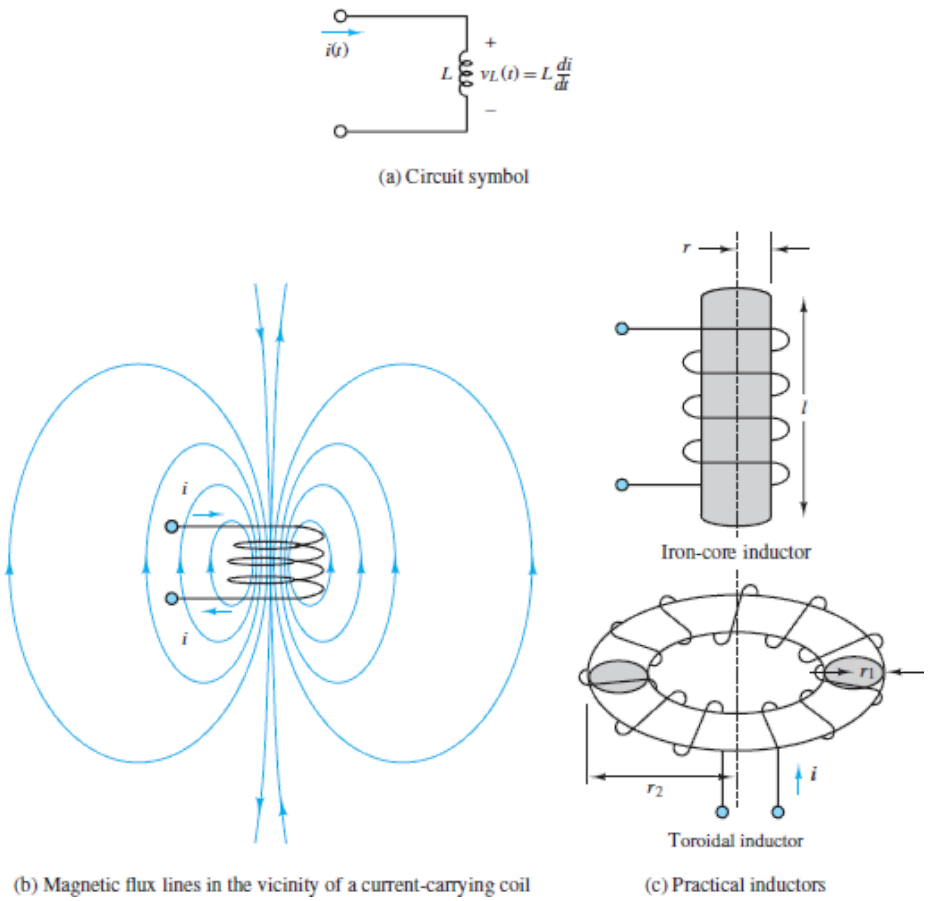


Figure 3.12 Inductance and practical inductors

The inductance of a coil is measured in henrys (H) where

$$1\,\text{H} = 1\,\text{V-s/A} \tag{3.16}$$

Page 177

Henrys are reasonable units for practical inductors although millihenrys (mH) are very common and microhenrys (μH) are occasionally found.

The *i*-*v* relationship for an inductor is derived directly from Faraday's law of induction but with the total flux $N\Phi$ replaced by Li from the definition of inductance L. The result is:

$$v_L(t) = L\frac{di_L(t)}{dt} \qquad i\text{-}v \text{ relation for inductor} \tag{3.17}$$

One immediate implication of equation 3.17 is that the voltage across an inductor in a DC circuit is zero. Why? Since the current through an inductor in a DC circuit must, by definition, be constant, the time derivative of the current must be zero. Thus, equation 3.17 requires the voltage across an inductor to also be zero.

An inductor in a DC circuit is equivalent to a short-circuit.

Page 178

Equation 3.17 can be integrated to yield an expression for the change in current through an inductor due to the voltage across it.

$$\Delta i_L = \frac{1}{L}\int^{t} v_L(\tau)\,d\tau \tag{3.18}$$

Equation 3.18 indicates that the change in current through an inductor depends on the history of the voltage across it. To calculate the current at a specific time it is necessary to know the current I_0 through the inductor at some previous time t_0.

$$i_L(t) = I_0 + \frac{1}{L}\int_{t_0}^{t} v_L(\tau)\,d\tau \qquad t \geq t_0 \tag{3.19}$$

Equivalent Inductance

Just as resistors in series or parallel can be represented by an equivalent resistance, so inductors in series or parallel can be represented by an equivalent inductance. For two inductors in series and parallel, the equivalent inductances are, respectively,

$$L_{eq} = L_1 + L_2 \qquad \text{and} \qquad L_{eq} = \frac{L_1 L_2}{L_1 + L_2} \tag{3.20}$$

Notice that the equivalent inductance of two inductors in series is simply the sum of the two, which is the same rule used for two resistors in series. Likewise, the rule for the equivalent inductance of two inductors in parallel is the product divided by the sum, which is the same rule used for two resistors in parallel. The more general rules are illustrated in Figure 3.13.

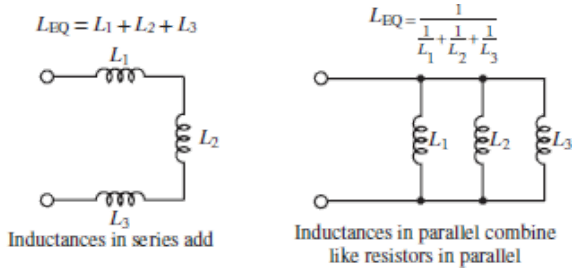


Figure 3.13 Equivalent inductance in a circuit

When calculating equivalent inductance, inductors in series combine like resistors in series and inductors in parallel combine like resistors in parallel.

Duality

All the relationships for capacitors and inductors exhibit duality, which means that the capacitor relations are mirror images of the inductor relations. Specifically, the Page 179roles played by voltage and current in a capacitor relation are reversed in the analogous inductor relation. For example, the *i*-*v* relationships for capacitors and inductors, respectively, are:

$$i = C\frac{dv}{dt} \quad \text{and} \quad v = L\frac{di}{dt}$$

Notice that the inductor relation is obtained from the capacitor relation by replacing i with v and v with i. It is also necessary, of course, to replace the capacitance C with the inductance L. Another example of duality is found in the energy storage relations for capacitors and inductors.

Duality is also at work in other relations not involving voltage and current explicitly. For example, consider the rules for calculating equivalent capacitance and equivalent inductance. Capacitors in series combine like inductors in parallel while capacitors in parallel combine like inductors in series. Another example of duality is seen in the DC behavior of capacitors and inductors. In a DC circuit, a capacitor acts like an open-circuit while an inductor acts like a short-circuit.

MAKE THE CONNECTION

Hydraulic Analog of an Inductor

Fluid inertance, which is caused by the inertial properties (i.e., the mass) of a fluid in motion, is analogous to inductance in an electric circuit. It is well known that a particle in motion possesses kinetic energy; likewise, a fluid in motion, which consists of a collection of particles, also possesses (i.e., stores) kinetic energy. Think of the water flowing out of a fire hose! The equations that define the analogy are:

$$\Delta p = p_1 - p_2 = I_f \frac{dq_f}{dt}$$

$$\Delta v = v_1 - v_2 = L \frac{di}{dt}$$

The figure below depicts the analogy. These analogies and the energy equations that apply to electric and fluid circuit elements are summarized in Table 3.3.

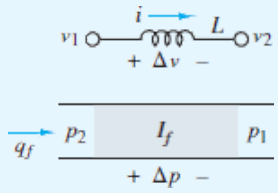


Table 3.3 Analogy between electric and fluid circuits

Property	Electric element or equation	Hydraulic analogy
Potential variable	Voltage or potential difference	Pressure difference
Flow variable	Current	Fluid volume flow rate
Resistance	Resistor R	Fluid resistor R_f
Capacitance	Capacitor C	Fluid capacitor C_f
Inductance	Inductor L	Fluid inertor I_f
Power dissipation	$P = i^2R$	$P_f = q_f^2 R_f$
Potential energy storage	$W_p = \frac{1}{2}Cv^2$	$W_p = \frac{1}{2}C_f p^2$
Kinetic energy storage	$W_k = \frac{1}{2}Li^2$	$W_k = \frac{1}{2}I_f q_f^2$

Analogy between fluid inertance and electrical inductance

Energy Storage in Inductors

The energy stored in an inductor $W_L(t)$ is derived from the definition of energy as the time integral of power, which is the product of voltage and current:

$$P_L(t) = i_L(t)v_L(t) = i_L(t)L\frac{di_L(t)}{dt} = \frac{d}{dt}\left[\frac{1}{2}Li_L^2(t)\right] \tag{3.21}$$

The total energy stored in the inductor is the integral of power over time.

$$W_L(t) = \int P_L(\tau)d\tau = \int \frac{d}{d\tau}\left[\frac{1}{2}Li_L^2(\tau)\right]d\tau$$

$$W_L = \frac{1}{2}Li_L^2 \quad \text{Energy stored in an inductor} \tag{3.22}$$

Note, once again, the duality with the expression for the energy stored in a capacitor, in equation 3.15.

Page 180

EXAMPLE 3.4 Calculating Inductor Voltage From Current

Problem

Calculate the voltage across an inductor from knowledge of its current.

Solution

Known Quantities: Inductor current; inductance value.

Find: Inductor voltage.

Schematics, Diagrams, Circuits, and Given Data:

$$i_L(t) = \begin{cases} 0\text{ mA} & t \le 1\text{ ms} \\ -\frac{0.1}{4} + \frac{0.1}{4}t \quad \text{mA} & 1 \le t \le 5\text{ ms} \\ 0.1\text{ mA} & 5 \le t \le 9\text{ ms} \\ 13 \times \frac{0.1}{4} - \frac{0.1}{4}t \quad \text{mA} & 9 \le t \le 13\text{ ms} \\ 0\text{ m} & t \ge 13\text{ ms} \end{cases}$$

$$L = 10\,\text{H}$$

where time t is in milliseconds. The inductor current is plotted in Figure 3.14.

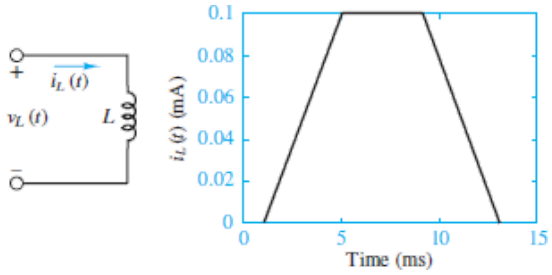


Figure 3.14

Assumptions: $i_L(t = 0) \le 0$.

Analysis: The voltage across the inductor is obtained by differentiating the current and multiplying by the inductance L.

$$v_L(t) = L\frac{di_L(t)}{dt}$$

Piecewise differentiating the expression for the inductor current, we obtain

$$v_L(t) = \begin{cases} 0\,\text{V} & t < 1\text{ ms} \\ 0.25\,\text{V} & 1 < t < 5\text{ ms} \\ 0\,\text{V} & 5 < t < 9\text{ ms} \\ -0.25\,\text{V} & 9 < t < 13\text{ ms} \\ 0\,\text{V} & t > 13\text{ ms} \end{cases}$$

The inductor voltage is plotted in Figure 3.15.

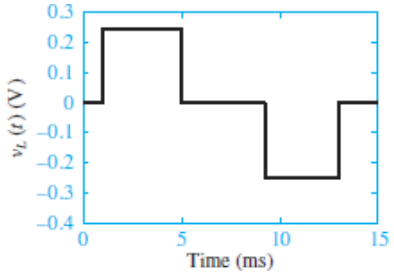


Figure 3.15

Comments: The inductor voltage can change instantaneously and thus be discontinuous!

Page 181

EXAMPLE 3.5 Calculating Inductor Current From Voltage

Problem

Use a time plot of the voltage across an inductor and its initial current to calculate the current through it as a function of time.

Solution

Known Quantities: Inductor voltage; initial condition (current at $t = 0$); inductance value.

Find: Inductor current.

Schematics, Diagrams, Circuits, and Given Data:

$$v(t) = \begin{cases} 0\,\text{V} & t < 0\,\text{s} \\ -10\,\text{mV} & 0 < t < 1\,\text{s} \\ 0\,\text{V} & t > 1\,\text{s} \end{cases}$$
$$L = 10\,\text{mH}; \qquad i_L(t = 0) = I_0 = 0\,\text{A}$$

The voltage across the inductor is plotted in Figure 3.16(a).

Analysis: Use the integral *i-v* relationship for an inductor to obtain the current through it:

$$i_L(t) = i_L(t_0) + \frac{1}{L}\int_{t_0}^{t} v(\tau)\,d\tau \qquad t \geq t_0$$
$$= \begin{cases} I_0 + \frac{1}{L}\int_0^t (-10 \times 10^{-3})\,d\tau = 0 + \frac{-10^{-2}}{10^{-2}}t = -t\,\text{A} & 0 \leq t \leq 1\,\text{s} \\ -1\,\text{A} & t \geq 1\,\text{s} \end{cases}$$

The inductor current is plotted in Figure 3.16b.

Comments: The inductor voltage can change instantaneously and thus be discontinuous!

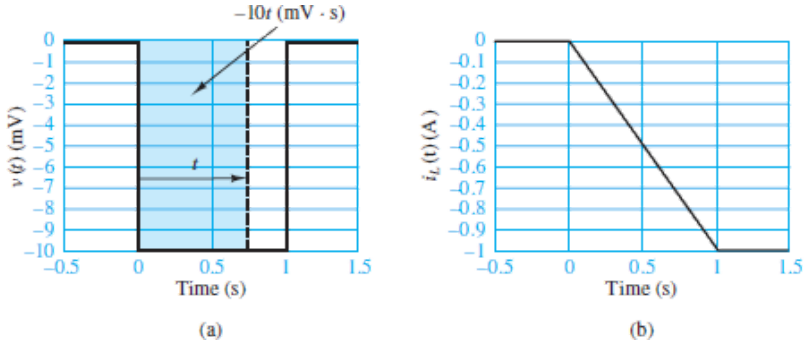


Figure 3.16

EXAMPLE 3.6 Energy Storage in an Ignition Coil

Problem

Determine the energy stored in an automotive ignition coil.

Page 182

Solution

Known Quantities: Inductor current; inductance value.

Find: Energy stored in inductor.

Schematics, Diagrams, Circuits, and Given Data: $L = 10$ mH; $i_L = 8$ A.

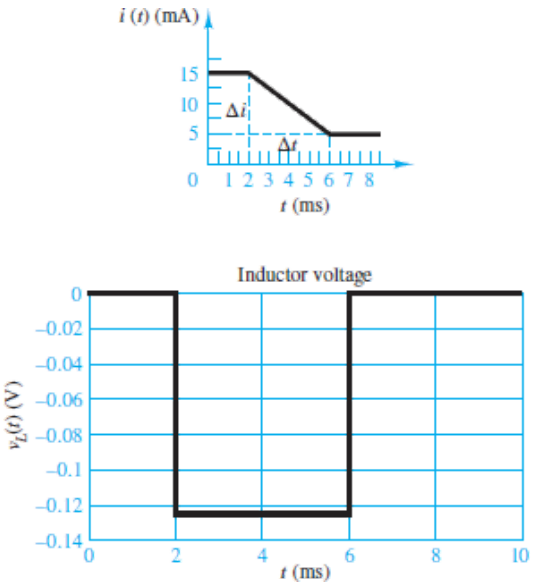
Analysis:

$$W_L = \frac{1}{2} L i_L^2 = \frac{1}{2} \times 10^{-2} \times 64 = 32 \times 10^{-2} = 320 \text{ mJ}$$

Comments: A more detailed analysis of an automotive ignition coil is presented in Chapter 4 to accompany the discussion of transient voltages and currents.

CHECK YOUR UNDERSTANDING

The waveform below shows the current through a 50-mH inductor. Plot the inductor voltage $v_L(t)$.



CHECK YOUR UNDERSTANDING

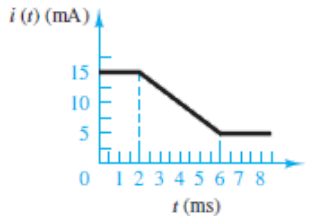
Find the maximum voltage across a 10-mH inductor when the inductor current is $i_L(t) = -2t(t-2)$ A for $0 \le t \le 2$ s and zero otherwise.

Answer: 40 mV

Page 183

CHECK YOUR UNDERSTANDING

Calculate and plot the inductor energy and power for a 50-mH inductor subject to the current waveform shown below. What is the energy stored at $t = 3$ ms?



Answer:

$$w(t = 3\text{ ms}) = 3.9\ \mu\text{J}$$

$$p(t) = \begin{cases} (20 \times 10^{-3} - 2.5t)(-0.125\ \text{W}) & 2 \le t < 6\text{ ms} \\ 0 & \text{otherwise} \end{cases}$$

$$w(t) = \begin{cases} 5.625 \times 10^{-6}\ \text{J} & 0 \le t < 2\text{ ms} \\ 0.156t^2 - (2.5 \times 10^{-3})t + 10^{-5} & 2 \le t < 6\text{ ms} \\ 0.625 \times 10^{-6} & t \ge 6\text{ ms} \end{cases}$$

3.3 TIME-DEPENDENT WAVEFORMS

Time-dependent periodic waveforms appear frequently in practical applications and are a useful approximation of many physical phenomena. For example, electric power worldwide is generated and delivered to industrial and household users in the form of periodic (i.e., 50- or 60-Hz sinusoidal) voltages and currents. In general, a periodic waveform $x(t)$ satisfies the equation

$$x(t) = x(t + nT) \qquad n = 1, 2, 3, \ldots \tag{3.23}$$

where T is the period of $x(t)$. Figure 3.17 illustrates a number of periodic waveforms that are typically encountered in the study of electric circuits. Waveforms such as the sine, triangle, square, pulse, and sawtooth waves are provided in the form of voltages (or, less frequently, currents) by commercially available signal generators.

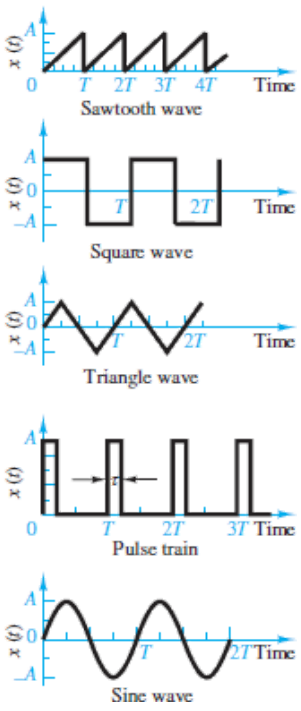


Figure 3.17 Periodic waveforms

In this chapter, time-varying voltages and currents and, in particular, sinusoidal (AC) waveforms are introduced. Figure 3.18 illustrates the convention employed to denote time-dependent sources.

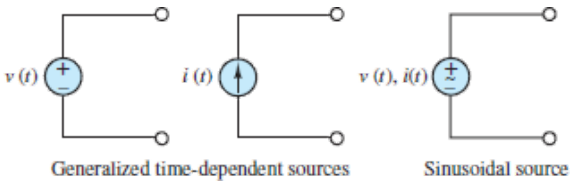


Figure 3.18 Time-dependent sources

Page 184

Sinusoids constitute the most important class of time-dependent waveforms. A generalized sinusoid is defined as

$$x(t) = A\cos(\omega t + \phi) \tag{3.24}$$

where A is the peak amplitude, ω the angular frequency, and ϕ the phase angle. Figure 3.19 summarizes the definitions of A, ω, and ϕ for the waveforms

$$x_1(t) = A\cos(\omega t) \qquad \text{and} \qquad x_2(t) = A\cos(\omega t + \phi)$$

where

$$f = \text{cyclical frequency} = \frac{1}{T} \quad \text{cycles/s, or Hz} \tag{3.25}$$

$$\omega = \text{angular frequency} = 2\pi f \quad \text{rad/s}$$

$$\phi = 2\pi \frac{\Delta t}{T} \quad \text{rad}$$

$$= 360 \frac{\Delta t}{T} \quad \text{deg}$$

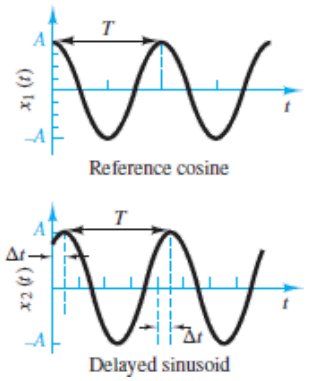


Figure 3.19 Sinusoidal waveforms have three characteristics: angular frequency ω, amplitude A, and phase angle ϕ. In general, $\omega T = 2\pi$ and $\phi = 2\pi\, \Delta t/t$, where T is the period and Δt is the time delay (positive or negative) relative to a reference waveform.

The value of the phase shift ϕ is a measure of the time delay of one sinusoid relative to a reference sinusoid, typically a cosine waveform. For example, a sine wave can be represented in terms of a cosine wave by introducing a phase shift of $\pi/2$ radians:

$$A \sin(\omega t) = A \cos\left(\omega t - \frac{\pi}{2}\right) \tag{3.26}$$

Notice that a negative phase angle represents a time shift to the right.

Although angular frequency ω, in units of radians per second, is commonly used to denote sinusoidal frequency, it is also common to employ the cyclical frequency f in units of cycles per second, or hertz (Hz). In music theory, a sinusoid is a pure tone; an A-440, for example, is a tone at a frequency of 440 Hz. The cyclical frequency is related to the angular frequency by the factor 2π.

$$\boxed{\omega = 2\pi f \quad \text{Angular frequency}} \tag{3.27}$$

Average (Mean) Value

Various measures exist for quantifying the amplitude of a time-varying electric signal. One of these measures is the average or mean value (also called the DC value). The average value of a waveform is computed by integrating it over a suitably chosen period, as shown in equation 3.28.

$$\langle x(t) \rangle = \frac{1}{T}\int_0^T x(\tau)\,d\tau \qquad \text{Average or mean value} \tag{3.28}$$

where T is the period of integration. Figure 3.20 illustrates the average amplitude of $x(t)$ over a period of T seconds. It can be shown that the average or mean value of a sinusoid is zero.

$$\langle A\cos(\omega t + \phi) \rangle = 0 \tag{3.29}$$

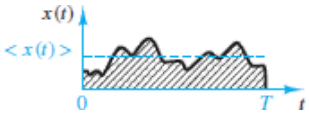


Figure 3.20 Average of a waveform

Page 185

This result might be surprising at first: If the average voltage across or current through an element is zero, is its average power also equal to zero? Clearly, the answer must be no. Otherwise, it would be impossible to illuminate households and streets and power industrial machinery with a 60-Hz sinusoidal voltage waveform!

Effective or RMS Value

A more useful measure of the amplitude of an AC waveform $x(t)$ is the *effective* or root-mean-square (rms) value, which takes into account fluctuations of a waveform about its mean, and which is defined as:

$$x_{\text{eff}} = x_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T x^2(\tau)d\tau} \qquad \text{Effective or rms value} \tag{3.30}$$

MAKE THE CONNECTION

Why Do We Use Units of Radians for the Phase Angle ϕ?

Engineers often find it more intuitive to express phase angle in units of degrees; however, to use consistent units in the argument (the quantity in the parentheses) of the expression $x(t) = A\sin(\omega t + \phi)$, we must express ϕ in units of radians, since the units of ωt are $[\omega] \cdot [t] = (\text{rad/s}) \cdot \text{s} = \text{rad}$. Thus, we will consistently use units of radians for the phase angle ϕ in all expressions of the form $x(t) = A\sin(\omega t + \phi)$. To be consistent is especially important when one is performing numerical calculations; if one used units of degrees for ϕ in calculating the value of $x(t) = A\sin(\omega t + \phi)$ at a given t, the answer would be incorrect.

Notice that the argument of the square root is the mean value of $x^2(t)$. Thus, the rms value is literally the square root of the mean of the square. Also note that the unit of the "mean of the square" is the unit of $x^2(t)$. Thus, the unit of the "root of the mean of the square" x_{rms} is the unit of $x(t)$.

Why are effective (rms) values useful? Consider two similar circuits, each with a resistor R connected to a source: one with a DC source, and one with an AC source, as shown in Figure 3.21. The effective value of the AC source is the value of a DC source such that the average power dissipated by the resistor R is the same in both circuits. Thus, the effective value of an AC source provides a comparable measure of the power associated with an element in an AC circuit. The AC power dissipated by a resistor can now be simply expressed as:

$$P_{\text{avg}} = I_{\text{eff}}^2 R = \frac{V_{\text{eff}}^2}{R} \tag{3.31}$$

In terms of AC current and voltage waveforms, the effective (rms) values are calculated as:

$$I_{\text{eff}} = I_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T i_{\text{ac}}^2(\tau)\,d\tau} \qquad \text{and} \qquad V_{\text{eff}} = V_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T v_{\text{ac}}^2(\tau)\,d\tau} \tag{3.32}$$

LO

The rms, or effective, value of an AC source is the DC value that produces the same average power dissipated by a resistive load.

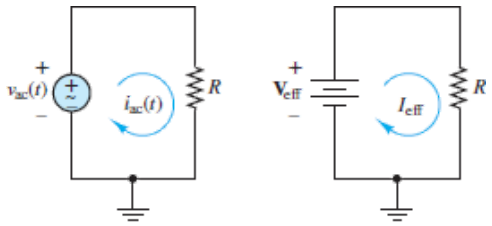


Figure 3.21 AC and DC circuits used to illustrate the concept of effective and rms values

The effective (or rms) value of a voltage or current is indicated by the notation V_{rms}, or $\tilde{V}$, and I_{rms}, or $\tilde{I}$. The ratio of the rms value of a sinusoid to its peak value is $1/\sqrt{2} \approx 0.707$. Table 3.4 lists the value of this ratio for other typical waveforms. The table also lists a Fourier sine series for each waveform to demonstrate that each is a summation of sine waves.

Page 186

Table 3.4 Ratio of RMS value to Peak value

Waveform	$x(t)$	$x_{\text{rms}}/x_{\text{pk}}$
Sinusoid	$A\sin(\omega t)$	$\frac{\sqrt{2}}{2} \approx 0.707$
Square	$\frac{8A}{\pi}\sum_{k=1}^{\infty}\frac{\sin[(2k-1)\omega t]}{2k-1}$	1
Triangle	$\frac{8A}{\pi^2}\sum_{k=1}^{\infty}(-1)^k\frac{\sin[(2k-1)\omega t]}{(2k-1)^2}$	$\frac{\sqrt{3}}{3} \approx 0.577$
Sawtooth	$\frac{2A}{\pi}\sum_{k=1}^{\infty}\frac{\sin(k\omega t)}{k}$	$\frac{\sqrt{3}}{3} \approx 0.577$

EXAMPLE 3.7 Average Value of Sinusoidal Waveform

Problem

Compute the average value of the signal $x(t) = 10\cos(100t)$.

Solution

Known Quantities: Functional form of the periodic signal $x(t)$.

Find: Average value of $x(t)$.

Analysis: The signal is periodic with period $T = 2\pi/\omega = 2\pi/100$. Integrate over one period to compute the average value:

$$\langle x(t)\rangle = \frac{1}{T}\int_0^T x(\tau)\,d\tau = \frac{100}{2\pi}\int_0^{2\pi/100} 10\cos(100t)\,dt$$

$$= \frac{10}{2\pi}\langle \sin(2\pi) - \sin(0)\rangle = 0$$

Comments: The mean value of a sinusoidal is zero, independent of its amplitude and frequency.

EXAMPLE 3.8 RMS Value of Sinusoidal Waveform

Problem

Compute the rms value of the sinusoidal current $i(t) = I\cos(\omega t)$.

Page 187

Solution

Known Quantities: Peak amplitude I and angular frequency ω of the periodic signal $i(t)$.

Find: rms value of $i(t)$.

Analysis: Applying the definition of rms value in equation 3.32, we compute

$$I_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T i^2(\tau)d\tau} = \sqrt{\frac{\omega}{2\pi}\int_0^{2\pi/\omega} I^2\cos^2(\omega\tau)d\tau}$$

$$= \sqrt{\frac{\omega}{2\pi}\int_0^{2\pi/\omega} I^2\left[\frac{1}{2}+\frac{1}{2}\cos(2\omega\tau)\right]d\tau}$$

$$= \sqrt{\frac{1}{2}I^2+\frac{\omega}{2\pi}\int_0^{2\pi/\omega}\frac{I^2}{2}\cos(2\omega\tau)d\tau}$$

Since the integral of a sinusoid over one period is equal to zero (see Example 3.7), the integral over two periods is also zero. (Notice that the argument of cosine function is $2\omega\tau$, which indicates that $T = 2\pi/(2\omega) = \pi/\omega$ such that $2\pi/\omega = 2T$.) Thus:

$$i_{\text{rms}} = \frac{I}{\sqrt{2}} = 0.707I$$

Comments: The rms value of a sinusoidal signal is independent of its frequency.

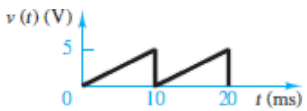
CHECK YOUR UNDERSTANDING

Express the voltage $v(t) = 155.6\sin(377t + \pi/6)$ in cosine form. Note that the angular frequency $\omega = 377$ rad/s is equivalent to the cyclical frequency 60 Hz, which is the frequency of the electric power generated in North America.

Answer: $v(t) = 155.6\cos(377t - \pi/3)$

CHECK YOUR UNDERSTANDING

Compute the mean (average) and rms values of the sawtooth waveform shown below.



Answer: $v_{\text{avg}} = 2.5$ V; $v_{\text{rms}} = 2.89$ V

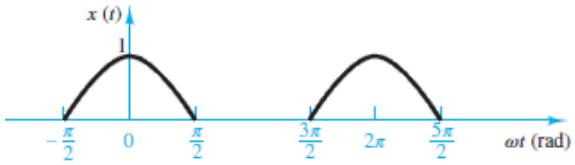
CHECK YOUR UNDERSTANDING

Compute the mean (average) and rms values of the triangle waveform shown below.

Answer: $v_{avg} = 1.5$ V; $v_{rms} = \sqrt{3}$ V

CHECK YOUR UNDERSTANDING

Compute the mean (average) and rms values of the clipped cosine waveform shown below.



$x(t) = \cos t \quad \text{for } \frac{-\pi}{2} \le \omega t < \frac{\pi}{2}$

$= 0 \quad \text{for } \frac{\pi}{2} \le \omega t < \frac{3\pi}{2} \qquad \omega = 1$

Answer: $x_{avg} = 1/\pi$; $x_{rms} = 0.5$

3.4 PHASOR SOLUTION OF CIRCUITS WITH SINUSOIDAL SOURCES

Any sinusoidal signal may be represented as a real function in the time domain:

$$v(t) = A\cos(\omega t + \phi)$$

or as a complex function in the frequency domain:

$$V(j\omega) = Ae^{j(\omega t+\phi)} = Ae^{j\phi}e^{j\omega t}$$

where A is the peak amplitude and ϕ is the phase shift relative to a reference sinusoid. It is important to keep in mind that phase shift in the frequency domain is equivalent to a delay in the time domain.

Since the sinusoidal frequencyω of each independent voltage or current source is common to all variables in an AC circuit, the complex exponential $e^{j\omega t}$ is usually not expressed explicitly in AC circuit analysis. However, the frequencyω of each sinusoidal source is an important parameter for characterizing the impact of capacitors and inductors in an AC circuit.

Page 189

Euler's Formula

Named after the famous Swiss mathematician Leonhard Euler, this formula is the basis of phasor notation. A phasor is similar to a vector in that it has an amplitude A and direction θ in the complex plane. Also, just as a vector can be decomposed into x and y components, a phasor can be decomposed into real and imaginary components. Euler's formula defines a complex exponential $e^{j\theta}$ as a *unit phasor* in the complex plane, with real and imaginary components given by:

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{3.33}$$

where $j \equiv \sqrt{-1}$ is the imaginary unit. The symbol θ is simply a place holder in Euler's formula. Any dimensionless quantity or expression can be substituted for θ in the formula. However, in AC circuit analysis, θ takes on the physical meaning of the phase shift of a sinusoid.

Leonhard Euler (1707–1783) (*Oxford Science Archive/ Heritage Images/The Print Collector/Alamy Stock Photo*)

The **dark** black arrow in Figure 3.22 represents a complex exponential in the complex plane. The real and imaginary components are cos θ and sin θ, respectively. These two components and the complex exponential itself form the three legs of a right triangle. The Pythagorean theorem requires:

$$|e^{j\theta}|^2 = \cos^2\theta + \sin^2\theta = 1 \tag{3.44}$$

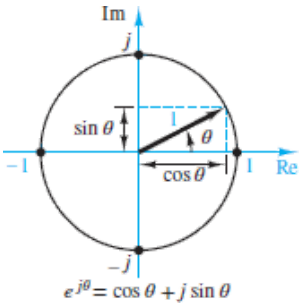


Figure 3.22 Euler's formula

Thus, the magnitude of $e^{j\theta}$ is unity, which is why it is also known as a *unit phasor*. The angle of inclination of the unit phasor is θ. As θ increases or decreases the unit phasor rotates counterclockwise or clockwise, respectively, about the origin of the complex plane.

It is difficult to overstate the power of the visualization presented in Figure 3.22. For example, when $\theta = \pi/2$, the unit phasor points straight up along the imaginary axis. Thus:

$$e^{j\pi/2} = 1\angle\frac{\pi}{2} = j \tag{3.55}$$

where the notation $1\angle\frac{\pi}{2}$ indicates a magnitude of 1 and a phase angle $\theta = \pi/2$. When $\theta = \pi$, the unit phasor points to the left along the negative real axis. Thus:

$$e^{j\pi} = 1\angle\pi = -1 \tag{3.36}$$

Likewise:

$$e^{j3\pi/2} = 1\angle\frac{3\pi}{2} = -j \quad \text{and} \quad e^{j2\pi} = 1\angle 2\pi = 1 \tag{3.37}$$

Each of these expressions equates the two polar forms on the left to the rectangular form on the far right side. In polar form, a phasor is represented by a magnitude (or

amplitude) and a phase angle, whether as $Ae^{j\theta}$ or $A\angle\theta$. In rectangular form, a phasor is represented by real and imaginary components. Table 3.5 lists a few other commonly encountered phasors in polar and rectangular forms.

Table 3.5 Polar and rectangular forms of common phasors

Complex exponential	Polar	Rectangular
$Ae^{\pm j\pi/6}$	$A\angle \pm \pi/6$	$A(\sqrt{3}/2 \pm j/2)$
$Ae^{\pm j\pi/4}$	$A\angle \pm \pi/4$	$A(\sqrt{2}/2 \pm j\sqrt{2}/2)$
$Ae^{\pm j\pi/3}$	$A\angle \pm \pi/3$	$A(1/2 \pm j\sqrt{3}/2)$
$Ae^{\pm j \arctan(3/4)}$	$A\angle \pm \arctan(3/4)$	$A(0.8 \pm j0.6)$
$Ae^{\pm j \arctan(4/3)}$	$A\angle \pm \arctan(4/3)$	$A(0.6 \pm j0.8)$

In general, the polar and rectangular forms are related by:

$$Ae^{j\theta} = A\angle\theta = A\cos\theta + jA\sin\theta \tag{3.38}$$

In effect, Euler's identity is simply a trigonometric relationship in the complex plane.

Page 190

Phasors

To see how complex numbers can be used to represent sinusoidal waveforms, rewrite the expression for a generalized sinusoid in light of Euler's equation:

$$A\cos(\omega t + \theta) = \mathrm{Re}\,(Ae^{j(\omega t+\theta)}) \tag{3.39}$$

Notice that it is possible to express any sinusoid as the real part of a complex exponential with an argument of$\omega t + \theta$ and a magnitude or amplitude of A. The expression can be further simplified by remembering that the angular frequencyω is common to all voltages and currents. Thus, the $e^{j\omega t}$ portion of the complex exponential is understood to be present in every phasor, but not written explicitly. The same perspective is taken with regard to the real part operator Re so that the complex exponential is simplified as shown in equation 3.40.

$$\mathrm{Re}\,\{Ae^{j(\omega t+\theta)}\} = \mathrm{Re}\,\{Ae^{j\omega t}e^{j\theta}\} \Rightarrow Ae^{j\theta} \tag{3.40}$$

In this expression, the relational operator $\Rightarrow$ indicates equality with the real part operator Re and the sinusoidal portion $e^{j\omega t}$ of the complex exponential hidden but understood implicitly. In general, this simplification will be used to express a phasor in polar and rectangular form as:

$$Ae^{j\theta} = A\angle\theta = A(\cos\theta + j\sin\theta) \qquad \text{Phasor notation} \tag{3.41}$$

There are five key rules of complex arithmetic that will be used to resolve complex multiplication and division:

1. The magnitude of the ratio of two phasors is the ratio of the individual magnitudes.

$$\left|\frac{\mathbf{V}}{\mathbf{I}}\right| = \frac{|\mathbf{V}|}{|\mathbf{I}|}$$

2. The phase angle of the ratio of two phasors is the difference of the individual phase angles.

$$\angle\left(\frac{\mathbf{V}}{\mathbf{I}}\right) = \angle\mathbf{V} - \angle\mathbf{I}$$

3. The complex conjugate $\bar{\mathbf{A}}$ of a phasor **A** is found by changing the sign of the imaginary number, *j*, everywhere in the phasor. The magnitude of the complex conjugate equals the magnitude of the phasor itself. The angle of the complex conjugate equals the negative of the angle of the phasor itself.
4. The product of a phasor and its complex conjugate is a real number equal to the square of the magnitude of the phasor, which is equal to the sum of the squares of the real and imaginary parts of the phasor.

Page 191

5. The angle of a phasor is the inverse tangent of the ratio of the imaginary part to the real part. That is, $\angle\mathbf{A} = \arctan[\text{Im}(\mathbf{A})/\text{Re}(\mathbf{A})]$.

A bold uppercase font indicates a phasor quantity.

Superposition of AC Signals

Consider the circuit depicted in Figure 3.23 with a load excited by two current sources in parallel.

$$\begin{aligned} i_1(t) &= A_1\cos(\omega_1 t + \theta_1) \\ i_2(t) &= A_2\cos(\omega_2 t + \theta_2) \end{aligned} \tag{3.42}$$

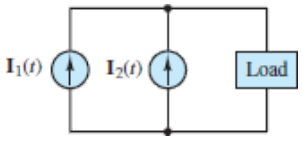


Figure 3.23 Superposition

By KCL, the load current is equal to the sum of the two source currents; that is,

$$i_{\text{load}}(t) = i_1(t) + i_2(t) \tag{3.43}$$

So far, so good. However, the expression in equation 3.43 *cannot* be expressed in phasor form without masking the fact that i_1 has a different frequency than that of i_2. In mathematical form:

$$\begin{aligned} \mathbf{I}_{\text{load}} &\neq \mathbf{I}_1 + \mathbf{I}_2 \\ &\neq A_1 e^{j\theta_1} + A_2 e^{j\theta_2} \end{aligned} \tag{3.44}$$

It is imperative to remember that the $e^{j\omega_1 t}$ and $e^{j\omega_2 t}$ terms are present implicitly in $\mathbf{I}_1$ and $\mathbf{I}_2$, respectively, as shown in equation 3.45.

$$\begin{aligned} i_1(t) &= \text{Re}\left(\mathbf{I}_1 e^{j\omega_1 t}\right) \\ i_2(t) &= \text{Re}\left(\mathbf{I}_2 e^{j\omega_2 t}\right) \end{aligned} \tag{3.45}$$

The two phasors of equation 3.44 cannot be added, but must be kept separate; the only unambiguous expression for the load current is equation 3.43. In general, sinusoidal waveforms of different frequencies must be analyzed separately.

3.5 IMPEDANCE

As phasors, the *i-v* relationships of resistors, capacitors and inductors take the form of a generalized Ohm's law:

$$\mathbf{V} = \mathbf{IZ}$$

where the quantity **Z** is known as impedance.

Series and parallel combinations of resistors, inductors, and capacitance can be represented by a single equivalent impedance of the form:

$$\mathbf{Z}(j\omega) = R(j\omega) + jX(j\omega) \qquad \text{units of } \Omega \text{ (ohms)}$$

where $R(j\omega)$ and $X(j\omega)$ are known as the "resistance" and "reactance", respectively, of the equivalent impedance **Z**.

In general, all the DC circuit relations and techniques introduced in Chapters 1 and 2 can be extended to AC circuits. Thus, it is not necessary to learn new methods to solve AC circuits; it is only necessary to apply the same methods with phasors.

Generalized Ohm's Law

The impedance concept reflects the fact that capacitors and inductors act as frequency-dependent resistors. Figure 3.24 depicts a generic AC circuit element with a sinusoidal voltage $\mathbf{V}_Z$ and an impedance load $\mathbf{Z}$, which is also a phasor and represents the effect of a generic network of resistors, capacitors, and inductors. The resulting current $\mathbf{I}_Z$ is a phasor determined by:

$$\mathbf{V}_Z = \mathbf{I}_Z\mathbf{Z} \quad \text{Generalized Ohm's law} \tag{3.46}$$

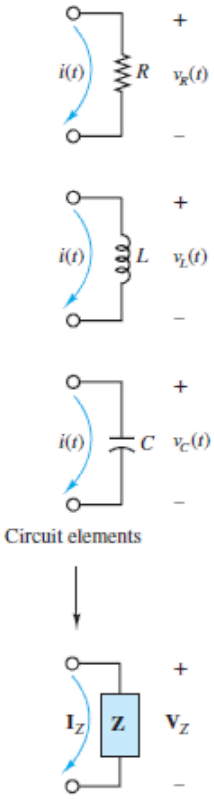


Figure 3.24 The impedance concept

The impedance $\mathbf{Z}$ for a specific network of resistors, capacitors, and inductors is determined by the definition of impedance:

$$\boxed{\mathbf{Z} \equiv \frac{\mathbf{V}}{\mathbf{I}} \quad \text{Definition of impedance}} \tag{3.47}$$

Once the impedances of the resistors, capacitors and inductors in a network are known, they can be combined in series and parallel (using the usual rules for resistors) to form equivalent impedances "seen" between nodes within the network.

Impedance of a Resistor

The *i-v* relationship for a resistor is, of course, Ohm's law, which in the case of sinusoidal sources is written as (see Figure 3.25):

$$v_R(t) = i_R(t)R \tag{3.48}$$

or, in phasor form,

$$\mathbf{V}_R e^{j\omega t} = \mathbf{I}_R e^{j\omega t} R$$

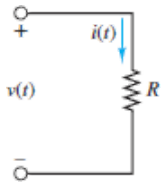


Figure 3.25 For a resistor, $v_R(t) = i_R(t)\,R$

where $\mathbf{V}_R = V_R e^{j\theta_v}$ and $\mathbf{I}_R = I_R e^{j\theta_i}$ are phasors.

Both sides of equation 3.48 can be divided by $e^{j\omega t}$ to yield:

$$\mathbf{V}_R = \mathbf{I}_R R \tag{3.49}$$

The impedance of a resistor is then determined from the definition of impedance:

$$\mathbf{Z}_R \equiv \frac{\mathbf{V}_R}{\mathbf{I}_R} = R \tag{3.50}$$

Thus:

$$\mathbf{Z}_R = R \quad \text{Impedance of a resistor} \tag{3.51}$$

The impedance of a resistor is a real number; that is, it has a magnitude R and a zero phase, as shown in Figure 3.26. The phase of the impedance is equal to the phase difference between the voltage across an element and the current through the same element. In the case of a resistor, the voltage is completely in phase with the current, which means that there is no time delay between the voltage waveform and the current waveform in the time domain.

Page 193

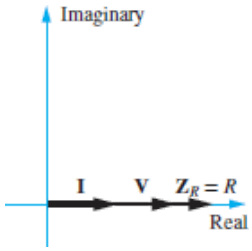


Figure 3.26 Phasor diagram of the impedance of a resistor. Remember that $\mathbf{Z} = \mathbf{V}/\mathbf{I}$.

Impedance of an Inductor

The *i-v* relationship for an inductor is (see Figure 3.27):

$$v_L(t) = L\frac{di_L(t)}{dt} \tag{3.52}$$

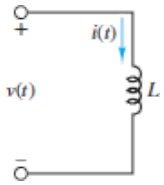


Figure 3.27 For an inductor, $v_L(t) = L\frac{di_L(t)}{dt}$

At this point, it is important to proceed carefully. The time-domain expression for the current through the inductor is

$$i_L(t) = I_L\cos(\omega t + \theta)$$

such that

$$\begin{aligned}\frac{d}{dt}i_L(t) &= -I_L\omega\sin(\omega t+\theta)\\ &= I_L\omega\cos(\omega t+\theta+\pi/2)\\ &= \mathrm{Re}\left(I_L\omega e^{j\pi/2}e^{j(\omega t+\theta)}\right)\\ &= \mathrm{Re}\left[I_L(j\omega)e^{j(\omega t+\theta)}\right]\end{aligned} \tag{3.53}$$

The net effect of the time derivative is to produce an extra ($j\omega$) term along with the complex exponential expression of $i_L(t)$.

Time domain	Frequency domain
$\frac{d}{dt}$	$j\omega$

Therefore, the phasor equivalent of the *i-v* relationship for an inductor is:

$$\mathbf{V}_L = L(j\omega)\mathbf{I}_L \tag{3.54}$$

The impedance of an inductor is then determined from the definition of impedance:

$$\mathbf{Z}_L \equiv \frac{\mathbf{V}_L}{\mathbf{I}_L} = j\omega L \tag{3.55}$$

Thus:

$$\mathbf{Z}_L = j\omega L = \omega L\angle\frac{\pi}{2} \qquad \text{Impedance of an inductor} \tag{3.56}$$

The impedance of an inductor is a positive, purely imaginary number; that is, it has a magnitude ofωL and a phase of $\pi/2$ radians or 90°, as shown in Figure 3.28. As before, the phase of the impedance is equal to the phase difference between the voltage across an element and the current through the same element. In the case of an inductor, the voltage *leads* the current by $\pi/2$ radians, which means that a feature (e.g., a zero crossing point) of the voltage waveform occurs *earlier* than the same feature of the current waveform.

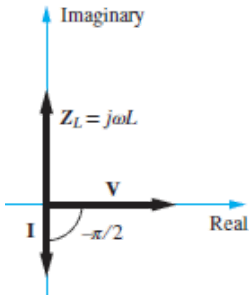


Figure 3.28 Phasor diagram of the impedance of an inductor. Remember that **Z** = **V**/**I**.

Note that the inductor behaves as a complex frequency-dependent resistor and that its magnitude ωL is proportional to the angular frequency ω. Thus, an inductor will "impede" current in proportion to the frequency of the source signal.

Page 194

> At low frequencies, an inductor acts like a short-circuit; at high frequencies, it acts like an open-circuit.

Impedance of a Capacitor

The principle of duality suggests that the procedure to derive the impedance of a capacitor should mirror the procedure for an inductor. The *i*-*v* relationship for a capacitor is (see Figure 3.29):

$$i_C(t) = C\frac{dv_C(t)}{dt} \tag{3.57}$$

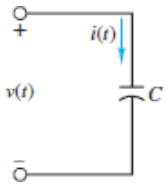


Figure 3.29 For a capacitor, $i_C(t) = C\frac{dv_C(t)}{dt}$

The time-domain expression for the voltage across the capacitor is

$$v_C(t) = V_C\cos(\omega t + \theta)$$

suchthat

$$\begin{aligned}\frac{d}{dt}v_C(t) &= -V_C\omega\sin(\omega t+\theta)\\ &= V_C\omega\cos(\omega t+\theta+\pi/2)\\ &= \mathrm{Re}\left(V_C\omega e^{j\pi/2}e^{j(\omega t+\theta)}\right)\\ &= \mathrm{Re}\left[V_C(j\omega)e^{j(\omega t+\theta)}\right]\end{aligned} \tag{3.58}$$

The *net* effect of the time derivative is to produce an extra ($j\omega$) term along with the complex exponential expression of $v_C(t)$. Therefore, the phasor equivalent of the *i*-*v* relationship for a capacitor is

$$\mathbf{I}_C = C(j\omega)\mathbf{V}_C \tag{3.59}$$

The impedance of an inductor is then determined from the definition of impedance:

$$\mathbf{Z}_C \equiv \frac{\mathbf{V}_C}{\mathbf{I}_C} = \frac{1}{j\omega C} = \frac{-j}{\omega C} \tag{3.60}$$

Thus:

$$\boxed{\mathbf{Z}_C = \frac{1}{j\omega C} = \frac{-j}{\omega C} = \frac{1}{\omega C}\angle\frac{-\pi}{2} \qquad \text{Impedance of a capacitor}} \tag{3.61}$$

The impedance of a capacitor is a negative, purely imaginary number; that is, it has a magnitude of $1/\omega C$ and a phase of $-\pi/2$ radians or $-90°$, as shown in Figure 3.30. As before, the phase of the impedance is equal to the phase difference between the voltage across an element and the current through the same element. In the case of a capacitor, the voltage *lags* the current by $\pi/2$ radians, which means that a feature (e.g., a zero crossing point) of the voltage waveform occurs *later* than the same feature of the current waveform.

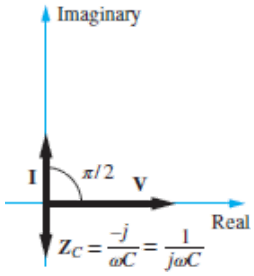


Figure 3.30 Phasor diagram of the impedance of a capacitor. Remember that **Z** = **V**/**I**.

Note that the capacitor also behaves as a complex frequency-dependent resistor, except that its magnitude $1/\omega C$ is inversely proportional to the angular frequencyω. Thus, a capacitor will "impede" current in inverse proportion to the frequency of the source.

Page 195

At low frequencies, a capacitor acts like an open-circuit; at high frequencies, it acts like a short-circuit.

Generalized Impedance

The impedance concept is very useful in solving AC circuit analysis problems. It allows network theorems developed for DC circuits to be applied to AC circuits. The only difference is that complex arithmetic, rather than scalar arithmetic, must be employed to find equivalent impedance and solve equations.

Figure 3.31 depicts $\mathbf{Z}_R(\,j\omega)$, $\mathbf{Z}_L(\,j\omega)$, and $\mathbf{Z}_C(\,j\omega)$ in the complex plane. It is important to emphasize that although the impedance of resistors is purely real and the impedance of capacitors and inductors is purely imaginary, the equivalent impedance between two terminals in an arbitrary circuit can be complex.

$$\mathbf{Z}(j\omega) = R + X(j\omega) \tag{3.62}$$

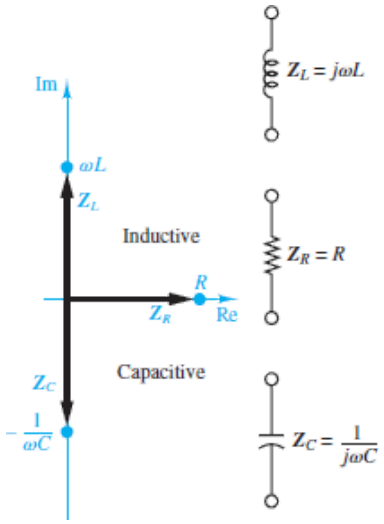


Figure 3.31 The impedances of R, L, and C are shown in the complex plane. Impedances in the upper right quadrant are inductive while those in the lower right quadrant are capacitive.

Here, R is *resistance* and X is *reactance*. The unit of R, X, and $\mathbf{Z}$ is the ohm.

Admittance

The solution of certain circuit analysis problems is handled more easily in terms of conductances than resistances. This is true, for example, in circuits with many parallel elements, since conductances in parallel add as resistors in series do. In AC circuit analysis, an analogous quantity may be defined—the reciprocal of complex impedance. Just as conductance G was defined as the inverse of resistance, admittance $\mathbf{Y}$ is defined as the inverse of impedance.

$$\mathbf{Y} \equiv \frac{1}{\mathbf{Z}} \qquad \text{unit of S (siemens)} \tag{3.63}$$

Whenever the impedance $\mathbf{Z}$ is purely real, the admittance $\mathbf{Y}$ is identical to the conductance G. In general, however, $\mathbf{Y}$ is complex.

$$\mathbf{Y} = G + jB \tag{3.64}$$

where G is the AC conductance and B is the susceptance, which is analogous to reactance. Clearly, G and B are related to R and X; however, the relationship is not a simple inverse. If $\mathbf{Z} = R + jX$, then the admittance is:

$$\mathbf{Y} = \frac{1}{\mathbf{Z}} = \frac{1}{R + jX} \tag{3.65}$$

Multiply the numerator and denominator by the complex conjugate $\mathbf{Z} = R - jX$:

$$\mathbf{Y} = \frac{\overline{\mathbf{Z}}}{\mathbf{Z}\overline{\mathbf{Z}}} = \frac{R - jX}{R^2 + X^2} \tag{3.66}$$

and conclude that

$$G = \frac{R}{R^2 + X^2} \tag{3.67}$$
$$B = \frac{-X}{R^2 + X^2}$$

Notice in particular that G is not the reciprocal of R in the general case!

EXAMPLE 3.9 Addition of Two Sinusoidal Sources Using Phasor Notation

Problem

Compute the phasor voltage across a series connection of two sinusoidal voltage sources (Figure 3.32).

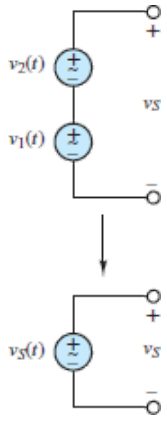


Figure 3.32

Solution

Known Quantities:

$$v_1(t) = 15\cos\left(377t + \frac{\pi}{4}\right) \quad \text{V}$$
$$v_2(t) = 15\cos\left(377t + \frac{\pi}{12}\right) \quad \text{V}$$

Find: Equivalent phasor voltage $v_S(t)$.

Analysis: Write the two voltages in phasor form:

$$\mathbf{V}_1(j\omega) = 15\angle\frac{\pi}{4} \quad \text{V}$$
$$\mathbf{V}_2(j\omega) = 15e^{j\pi/12} = 15\angle\frac{\pi}{12} \quad \text{V}$$

The phasor diagram of Figure 3.33 shows $\mathbf{V}_1$ and $\mathbf{V}_2$ in the complex plane. Convert the phasor voltages from polar to rectangular form:

$$\mathbf{V}_1(j\omega) = 10.61 + j10.61 \quad \text{V}$$
$$\mathbf{V}_2(j\omega) = 14.49 + j3.88 \quad \text{V}$$

Then, by KVL:

$$\mathbf{V}_S(j\omega) = \mathbf{V}_1(j\omega) + \mathbf{V}_2(j\omega) = 25.10 + j14.49 = 28.98\,e^{j\pi/6} = 28.98\angle\frac{\pi}{6} \quad \text{V}$$

Finally, convert $\mathbf{V}_S(j\omega)$ to its time-domain form:

$$v_S(t) = 28.98\cos\left(377t + \frac{\pi}{6}\right) \quad \text{V}$$

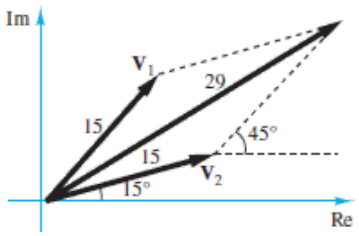


Figure 3.33 Phasor diagram showing the addition of two voltage phasors.

Comments: The same result could have been obtained by adding the two sinusoids in the time domain, using trigonometric identities:

$$v_1(t) = 15\cos\left(377t + \frac{\pi}{4}\right) = 15\cos\frac{\pi}{4}\cos(377t) - 15\sin\frac{\pi}{4}\sin(377t) \quad \text{V}$$

$$v_2(t) = 15\cos\left(377t + \frac{\pi}{12}\right) = 15\cos\frac{\pi}{12}\cos(377t) - 15\sin\frac{\pi}{12}\sin(377t) \quad \text{V}$$

Page 197

Combine like terms to obtain:

$$\begin{aligned} v_1(t) + v_2(t) &= 15\left(\cos\frac{\pi}{4} + \cos\frac{\pi}{12}\right)\cos(377t) - 15\left(\sin\frac{\pi}{4} + \sin\frac{\pi}{12}\right)\sin(377t) \\ &= 15[1.673\cos(377t) - 0.966\sin(377t)] \\ &= 15\sqrt{(1.673)^2 + (0.966)^2} \times \cos\left[377t + \arctan\left(\frac{0.966}{1.673}\right)\right] \\ &= 15\left[1.932\cos\left(377t + \frac{\pi}{6}\right)\right] = 28.98\cos\left(377t + \frac{\pi}{6}\right) \quad \text{V} \end{aligned}$$

The above expression is, of course, identical to the one obtained using phasor notation, but it required more computation. Phasor analysis often simplifies calculations.

EXAMPLE 3.10 Impedance of a Practical Capacitor

Problem

A practical capacitor is often modeled as an ideal capacitor in parallel with a resistor as shown in Figure 3.34. The parallel resistance represents leakage losses in the capacitor that can be quite significant. Find the impedance of a practical capacitor at the radian frequency ω = 377 rad/s (60 Hz). How will the impedance change if the capacitor is used at a much higher frequency, say, 800 kHz?

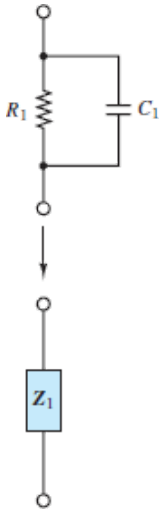


Figure 3.34

Solution

Known Quantities: Figure 3.34; C_1 = 1.0 nF; R_1 = 1 MΩ; ω = 377 rad/s.

Find: The equivalent impedance $\mathbf{Z}_1$ across the parallel elements.

Analysis: Combine the two impedances in parallel to determine the equivalent impedance.

$$\mathbf{Z}_1 = R_1 \left\| \frac{1}{j\omega C_1} = \frac{R_1(1/j\omega C_1)}{R_1 + 1/j\omega C_1} = \frac{R_1}{1 + j\omega C_1 R_1} \right.$$

Substitute numerical values to find:

$$\mathbf{Z}_1(\omega = 377) = \frac{10^6}{1 + j377 \times 10^{-9} \times 10^6} = \frac{10^6}{1 + j0.377}$$

$$= 9.36 \times 10^5 \angle(-0.36)\ \Omega$$

The impedance of the capacitor alone at $\omega = 377$ rad/s is

$$\mathbf{Z}_{C1}(\omega = 377) = \frac{1}{j377 \times 10^{-9}} = 2.65 \times 10^{6} \angle(-1.57)\ \Omega$$

When the frequency is increased to 800 kHz, or $1600\pi \times 10^3$ rad/s—a radio frequency in the AM range—the impedance changes to:

$$\mathbf{Z}_1(\omega = 1600\pi \times 10^3) = \frac{10^6}{1 + j1600\pi \times 10^3 \times 10^{-9} \times 10^6}$$

$$= \frac{10^6}{1 + j1600\pi} = 198.9 \angle(-1.57)\ \Omega$$

Page 198

The impedance of the capacitor alone at $\omega = 1600\pi \times 10^3$ rad/s is

$$\mathbf{Z}_{C1}(\omega = 1600\pi \times 10^3) = \frac{1}{j1600\pi \times 10^3 \times 10^{-9}} = 198.9 \angle(-1.57)\ \Omega$$

Now, the impedances $\mathbf{Z}_1$ and $\mathbf{Z}_{C1}$ are virtually identical. Thus, the effect of the parallel resistance is negligible at high frequencies.

Comments: For elements in parallel, the element with the smallest impedance tends to dominate the equivalent impedance across two nodes. At the lower frequency (corresponding to the well-known 60-Hz AC power frequency) the impedance of the resistor is roughly 38 percent smaller than that of the ideal capacitor. Thus, the resistor tends to dominate the equivalent impedance at 60-Hz; in fact, at that frequency the equivalent impedance is only 6.5 percent smaller than the resistance and so the practical and ideal capacitors are substantially different. At the higher frequency, the impedance of the ideal capacitor is much smaller than the resistance. The equivalent impedance is dominated by the ideal capacitor. At frequencies above and below $\omega = 1/RC$, the network is capacitive and resistive, respectively. This example suggests that the behavior of a network may depend heavily on frequency.

EXAMPLE 3.11 Impedance of a Practical Inductor

Problem

Figure 3.35 shows a toroidal (doughnut-shaped) inductor. A practical inductor is often modeled as an ideal inductor in series with a resistor, as shown in Figure 3.36. The series resistance represents the resistance of the wire. Find the range of frequencies over which the impedance of the practical inductor is largely inductive. Consider the impedance to be inductive if it is at least 10 times larger than the resistance.

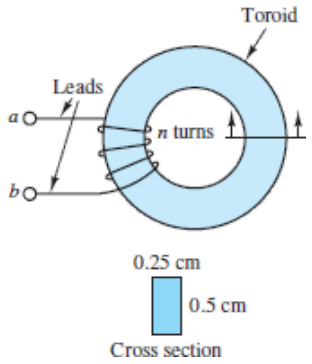


Figure 3.35 A practical inductor

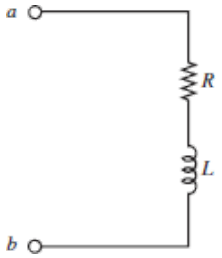


Figure 3.36

Solution

Known Quantities: $L = 0.098$ H; lead length = 2×10 cm; $n = 250$ turns; wire is 30 gauge. Resistance of 30-gauge wire = 0.344 Ω/m.

Find: The range of frequencies over which the practical inductor acts nearly as an ideal inductor.

Analysis: To determine the equivalent resistance of the wire, use the cross section of the toroid to estimate its length l_w:

$$l_w = 250(2 \times 0.25 + 2 \times 0.5) = 375\,\text{cm}$$
$$\text{Total length} = 375 + 20 = 395\,\text{cm}$$

Thus, the total resistance is

$$R = 0.344\ \Omega/\text{m} \times 3.95\ \text{m} = 1.36\ \Omega$$

To determine the range of frequencies, ω, over which the impedance $j\omega L$ of the ideal inductor is 10× greater than 1.36 Ω:

$$\omega L > 13.6 \quad \text{or} \quad \omega > \frac{13.6}{L} = \frac{13.6}{0.098} = 139\ \text{rad/s}$$

In terms of cyclical frequency, the range is $f = \omega/2\pi > 22$ Hz.

Page 199

Comments: For elements in series, the element with the largest impedance tends to dominate the equivalent impedance across two nodes. At frequencies above 139 rad/s the impedance of the inductor is at least 10× greater than the resistance and the resistance is insignificant. (Remember the 10:1 rule.) At lower frequencies, the resistance is significant; at very low frequencies ($\omega L << R$), the impedance of the ideal inductor effectively acts as a short-circuit and is negligible. At high frequencies, the separation between the insulated coil wires begins to exhibit significant capacitance and so the model should be modified accordingly.

EXAMPLE 3.12 Impedance of a Series-Parallel Network

Problem

Find the equivalent impedance of the circuit shown in Figure 3.37.

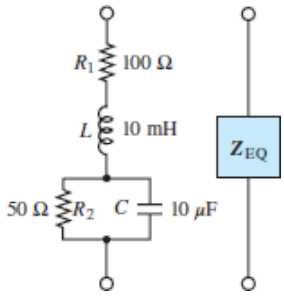


Figure 3.37

Solution

Known Quantities: $\omega = 10^4$ rad/s; $R_1 = 100\ \Omega$; $L = 10$ mH; $R_2 = 50\ \Omega$; $C = 10\ \mu$F.

Find: The equivalent impedance of the series-parallel circuit.

Analysis: The equivalent impedance of R_2 in parallel with C is

$$\begin{aligned}\mathbf{Z}_{\|} &= R_2 \left\| \frac{1}{j\omega C} = \frac{R_2(1/j\omega C)}{R_2 + 1/j\omega C} = \frac{R_2}{1 + j\omega C R_2} \right. \\ &= \frac{50}{1 + j\,10^4 \times 10 \times 10^{-6} \times 50} = \frac{50}{1 + j5} = 1.92 - j9.62\ \Omega \\ &= 9.81\angle(-1.3734)\ \Omega\end{aligned}$$

To determine the equivalent impedance $\mathbf{Z}_{eq}$ across the entire network:

$$\begin{aligned}\mathbf{Z}_{eq} &= R_1 + j\omega L + \mathbf{Z}_{\|} = 100 + j\,10^4 \times 10^{-2} + 1.92 - j9.62 \\ &= 101.92 + j90.38 = 136.2\angle 0.725\ \Omega\end{aligned}$$

Comment: At $\omega = 10^4$ rad/s, the impedance across the network is inductive since the reactance is positive (or, equivalently, the phase angle is positive). (See Figure 3.31.)

EXAMPLE 3.13 Admittance

Problem

Find the equivalent admittance across each of the two networks shown in Figure 3.38.

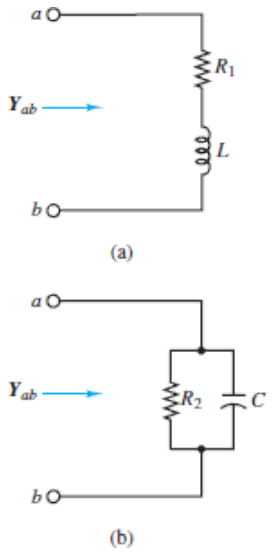


Figure 3.38

Solution

Known Quantities: $\omega = 2\pi \times 10^3$ rad/s; $R_1 = 50\ \Omega$; $L = 16$ mH; $R_2 = 100\ \Omega$; $C = 3\ \mu$F.

Find: The equivalent admittance across each of the two networks.

Analysis: Network (a): First, determine the equivalent impedance across the network *ab:*

$$\mathbf{Z}_{ab} = R_1 + j\omega L$$

Page 200

To obtain the admittance, compute the inverse of $\mathbf{Z}_{ab}$ by multiplying the numerator and denominator by the complex conjugate of the denominator:

$$\mathbf{Y}_{ab} = \frac{1}{\mathbf{Z}_{ab}} = \frac{1}{R_1 + j\omega L} = \frac{R_1 - j\omega L}{R_1^2 + (\omega L)^2}$$

Substitute numerical values to find:

$$\mathbf{Y}_{ab} = \frac{1}{50 + j2\pi \times 10^3 \times 0.016} = \frac{50 - j(2\pi \times 10^3)(0.016)}{50^2 + (2\pi \times 10^3)^2 (0.016)^2}$$
$$\approx 4.0 \times 10^{-3} - j8.0 \times 10^{-3} \quad \text{S}$$

Network (b): First, determine the equivalent impedance across the network *ab*:

$$\mathbf{Z}_{ab} = R_2 \left\| \frac{1}{j\omega C} = \frac{R_2(1/j\omega C)}{R_2 + (1/j\omega C)} \right.$$

Multiply the numerator and denominator by $j\omega C$ to find:

$$\mathbf{Z}_{ab} = \frac{R_2}{1 + j\omega R_2 C}$$

The inverse of $\mathbf{Z}_{ab}$ is the admittance:

$$\mathbf{Y}_{ab} = \frac{1}{\mathbf{Z}_{ab}} = \frac{1 + j\omega R_2 C}{R_2} = \frac{1}{R_2} + j\omega C = 0.01 + j0.019 \quad \text{S}$$

Comment: The units of admittance and conductance are the same, siemens (S).

CHECK YOUR UNDERSTANDING

Add the sinusoidal voltages $v_1(t) = A\cos(\omega t + \phi)$ and $v_2(t) = B\cos(\omega t + \theta)$ using phasor notation, and then convert back to time-domain form.

a. $A = 1.5$ V, $\phi = 10°$; $B = 3.2$ V, $\theta = 25°$.

b. $A = 50$ V, $\phi = -60°$; $B = 24$ V, $\theta = 15°$.

Answer: (a) $v_1 + v_2 = 4.67\cos(\omega t + 0.353$ rad); (b) $v_1 + v_2 = 60.8\cos(\omega t -$ 0.656 rad)

CHECK YOUR UNDERSTANDING

Add the sinusoidal currents $i_1(t) = A\cos(\omega t + \phi)$ and $i_2(t) = B\cos(\omega t + \theta)$ for

a. $A = 0.09$ A, $\phi = 72°$; $B = 0.12$ A, $\theta = 20°$.

b. $A = 0.82$ A, $\phi = -30°$; $B = 0.5$ A, $\theta = -36°$.

Answer: (a) $i_1 + i_2 = 0.19\cos(\omega t + 0.733)$; (b) $i_1 + i_2 = 1.32\cos(\omega t -$ 0.5633)

CHECK YOUR UNDERSTANDING

Compute the equivalent impedance across the network of Example 3.12 for ω = 1,000 and 100,000 rad/s.

Find the reactance across the parallel R_2C network of Example 3.12 at the frequency ω = 10 rad/s and calculate its equivalent capacitance.

Answer: $\mathbf{Z}(1{,}000) = 140 - j10;\ \mathbf{Z}(100{,}000) = 100 + j999;\ |X| = 0.25;\ C =$ 0.4 F

CHECK YOUR UNDERSTANDING

Compute the equivalent admittance across the network of Example 3.12.

Answer: $Y_{eq} = 5.492 \times 10^{-3} - j4.871 \times 10^{-3}$

3.6 AC CIRCUIT ANALYSIS

Phasors and the concept of impedance, in particular, facilitate the solution of AC circuits by making it possible to use the same solution methods developed in Chapter 2 for DC circuits. Those methods are explored for AC circuits containing linear passive circuit elements (R, L, C) excited by a sinusoidal source. Figure 3.39 depicts one such circuit, represented in both conventional time-domain and phasor-impedance forms. Notice the differences in notation.

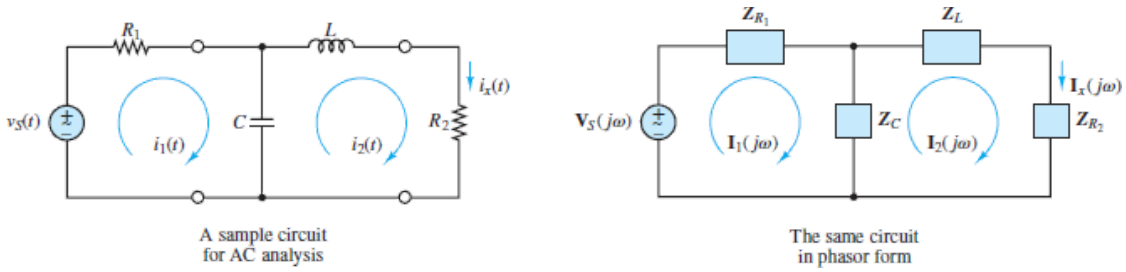


Figure 3.39 An AC circuit

The first step in AC circuit analysis is to convert all sources to phasor form and use the frequency of excitation to determine the impedance of each passive element. Each passive impedance element will have an amplitude and a phase, both of which may depend upon the excitation frequencyω.

The second step is to apply the same solution methods previously explored in Chapters 1 and 2, treating each impedance element in the same manner as resistors were treated in those chapters. The only difference is that AC circuit phasor analysis involves complex arithmetic whereas DC circuit analysis involves scalar Page 202arithmetic. For example, voltage division for two impedance elements in series will have the form:

$$\frac{\mathbf{V}_1}{\mathbf{V}_2} = \frac{\mathbf{Z}_1}{\mathbf{Z}_2}$$

Similarly, current division for two impedance elements in parallel will have the form:

$$\frac{\mathbf{I}_1}{\mathbf{I}_2} = \frac{\mathbf{Z}_2}{\mathbf{Z}_1}$$

Compare these expressions to those found in Chapter 2 for resistors in series and parallel. The applications of KVL, KCL, Ohm's law, the node voltage and mesh current methods, Thévenin's and Norton's theorems, superposition, and source transformations to AC circuits are all identical to their applications in DC circuits, except that impedance **Z** and source phasors take the place of resistance and source scalars.

The solution of an AC circuit problem will be a phasor, in general. Thus, the third and last step is to convert the solution to its time-domain form. In effect, the use of phasor notation is but an intermediate step that facilitates the computation of the final answer.

It is worth noting that although it is possible to extend the node voltage and mesh current methods to AC circuits, the resulting simultaneous complex equations will usually be difficult to solve without the aid of a scientific calculator or computer,

even for relatively simple circuits. In addition, these methods lend relatively little insight into the nature of the circuit. On the other hand, it is very useful to extend the concept of equivalent networks to the AC case and to make use of complex Thévenin and Norton equivalent impedances.

FOCUS ON PROBLEM SOLVING

AC CIRCUIT ANALYSIS

1. Identify the sinusoidal sources in a circuit and note the excitation frequencies
2. Convert the sources to phasor form.
3. Use the excitation frequency to determine the impedance of each pas element.
4. Solve the resulting phasor circuit, using an appropriate solution method, suc Thévenin's theorem, Norton's theorem, superposition, source transforma and the node voltage and mesh current methods. Take care to execute complex arithmetic properly. Represent the solution as a phasor.
5. Convert the phasor solution to its time-domain form.

Page 203

AC Equivalent Circuits

The concept of an equivalent circuit is equally useful in AC and DC circuit analyses. Figure 3.40(a) depicts the one-port source-load perspective first introduced in Chapter 2. In the figure, the overall circuit is divided into two parts: a load and a source. Typically, the load is the element or circuit segment of interest to the analyst. The source is everything else not included in the load. The source and load are connected at two terminals *a* and *b*.

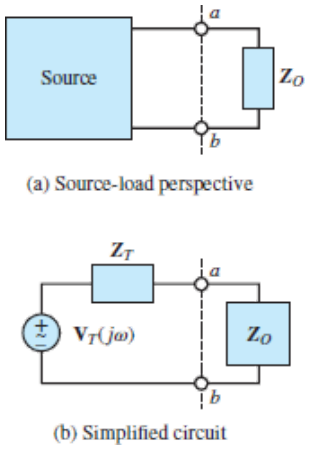


Figure 3.40 AC circuit simplification using Thévenin's theorem

Thévenin's or Norton's theorem can be used to simplify the source network as shown in Figure 3.40(b). Notice that the Thévenin equivalent source is composed of two phasors: an independent voltage source $\mathbf{V}_T(j\omega)$ in series with an equivalent impedance $\mathbf{Z}_T(j\omega)$. The voltage $\mathbf{V}_o$ across the load $\mathbf{Z}_o(j\omega)$ can be found by voltage division.

$$\frac{\mathbf{V}_o}{\mathbf{V}_T} = \frac{\mathbf{Z}_o}{\mathbf{Z}_o + \mathbf{Z}_T}$$

Notice that the approach and form of this solution is exactly the same as that previously presented in Chapter 2 for DC circuits. The only difference is the use of impedance instead of resistance. The independent Thévenin voltage source $\mathbf{V}_T$ is the open-circuit voltage $\mathbf{V}_{OC}$ across the terminals a and b of the source network. The Thévenin equivalent impedance $\mathbf{Z}_T$ is found by setting to zero all independent voltage and current sources in the source network and then finding the equivalent impedance $\mathbf{Z}_{ab}$ between the source network terminals. Once again, a zero voltage source acts as a short-circuit and a zero current source acts as an open-circuit. The rules for finding the equivalent impedance and admittance of series and parallel networks are shown in Figure 3.41.

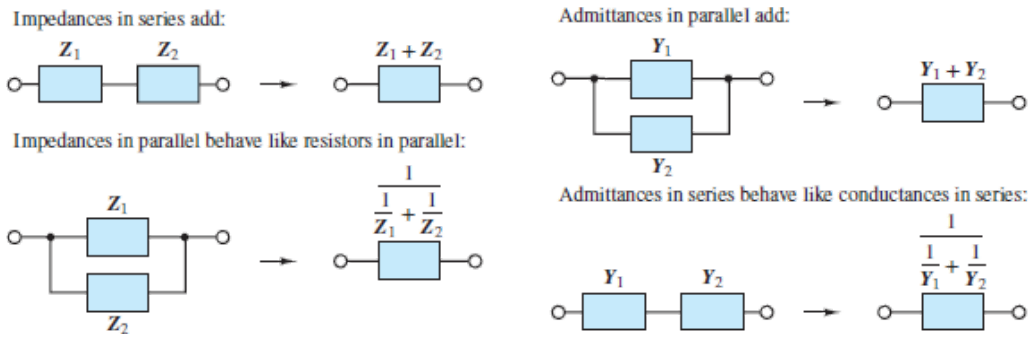


Figure 3.41 Rules for impedance and admittance reduction

The Norton equivalent source is also composed of two phasors: an independent current source $\mathbf{I}_N$ in parallel with the same equivalent impedance $\mathbf{Z}_N(j\omega) = \mathbf{Z}_T(j\omega)$. Figure 3.42 depicts a somewhat complicated AC circuit partitioned into source and load networks, and how $\mathbf{V}_T$, $\mathbf{I}_N$, and $\mathbf{Z}_T$ are found. The example problems that follow clarify some of the finer points in the calculation of such equivalent circuits. The details of the complex arithmetic are also explored.

Page 204

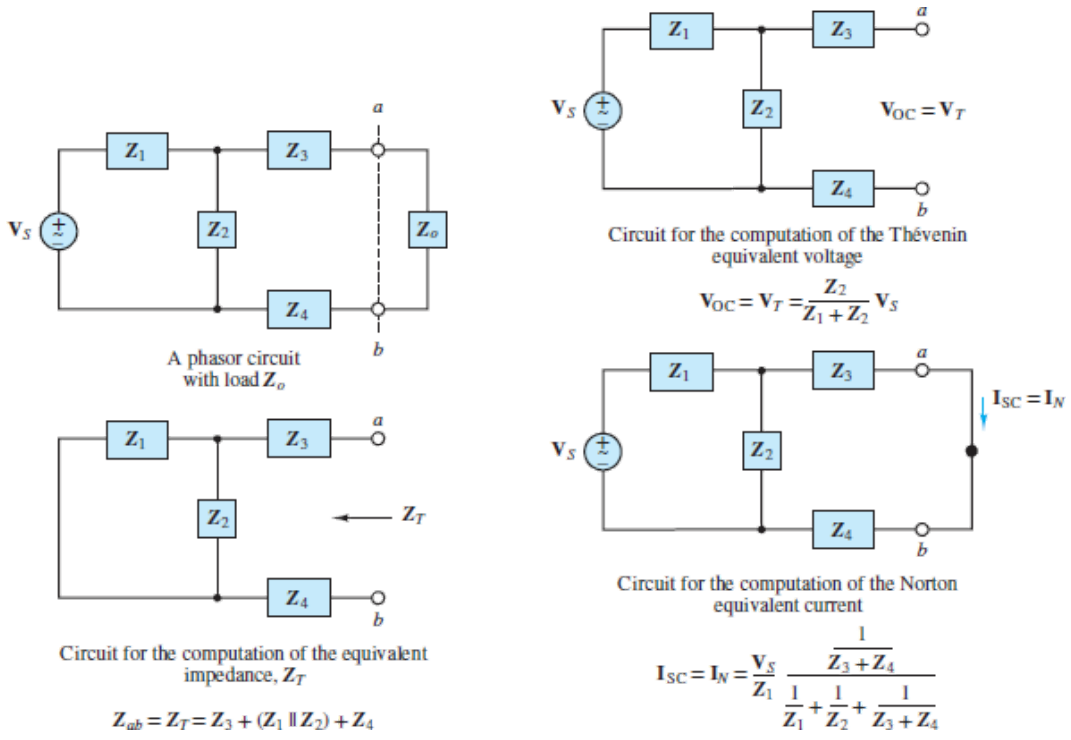


Figure 3.42 Reduction of AC circuit to equivalent form

FOCUS ON MEASUREMENTS

Capacitive Displacement Transducer

As introduced in the previous Focus on Measurements section, a displacement transducer consists of a parallel-plate capacitor with a variable separation distance x. The capacitance was shown to be:

$$C = \frac{8.854 \times 10^{-3} A}{x} \quad \text{pF}$$

where C is in picofarads, the area of the plates A is in square millimeters, and x is in millimeters. The impedance of the capacitor is

$$\mathbf{Z}_C = \frac{1}{j\omega C} = \frac{x}{j\omega(8.854 \times 10^{-3})A} \quad \text{T}\Omega$$

Thus, at a given frequency ω, the impedance of the capacitor varies linearly with the separation distance. This result can be exploited in a bridge circuit, as shown in Figure 3.7 where half of the bridge is a differential pressure transducer in which a thin diaphragm (plate) is situated between two fixed plates and subject to variations in pressure across the diaphragm. The result is that when the capacitance of one leg of the bridge, shown here again as Figure 3.43, increases, the capacitance of the other leg decreases. Assume the bridge is excited by a sinusoidal source.

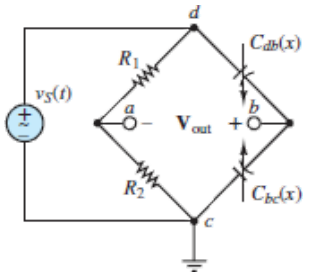


Figure 3.43 Bridge circuit for capacitive displacement transducer

Apply voltage division and KVL to express the output voltage in phasor notation as:

$$\mathbf{V}_{out}(j\omega) = \mathbf{V}_S(j\omega)\left(\frac{\mathbf{Z}_{C_{bc}}(x)}{\mathbf{Z}_{C_{db}}(x) + \mathbf{Z}_{C_{bc}}(x)} - \frac{R_2}{R_1 + R_2}\right)$$

When the diaphragm is not displaced from its center position, the nominal capacitance of each half of the transducer is given by:

$$C_0 = \frac{\varepsilon A}{d}$$

where d is the nominal separation distance between the diaphragm and the fixed surfaces (in millimeters). Thus, when the diaphragm is displaced an effective distance Δx, the capacitance of each leg of the bridge is given by:

$$C_{db} = \frac{\varepsilon A}{d - \Delta x} \quad \text{and} \quad C_{bc} = \frac{\varepsilon A}{d + \Delta x}$$

Therefore, the corresponding impedance of each leg is:

$$\mathbf{Z}_{C_{db}} = \frac{d - \Delta x}{j\omega(8.854 \times 10^{-3})A} \quad \text{and} \quad \mathbf{Z}_{C_{bc}} = \frac{d + \Delta x}{j\omega(8.854 \times 10^{-3})A}$$

such that the phasor output voltage is:

$$\begin{aligned}\mathbf{V}_{out}(j\omega) &= \mathbf{V}_S(j\omega)\left(\frac{\dfrac{d + \Delta x}{j\omega(8.854 \times 10^{-3})A}}{\dfrac{d - \Delta x}{j\omega(8.854 \times 10^{-3})A} + \dfrac{d + \Delta x}{j\omega(8.854 \times 10^{-3})A}} - \frac{R_2}{R_1 + R_2}\right) \\ &= \mathbf{V}_S(j\omega)\left(\frac{1}{2} + \frac{\Delta x}{2d} - \frac{R_2}{R_1 + R_2}\right) \\ &= \mathbf{V}_S(j\omega)\frac{\Delta x}{2d} \qquad \text{(assuming } R_1 = R_2\text{)}\end{aligned}$$

Thus, the output voltage will vary as a scaled version of the input voltage in proportion to the displacement. A typical $v_{out}(t)$ is displayed in Figure 3.44 for a 0.05-mm "triangular" diaphragm displacement, with d = 0.5 mm and $\mathbf{V}_S$ a 25-Hz sinusoid with 1-V amplitude.

Page 206

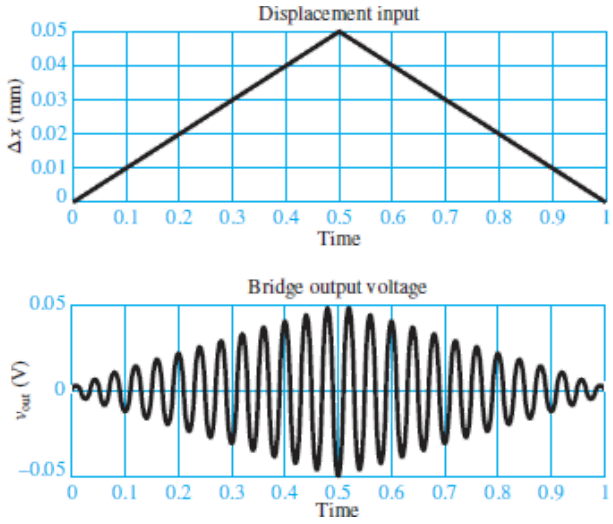


Figure 3.44 Displacement input and bridge output voltage for capacitive displacement transducer

EXAMPLE 3.14 Phasor Analysis of an AC Circuit

Problem

Apply the phasor analysis method to the circuit of Figure 3.45 to determine the source current.

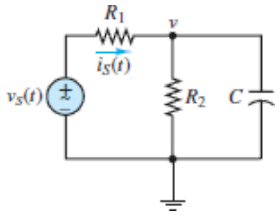


Figure 3.45

Solution

Known Quantities: Figures 3.45, 3.46, $v_S(t) = 10 \cos \omega t$ V; $\omega = 377$ rad/s; $R_1 = 50\ \Omega$; $R_2 = 200\ \Omega$; $C = 100\ \mu$F.

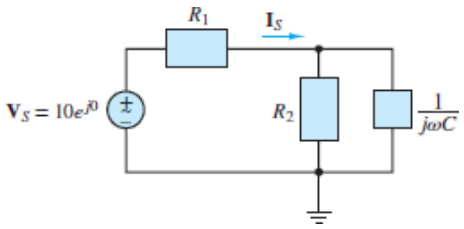


Figure 3.46

Find: The source current $i_S(t)$.

Analysis: Define the voltage v at the top right node, and use the node voltage method to determine v. Then observe that

$$i_S(t) = \frac{v_S(t) - v(t)}{R_1}$$

Page 207

Next, we follow the steps of the Focus on Problem Solving box "AC Circuit Analysis."

Step 1: $v_S(t) = 10 \cos \omega t$ V $\quad \omega = 377$ rad/s $\quad (f = 60$ Hz$)$

Step 2: $\mathbf{V}_S = 10\angle 0$ V

Step 3: $\mathbf{Z}_{R_1} = R_1 \quad \mathbf{Z}_{R_2} = R_2 \quad \mathbf{Z}_C = \frac{1}{j\omega C}$

The resulting phasor circuit is shown in Figure 3.46.

Step 4: Solve for the source current using the node voltage method. Apply KCL at the upper right node to find:

$$\frac{\mathbf{V}_S - \mathbf{V}}{\mathbf{Z}_{R_1}} = \frac{\mathbf{V}}{\mathbf{Z}_{R_2} \| \mathbf{Z}_C}$$

$$\frac{\mathbf{V}_S}{\mathbf{Z}_{R_1}} = \mathbf{V}\left(\frac{1}{\mathbf{Z}_{R_2} \| \mathbf{Z}_C} + \frac{1}{\mathbf{Z}_{R_1}}\right) = \mathbf{V}\left(\frac{1}{\frac{R_2 \cdot (1/j\omega C)}{R_2 + (1/j\omega C)}} + \frac{1}{R_1}\right)$$

$$= \mathbf{V}\left(\frac{j\omega C R_2 + 1}{R_2} + \frac{1}{R_1}\right) = \mathbf{V}\left[\frac{(j\omega C R_2 R_1 + R_1) + R_2}{R_1 R_2}\right]$$

Thus:

$$\begin{aligned}\mathbf{V} &= \left[\frac{R_1R_2}{(j\omega C R_2R_1 + R_1) + R_2}\right]\frac{\mathbf{V}_S}{R_1} \\ &= \left[\frac{50 \times 200}{(j377 \times 10^{-4} \times 50 \times 200 + 50) + 200}\right]\frac{\mathbf{V}_S}{50} \\ &= 0.44\angle(-0.99)\,\mathbf{V}_S = 4.4\angle(-0.99)\text{ V}\end{aligned}$$

Then, compute $\mathbf{I}_S$:

$$\mathbf{I}_S = \frac{\mathbf{V}_S - \mathbf{V}}{\mathbf{Z}_{R_1}} = \frac{10\angle 0 - 4.4\angle(-0.99)}{50} = 0.17\angle(0.45)\text{ A}$$

Step 4: Finally, convert the solution to its time-domain form:

$$i_S(t) = 0.17\cos(377t + 0.45)\text{ A}$$

EXAMPLE 3.15 An AC Circuit With an Arbitrary Sinusoidal Input

Problem

Determine the general solution of Example 3.14 for any sinusoidal source, $A\cos(\omega t + \phi)$.

Solution

Known Quantities: $R_1 = 50\ \Omega$; $R_2 = 200\ \Omega$, $C = 100\ \mu$F.

Find: The phasor source current $\mathbf{I}_S(j\omega)$.

Analysis: Since the radian frequency is arbitrary, it will be impossible to determine a numerical answer. The answer will be a function of ω. The source phasor is: $\mathbf{V}_S(j\omega) = A\angle\phi$. The impedances will be $\mathbf{Z}_{R_1} = 50\ \Omega$; $\mathbf{Z}_{R_2} = 200\ \Omega$; $\mathbf{Z}_C = -j10^4/\omega\Omega$. Note that the impedance of the capacitor is a function of ω.

Observe that the source current is given by:

$$\mathbf{I}_S = \frac{\mathbf{V}_S}{\mathbf{Z}_{R_1} + \mathbf{Z}_{R_2} \| \mathbf{Z}_C}$$

The parallel impedance $\mathbf{Z}_{R_2}||\mathbf{Z}_C$ is given by:

$$\mathbf{Z}_{R_2}||\mathbf{Z}_C = \frac{\mathbf{Z}_{R_2} \times \mathbf{Z}_C}{\mathbf{Z}_{R_2} + \mathbf{Z}_C} = \frac{200 \times 10^4/j\omega}{200 + 10^4/j\omega} = \frac{2 \times 10^6}{10^4 + j\omega 200} \quad \Omega$$

Thus, the total series impedance is:

$$\mathbf{Z}_{R_1} + \mathbf{Z}_{R_2}||\mathbf{Z}_C = 50 + \frac{2 \times 10^6}{10^4 + j\omega 200} = \frac{2.5 \times 10^6 + j\omega 10^4}{10^4 + j\omega 200} \quad \Omega$$

and the phasor source current is:

$$\mathbf{I}_S = \frac{\mathbf{V}_S}{\mathbf{Z}_{R_1} + \mathbf{Z}_{R_2}||\mathbf{Z}_C} = A\angle\phi \frac{10^4 + j\omega 200}{2.5 \times 10^6 + j\omega\, 10^4} \quad \text{A}$$

Comment: The expression obtained in this example can be evaluated for a specific sinusoidal excitation, by substituting numerical values for A, ϕ, and ω. Use the values from Example 3.14 (A = 10 V, ϕ = 0 rad, ω = 377 rad/s) and verify that the same answer is obtained.

EXAMPLE 13.16 Solving an AC Circuit by the Node Voltage Method

Problem

The fundamental electrical characteristics of electric motors can be approximated by a series *RL* circuit. In this problem, a voltage source provides current to two different motors (Figure 3.47).

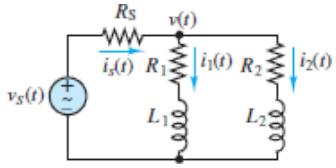


Figure 3.47 AC circuit used to demonstrate node analysis

Solution

Known Quantities: $R_S = 0.5\ \Omega$; $R_1 = 2\ \Omega$; $R_2 = 0.2\ \Omega$; $L_1 = 0.1$ H; $L_2 = 20$ mH; $v_S(t) = 155\cos(377t)$ V.

Find: The motor load currents $i_1(t)$ and $i_2(t)$.

Analysis: First, calculate the impedances of the source and of each motor:

$$\mathbf{Z}_S = 0.5\ \Omega$$
$$\mathbf{Z}_1 = 2 + j377 \times 0.1 = 2 + j37.7 = 37.8\angle 1.52\ \Omega$$
$$\mathbf{Z}_2 = 0.2 + j377 \times 0.02 = 0.2 + j7.54 = 7.54\angle 1.54\ \Omega$$

The source voltage is $\mathbf{V}_S = 155\angle 0$ V.

Page 209

Next, apply KCL at the top right node, with the aim of solving for the node voltage **V**:

$$\frac{\mathbf{V}_S - \mathbf{V}}{\mathbf{Z}_S} = \frac{\mathbf{V}}{\mathbf{Z}_1} + \frac{\mathbf{V}}{\mathbf{Z}_2}$$
$$\frac{\mathbf{V}_S}{\mathbf{Z}_S} = \frac{\mathbf{V}}{\mathbf{Z}_S} + \frac{\mathbf{V}}{\mathbf{Z}_1} + \frac{\mathbf{V}}{\mathbf{Z}_2} = \mathbf{V}\left(\frac{1}{\mathbf{Z}_S} + \frac{1}{\mathbf{Z}_1} + \frac{1}{\mathbf{Z}_2}\right)$$
$$\mathbf{V} = \left(\frac{1}{0.5} + \frac{1}{2 + j37.7} + \frac{1}{0.2 + j7.54}\right)^{-1} \frac{\mathbf{V}_S}{0.5}$$
$$= 154\angle 0.079\,\mathrm{V}$$

It is now easy to find the phasor motor currents, $\mathbf{I}_1$ and $\mathbf{I}_2$, from the phasor node voltage **V**.

$$\mathbf{I}_1 = \frac{\mathbf{V}}{\mathbf{Z}_1} = \frac{154\angle 0.079}{2 + j37.7} = 4.1\angle -1.44$$
$$\mathbf{I}_2 = \frac{\mathbf{V}}{\mathbf{Z}_2} = \frac{154\angle 0.079}{0.2 + j7.54} = 20.4\angle -1.47$$

Finally, write the time-domain expressions for the currents:

$$i_1(t) = 4.1\cos(377t - 1.44)\,\mathrm{A}$$
$$i_2(t) = 20.4\cos(377t - 1.47)\,\mathrm{A}$$

Figure 3.48 depicts the source voltage (scaled down by a factor of 10) and the two motor currents.

Comment: Note the phase shift between the source voltage and the two motor currents.

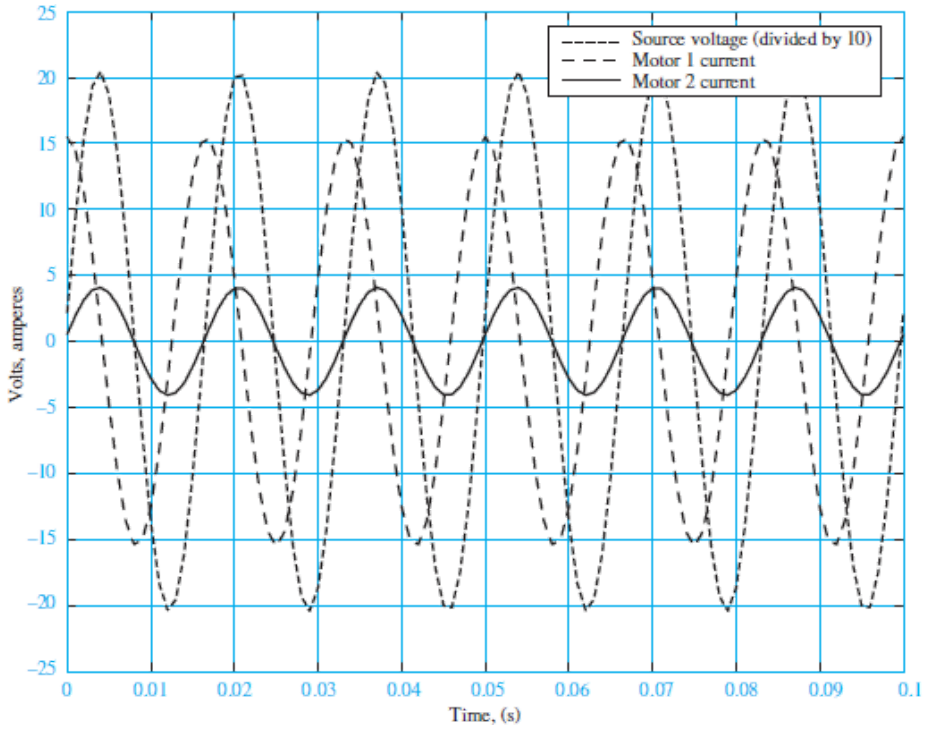


Figure 3.48 Source voltage and motor currents for Example 3.16

Page 210

LO

EXAMPLE 3.17 AC Superposition

Problem

Compute the voltages $v_1(t)$ and $v_2(t)$ in the circuit of Figure 3.49.

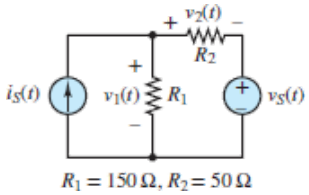


Figure 3.49

Solution

Known Quantities:

$$i_S(t) = 0.5\cos[2\pi(100t)] \quad \text{A}$$
$$v_S(t) = 20\cos[2\pi(1{,}000t)] \quad \text{V}$$

Find: $v_1(t)$ and $v_2(t)$.

Analysis: Since the two sources are at different frequencies, compute a separate solution for each. Consider the current source first, with the voltage source set to zero (short-circuit) as shown in Figure 3.50. The resulting circuit is a simple current divider. Write the source current in phasor notation:

$$\mathbf{I}_S(j\omega) = 0.5e^{j0} = 0.5\angle 0 \text{ A} \qquad \omega = 200\pi \text{ rad/s}$$

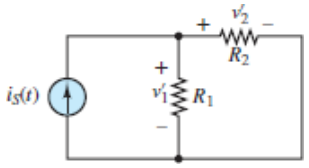


Figure 3.50

Then

$$\mathbf{V}_1' = \mathbf{I}_S \frac{R_2}{R_1 + R_2} R_1 = 0.5\angle 0 \left(\frac{50}{150+50}\right) 150 = 18.75\angle 0 \text{ V}$$
$$\mathbf{V}_2' = \mathbf{I}_S \frac{R_1}{R_1 + R_2} R_2 = 0.5\angle 0 \left(\frac{150}{150+50}\right) 50 = 18.75\angle 0 \text{ V}$$

Next, consider the voltage source, with the current source set to zero (equivalent to an open-circuit), as shown in Figure 3.51. First, write the source voltage in phasor notation:

$$\mathbf{V}_S(j\omega) = 20e^{j0} = 20\angle 0 \text{ V} \qquad \omega = 2{,}000\pi \text{ rad/s}$$

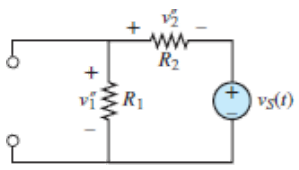


Figure 3.51

Then, apply the voltage divider law, to obtain:

$$\mathbf{V}_1'' = \mathbf{V}_S \frac{R_1}{R_1 + R_2} = 20\angle 0\left(\frac{150}{150 + 50}\right) = 15\angle 0 \text{ V}$$

$$\mathbf{V}_2'' = -\mathbf{V}_S \frac{R_2}{R_1 + R_2} = -20\angle 0\left(\frac{50}{150 + 50}\right) = -5\angle 0 = 5\angle \pi \text{ V}$$

The voltage across each resistor is obtained by adding the contributions from each source and converting the equivalent phasor to the time domain:

$$\mathbf{V}_{R1} = \mathbf{V}_1' + \mathbf{V}_1''$$
$$v_1(t) = 18.75\cos[2\pi(100t)] + 15\cos[2\pi(1{,}000t)] \quad \text{V}$$

and

$$\mathbf{V}_{R2} = \mathbf{V}_2' + \mathbf{V}_2''$$
$$v_2(t) = 18.75\cos[2\pi(100t)] + 5\cos[2\pi(1{,}000t) + \pi] \quad \text{V}$$

Comment: It is impossible to further simplify the final expression because the two components are at different frequencies.

Page 211

EXAMPLE 3.18 Using Thévenin's Theorem to Solve an AC Circuit Problem

Compute the Thévenin equivalent network seen by the load $\mathbf{Z}_o$ in Figure 3.52.

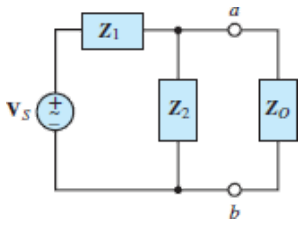


Figure 3.52

Solution

Known Quantities: $\mathbf{Z}_1 = 5\ \Omega$; $\mathbf{Z}_2 = j20\ \Omega$, $v_S(t) = 110\cos(377t)$ V.

Find: Thévenin equivalent network seen by the load $\mathbf{Z}_o$.

Analysis: Find the Thévenin equivalent impedance seen by the load $\mathbf{Z}_o$. Remove the load, replace the independent voltage source with a short-circuit, and compute the equivalent impedance between terminals *a* and *b* shown in Figure 3.53.

$$\mathbf{Z}_T = \mathbf{Z}_1 \| \mathbf{Z}_2 = \frac{\mathbf{Z}_1 \times \mathbf{Z}_2}{\mathbf{Z}_1 + \mathbf{Z}_2} = \frac{5 \times j20}{5 + j20} = 4.71 + j1.176\ \Omega$$

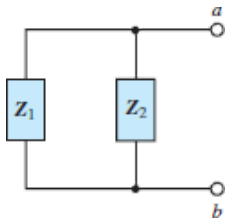


Figure 3.53

Next, apply voltage division to find the open-circuit (Thévenin) voltage across terminals *a* and *b*:

$$\mathbf{V}_T = \frac{\mathbf{Z}_2}{\mathbf{Z}_1 + \mathbf{Z}_2}\mathbf{V}_S = \frac{j20}{5 + j20}110\angle 0 = \frac{20\angle \pi/2}{20.6\angle 1.326}110\angle 0 = 106.7\angle 0.245\ \text{V}$$

The complete simplified circuit is shown in Figure 3.54.

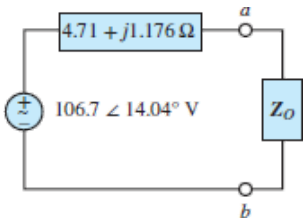


Figure 3.54

Comments: The procedure to simplify the circuit is identical to that used for resistive circuits, the only difference being the use of complex impedances in place of resistances.

EXAMPLE 3.19 Thévenin's Theorem Applied to an AC Circuit

Problem

Determine the Thévenin equivalent network seen by the load R_o in the circuit of Figure 3.55 when the input sinusoidal voltage is (a) at a frequency of 10^3 Hz and (b) at a frequency of 10^6 Hz.

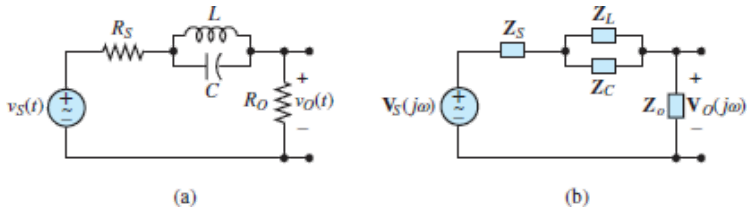


Figure 3.55 (a) Circuit for Example 3.19; (b) same circuit ready for phasor analysis

Solution

Known Quantities: $R_S = R_o = 50\ \Omega$, $C = 0.1\ \mu$F; $L = 10$ mH.

Page 212

Analysis: First, convert the circuit to phasor form, as shown in Figure 3.55(b). Next, compute the Thévenin equivalent impedance with the load removed:

$$\mathbf{Z}_T = \mathbf{Z}_S + \mathbf{Z}_L \| \mathbf{Z}_C = R_S + \frac{j\omega L \times 1/j\omega C}{j\omega L + 1/j\omega C} \cdot \frac{j\omega C}{j\omega C}$$

$$= R_S + \frac{j\omega L}{j\omega L \times j\omega C + 1} = R_S + j\frac{\omega L}{1 - \omega^2 LC}$$

Observe that the Thévenin equivalent voltage is equal to the source voltage, since once the load is removed, the current is zero in the resulting network and the voltage drop across the remaining impedances is zero. Thus:

$$\mathbf{V}_T = \mathbf{V}_S$$

Next, evaluate the Thévenin equivalent at each of the two frequencies.

a. Let $f = 10^3$ Hz. Then $\omega = 2\pi \times 10^3$ rad/s. At this frequency,

$$\mathbf{Z}_T = R_S + j\frac{\omega L}{1 - \omega^2 LC} = 50 + j65.4 = 82.3\angle 0.92$$

b. Let $f = 10^6$ Hz. Then $\omega = 2\pi \times 10^6$ rad/s. At this frequency,

$$\mathbf{Z}_T = R_S + j\frac{\omega L}{1-\omega^2 LC} = 50 - j1.59 = 50\angle(-0.032)$$

Comments: Note that at the higher frequency the equivalent impedance is very close to that of the resistor R_S. This result is due to the action of the capacitor and inductor, which behave like short- and open-circuits, respectively, at high frequency. The very small impedance of the capacitor at high frequency dominates the parallel equivalent impedance.

EXAMPLE 3.20 Solution of an AC Circuit by the Mesh Current Method

Problem

Determine the currents $i_1(t)$ and $i_2(t)$ in the circuit of Figure 3.56, using the mesh current method.

Solution

Known Quantities: $R_1 = 100\ \Omega$; $R_2 = 75\ \Omega$; $C = 1\ \mu$F; $L = 0.5$ H; $v_S(t) = 15\cos(1{,}500t)$ V.

Analysis: Follow the steps of the Focus on Problem Solving box "AC Circuit Analysis."

Step 1: $v_S(t) = 15\cos(1{,}500t)\,\text{V} \qquad \omega = 1{,}500$ rad/s

Step 2: $\mathbf{V}_S = 15\angle 0\ \text{V}$

Step 3: $\mathbf{Z}_{R_1} = R_1 \qquad \mathbf{Z}_{R_2} = R_2 \qquad \mathbf{Z}_C = \dfrac{1}{j\omega C} \qquad \mathbf{Z}_L = j\omega L$

The resulting phasor circuit is shown in Figure 3.56(b).

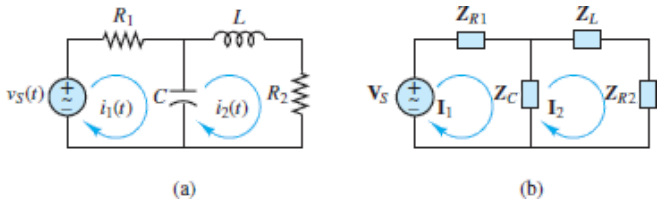


Figure 3.56 (a) Circuit for Example 3.20; (b) same circuit ready for phasor analysis

Step 4: The mesh equations are:

$$\begin{aligned} \mathbf{V}_S - \mathbf{Z}_{R_1}\mathbf{I}_1 - \mathbf{Z}_C[\mathbf{I}_1 - \mathbf{I}_2] = 0 \qquad & \text{mesh 1} \\ -\mathbf{Z}_C[\mathbf{I}_2 - \mathbf{I}_1] - \mathbf{Z}_L\mathbf{I}_2 - \mathbf{Z}_{R_2}\mathbf{I}_2 = 0 \qquad & \text{mesh 2} \end{aligned}$$

Multiply both sides of the mesh 2 equation by −1 and express the equations in matrix form.

$$\begin{bmatrix} \mathbf{Z}_{R_1} + \mathbf{Z}_C & -\mathbf{Z}_C \\ -\mathbf{Z}_C & \mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2} \end{bmatrix} \begin{bmatrix} \mathbf{I}_1 \\ \mathbf{I}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{V}_S \\ 0 \end{bmatrix}$$

Cramer's rule can be used to solve for the two currents:

$$\mathbf{I}_1 = \frac{\begin{vmatrix} \mathbf{V}_S & -\mathbf{Z}_C \\ 0 & \mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2} \end{vmatrix}}{\begin{vmatrix} \mathbf{Z}_{R_1} + \mathbf{Z}_C & -\mathbf{Z}_C \\ -\mathbf{Z}_C & \mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2} \end{vmatrix}} = \frac{\mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2}}{(\mathbf{Z}_{R_1} + \mathbf{Z}_C)(\mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2}) - \mathbf{Z}_C^2}\mathbf{V}_S$$

$$\mathbf{I}_2 = \frac{\begin{vmatrix} \mathbf{Z}_{R_1} + \mathbf{Z}_C & \mathbf{V}_S(j\omega) \\ -\mathbf{Z}_C & 0 \end{vmatrix}}{\begin{vmatrix} \mathbf{Z}_{R_1} + \mathbf{Z}_C & -\mathbf{Z}_C \\ -\mathbf{Z}_C & \mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2} \end{vmatrix}} = \frac{\mathbf{Z}_C}{(\mathbf{Z}_{R_1} + \mathbf{Z}_C)(\mathbf{Z}_C + \mathbf{Z}_L + \mathbf{Z}_{R_2}) - \mathbf{Z}_C^2}\mathbf{V}_S$$

Step 5: Substitute the impedance values in the previous expressions to find:

$$\begin{aligned} \mathbf{I}_1 &= \frac{1/j\omega C + j\omega L + R_2}{(R_1 + 1/j\omega C)(1/j\omega C + j\omega L + R_2) - (1/j\omega C)^2}\mathbf{V}_S \\ &= \frac{j\omega C + (j\omega C)^2(j\omega L) + (j\omega C)^2 R_2}{(j\omega C R_1 + 1)[1 + (j\omega C)(j\omega L) + j\omega C R_2] - 1}\mathbf{V}_S \\ \mathbf{I}_2 &= \frac{1/j\omega C}{(R_1 + 1/j\omega C)(1/j\omega C + j\omega L + R_2) - (1/j\omega C)^2}\mathbf{V}_S \\ &= \frac{j\omega C}{(j\omega C R_1 + 1)[1 + (j\omega C)(j\omega L) + j\omega C R_2] - 1}\mathbf{V}_S \end{aligned}$$

Plug in numerical values to obtain:

$$\mathbf{I}_1 \approx 0.0033\angle 0.92 = 0.0033\angle 53°\text{A}$$
$$\mathbf{I}_2 \approx 0.020\angle -1.5 = 0.020\angle -85°\text{A}$$

Step 6: Finally, express the resulting phasor currents in time-domain form:

$$i_1(t) \approx 3.3\cos(1{,}500t + 0.92) = 3.3\cos(1{,}500t + 53°)\text{ mA}$$
$$i_2(t) \approx 20.0\cos(1{,}500t - 1.5) = 20.0\cos(1{,}500t - 85°)\text{ mA}$$

Page 214

Comments: Note that the derivation of the symbolic equations for a circuit in phasor-impedance form using matrix techniques is no more involved than it would be for a resistive circuit. The only difference is the complex algebra manipulations. Complex matrix equations can be solved numerically using MatLab.

CHECK YOUR UNDERSTANDING

Compute the magnitude of the current $\mathbf{I}_S$ of Example 3.15 if $A = 10$ and $\phi = 0$, for $\omega = 10, 10^2, 10^3, 10^4$, and 10^5 rad/s. Are these results intuitively satisfying?

Answer: $|\mathbf{I}_S| = 0.041$ A; 0.083 A; 0.194 A; 0.2 A; 0.2 A. As the frequency increases, the impedance of the capacitor decreases. In the limit $\omega \to \infty$ the capacitor acts as a short-circuit. Thus, at sufficiently high frequency, $|\mathbf{I}_S| \approx |\mathbf{V}_S| / R_1 = 0.2$ A.

CHECK YOUR UNDERSTANDING

In the circuit of Example 3.16, assume that the branch R_2 in series with L_2 is the load. Find the Norton equivalent of the network seen by the load.

Answer: Short-circuit current is $310e^{j0}$ A. Norton equivalent impedance is $\approx 0.5e^{j0.013}\ \Omega$.

CHECK YOUR UNDERSTANDING

In Example 3.19, determine the value of the capacitor and inductor impedance at the two given frequencies. Compare these values to R_S. Do the results confirm what is stated in the Comments section?

Answer: At $\omega = 2\pi \times 10^3$, $\mathbf{Z}_L = j62.832\ \Omega$, $Z_C = -j1.5915 \times 10^3\ \Omega$. At $\omega = 2\pi \times 10^6$, $\mathbf{Z}_L = j6.2832 \times 10^4\ \Omega$, $\mathbf{Z}_C = -j1.5915\ \Omega$.

CHECK YOUR UNDERSTANDING

In Example 3.20, find the equivalent impedance "seen" by the independent voltage source to calculate its current. Show that the result is equal to the mesh current $i_1(t)$.

3.7 INSTANTANEOUS AND AVERAGE POWER

The fundamental concepts of single-phase AC power are average power, power factor, and the power triangle, which are extensions of concepts developed earlier in this chapter. An understanding of single-phase AC power is essential to the study of more complex topics, such as three-phase power, electric power generation and distribution, and electric machines.

When a linear electric circuit is excited by a sinusoidal source, all voltages and currents in the circuit are also sinusoids of the same frequency as the source. Page 215Figure 3.57 depicts a simple linear AC circuit. The most general expressions for the voltage and current delivered to an arbitrary load are as follows:

$$\begin{aligned} v(t) &= V\cos(\omega t + \theta_V) \\ i(t) &= I\cos(\omega t + \theta_I) \end{aligned} \tag{3.68}$$

where V and I are the peak amplitudes of the sinusoidal voltage and current, respectively, and θ_V and θ_I are their phase angles. Two such waveforms are plotted in Figure 3.58, with unit amplitude, angular frequency 150 rad/s, and phase angles $\theta_V = 0$ and $\theta_I = \pi/3$. Notice that a positive phase shift moves the time waveform to the left such that the current *leads* the voltage; or equivalently, the voltage *lags* the current. Keep in mind that all phase angles are relative to some reference. The reference

phase angle is freely chosen and therefore usually set to zero for simplicity. Also keep in mind that a phase angle represents a *time delay* of one sinusoid relative to its reference sinusoid.

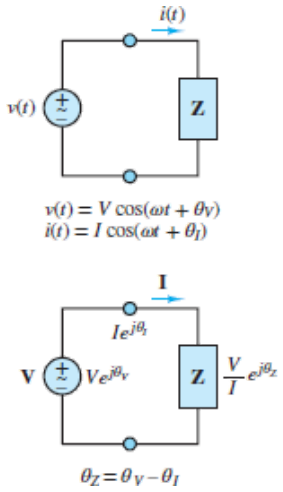


Figure 3.57 Time and frequency domain representations of an AC circuit. The phase angle of the load is $\theta_Z = \theta_V - \theta_I$.

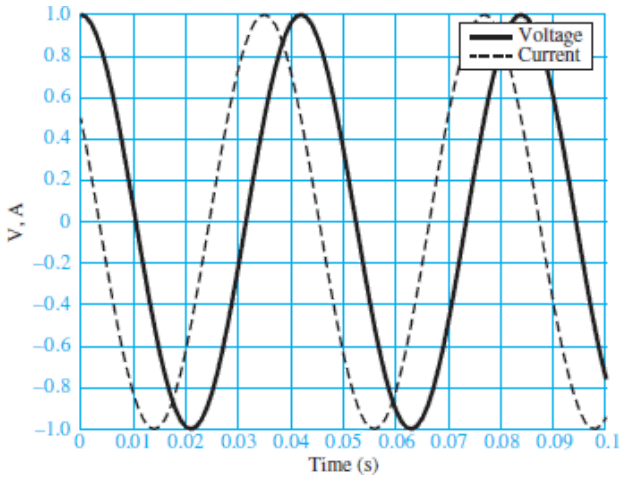


Figure 3.58 Current and voltage waveforms with unit amplitude and a phase shift of 60°

The **instantaneous power** dissipated by any element is the product of its instantaneous voltage and current.

$$p(t) = v(t)i(t) = VI\cos(\omega t + \theta_V)\cos(\omega t + \theta_I) \tag{3.69}$$

This expression is further simplified with the aid of the trigonometric identity:

$$2\cos(x)\cos(y) = \cos(x+y) + \cos(x-y) \tag{3.70}$$

Let $x = \omega t + \theta_V$ and $y = \omega t + \theta_I$ to yield:

$$\begin{aligned} p(t) &= \frac{VI}{2}[\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_V - \theta_I)] \\ &= \frac{VI}{2}[\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_Z)] \end{aligned} \tag{3.71}$$

The time average of the instantaneous power is defined by:

$$P = P_{\text{avg}} \equiv \frac{1}{T}\int_{t_0}^{t_0+T} p(t)dt \tag{3.72}$$

Page 216

where T is one period of $p(t)$. Use equation 3.71 to substitute for $p(t)$ to yield:

$$\begin{aligned} P = P_{\text{avg}} &= \frac{1}{T}\int_{t_0}^{t_0+T} \frac{VI}{2}[\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_Z)]\,dt \\ &= \frac{VI}{2T}\int_{t_0}^{t_0+T} [\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_Z)]\,dt \end{aligned} \tag{3.73}$$

The integral of the first part $\cos(2\omega t + \theta_V + \theta_I)$ is zero while the integral of the second part (a constant) is $T\cos(\theta_Z)$. Thus, the time averaged power P_{avg} is:

$$P = P_{\text{avg}} = \frac{VI}{2}\cos(\theta_Z) = \frac{1}{2}\frac{V^2}{|\mathbf{Z}|}\cos(\theta_Z) = \frac{1}{2}I^2|\mathbf{Z}|\cos(\theta_Z) \tag{3.74}$$

where

$$|\mathbf{Z}| = \frac{|\mathbf{V}|}{|\mathbf{I}|} = \frac{V}{I} \qquad \text{and} \qquad \theta_Z = \theta_V - \theta_I \tag{3.75}$$

Figure 3.59 shows the instantaneous and average power corresponding to the voltage and current signals of Figure 3.58.

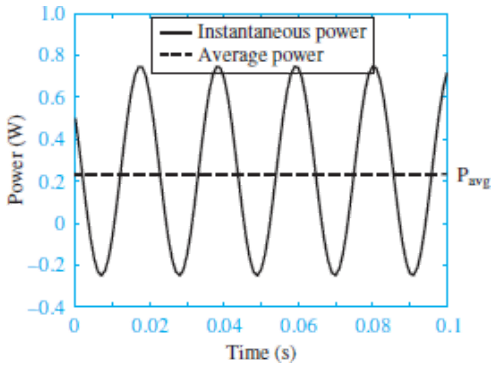


Figure 3.59 Instantaneous and average power corresponding to the signals in Figure 3.58

Effective Values

In North America, AC power systems operate at a fixed frequency of 60 cycles per second, or hertz (Hz), which corresponds to an angular (radian) frequency ω given by:

$$\omega = 2\pi \cdot 60 = 377 \text{ rad/s} \qquad \text{AC power frequency} \tag{3.76}$$

In Europe and many other parts of the world, the AC power frequency is 50 Hz.

It is customary in AC power analysis to employ the effective or root-mean-square (rms) amplitude rather than the peak amplitude for AC voltages and currents. In the case of a sinusoidal waveform, the effective voltage $\tilde{V} \equiv V_{rms}$ is related to the peak voltage V by:

$$\tilde{V} = V_{rms} = \frac{V}{\sqrt{2}} \tag{3.77}$$

Page 217

Likewise, the effective current $\tilde{I} \equiv I_{rms}$ is related to the peak current I by:

$$\tilde{I} = I_{rms} = \frac{I}{\sqrt{2}} \tag{3.78}$$

The rms, or effective, value of an AC source is the DC value that produces the same average power dissipated by a resistive load.

The average power can be expressed in terms of effective voltage and current by plugging $v = \sqrt{2}\tilde{v}$ and $i = \sqrt{2}\tilde{i}$ into equation 3.74 to find:

$$P = P_{\text{avg}} = \tilde{V}\tilde{I}\cos(\theta_Z) = \frac{\tilde{V}^2}{|\mathbf{Z}|}\cos(\theta_Z) = \tilde{I}^2|\mathbf{Z}|\cos(\theta_Z) \tag{3.79}$$

Voltage and current phasors are also represented with effective amplitudes by the notation:

$$\tilde{\mathbf{V}} = \tilde{V}e^{j\theta_V} = \tilde{V}\angle\theta_V \tag{3.80}$$

and

$$\tilde{\mathbf{I}} = \tilde{I}e^{j\theta_I} = \tilde{I}\angle\theta_I \tag{3.81}$$

It is critical to pay close attention to the *mathematical notation,* namely that complex quantities, such as **V**, **I**, and **Z** are boldface. On the other hand, scalar quantities, such as v, i, $\tilde{V}$, and $\tilde{I}$ are italicized. The relationship between these quantities is $V = |\mathbf{V}|$ and $\tilde{V} = |\tilde{\mathbf{V}}|$.

Power Factor

The phase angle θ_Z of the load impedance plays a very important role in AC power circuits. As shown in equation 3.79, the average power dissipated by an AC load is proportional to $\cos(\theta_Z)$. For this reason, $\cos(\theta_Z)$ is known as the **power factor (pf)**. For purely resistive loads:

$$\theta_Z = 0 \quad \rightarrow \quad \text{pf} = 1 \quad \text{Resistive load} \tag{3.82}$$

For purely inductive or capacitive loads:

$$\theta_Z = +\pi/2 \quad \rightarrow \quad \text{pf} = 0 \quad \text{Inductive load} \tag{3.83}$$

$$\theta_Z = -\pi/2 \quad \rightarrow \quad \text{pf} = 0 \quad \text{Capacitive load} \tag{3.84}$$

For loads with nonzero resistive (real) and reactive (imaginary) parts:

$$0 < |\theta_Z| < \pi/2 \quad \rightarrow \quad 0 < \text{pf} < 1 \quad \text{Complex load} \tag{3.85}$$

Using the definition $\text{pf} \equiv \cos(\theta_Z)$ the average power can be expressed as:

$$P = P_{\text{avg}} = \tilde{V}\tilde{I}\,\text{pf} \tag{3.86}$$

Thus, the average power dissipated by an ideal resistor is

$$(P_{\text{avg}})_R = \tilde{V}_R\tilde{I}_R\,\text{pf}_R = \tilde{V}_R\tilde{I}_R \tag{3.87}$$

Page 218

because $\text{pf}_R = 1$. By contrast, the average power dissipated by an ideal capacitor or inductor is

$$(P_{\text{avg}})_X = \tilde{V}_X\tilde{I}_X\,\text{pf}_X = 0 \tag{3.88}$$

because $\text{pf}_X = 0$, where the subscript X indicates a reactive element (i.e., either a capacitor or inductor). It is important to note that although ideal capacitors and inductors are *lossless* (i.e., they store and release energy but do not dissipate energy), they do influence power dissipation in a circuit by affecting the voltage across and the current through resistors in the circuit.

When θ_Z is positive, the load is *inductive* and the power factor is said to be *lagging*; when θ_Z is negative, the load is *capacitive* and the power factor is said to be *leading*. It is important to keep in mind that $\text{pf} = \cos(\theta_Z) = \cos(-\theta_Z)$ because the cosine is an even function. Thus, while it may be important to know whether a load is inductive or capacitive, the value of the power factor only indicates the extent to which a load is inductive or capacitive. To know whether a load is inductive or capacitive, one must know whether the power factor is leading or lagging.

EXAMPLE 3.21 Computing Average and Instantaneous AC Power

Problem

Compute the average and instantaneous power dissipated by the load of Figure 3.60.

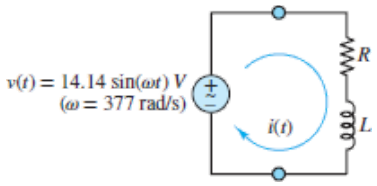


Figure 3.60

Solution

Known Quantities: Source voltage and frequency, load resistance and inductance values.

Find: P_{avg} and $p(t)$ for the *RL* load.

Schematics, Diagrams, Circuits, and Given Data: $v(t) = 14.1 \sin(377t)$ V; $R = 4\ \Omega$; $L = 8$ mH.

Assumptions: None.

Analysis: The source voltage is expressed in terms of $\sin(377t)$. By convention, all time- domain sinusoids are expressed as cosines. To convert $\sin(377t)$ to $\cos(377t + \theta_V)$ recall that a sine equals a cosine shifted forward in time (to the right) by $\pi/2$ rad; that is, $\sin(377t) = \cos(377t - \pi/2)$. Thus, at the angular frequency $\omega = 377$ rad/s the source effective voltage is

$$\tilde{\mathbf{V}} \approx 10\angle\left(-\frac{\pi}{2}\right) \text{V rms}$$

where 14.1 V ≈ 10 V rms.

The equivalent impedance of the load is

$$\mathbf{Z} = R + j\omega L \approx 4 + j3 = 5\angle(37°) = 5\angle(0.65\,\text{rad})\,\Omega$$

The current in the loop is

$$\tilde{\mathbf{I}} = \frac{\tilde{\mathbf{V}}}{\mathbf{Z}} \approx \frac{10\angle(-\pi/2)}{5\angle(0.65)} \approx 2\angle(-2.2)\,\text{A rms}$$

It is instructive to compute the average power dissipated in the circuit in two ways:

1. The most straightforward approach is to compute:

$$P_{avg} = \tilde{V}\tilde{I}\cos(\theta_Z) = 10 \times 2 \times \cos(0.65) \approx 16\,\text{W}$$

2. Another approach is to realize that the average power dissipated by the inductor is zero. Thus, the total average power dissipated equals the average power dissipated by the resistor. Thus:

$$(P_{avg})_R = \tilde{I}^2 R \, \text{pf}_R = \tilde{I}^2 R \approx (2)^2 \times 4 = 16\,\text{W}$$

The instantaneous power supplied by the voltage source and dissipated by the complex load is given by:

$$p(t) = v(t) \times i(t) = \sqrt{2} \times 10 \sin(377t) \times \sqrt{2} \times 2 \cos(377t - 2.2)\,\text{W}$$

The instantaneous voltage and current waveforms and the instantaneous and average power are plotted in Figure 3.61.

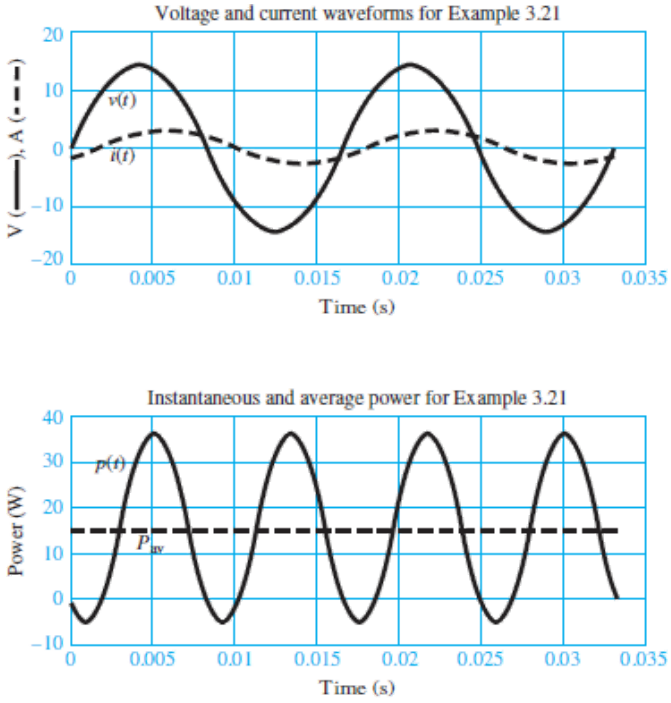


Figure 3.61

Comment: It is standard procedure in electrical engineering practice to use rms values in power calculations. Also, note that the instantaneous power can be negative at times even though the average power is positive. This result reflects the fact that although the average power of an inductor is identically zero, the instantaneous power of an inductor can be positive or negative as the inductor charges or discharges with the sinusoidal source.

EXAMPLE 3.22 Computing Average AC Power

Problem

Compute the average power dissipated by the load of Figure 3.62.

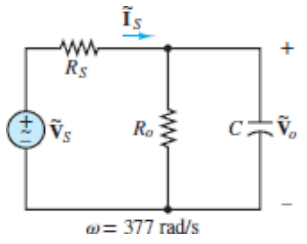


Figure 3.62

Solution

Known Quantities: Source voltage, internal resistance, load resistance, capacitance, and frequency.

Find: P_{avg} for the R_o // C load.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 110\angle 0°$ V rms; $R_S = 2\ \Omega$; $R_o = 16\ \Omega$; $C = 100\ \mu$F; $\omega = 377$ rad/s.

Assumptions: None.

Analysis: First, compute the impedance of the load at the angular frequency $\omega = 377$ rad/s:

$$\mathbf{Z}_o = R_o \| \frac{1}{j\omega C} = \frac{R_o}{1 + j\omega C R_o} = \frac{16}{1 + j0.603} = 13.7\angle(-0.543)\,\Omega$$

where the angle is given in radians. Next, apply voltage division to compute the load voltage:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{13.7\angle(-0.543)}{2 + 13.7\angle(-0.543)}110\angle 0 = 97.6\angle(-0.067)\text{ V rms}$$

Finally, compute the average power using equation 3.79:

$$P_{avg} = \frac{|\tilde{\mathbf{V}}_o|^2}{|\mathbf{Z}_o|}\cos(\theta_{Z_o}) = \frac{97.6^2}{13.7}\cos(-0.543) = 595 \text{ W}$$

Alternatively, compute the source current $\tilde{\mathbf{I}}_S$ and then use equation 3.79 to compute the average power:

$$\tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = 7.12\angle 0.476 \text{ A rms}$$

$$P_{avg} = |\tilde{\mathbf{I}}_S|^2 |\mathbf{Z}_o| \cos(\theta_{Z_o}) = 7.12^2 \times 13.7 \times \cos(-0.543) = 595 \text{ W}$$

EXAMPLE 3.23 Computing Average AC Power

Problem

Compute the average power dissipated by the load of Figure 3.63.

Solution

Known Quantities: Source voltage, internal resistance, load resistance, capacitance and inductance values, and frequency.

Find: P_{avg} for the complex load.

Page 221

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_s = 110\angle 0$ V rms; $R = 10\ \Omega$; $L = 0.05$ H; $C = 470\ \mu$F; $\omega = 377$ rad/s. Figure 3.63. Figure 3.63.

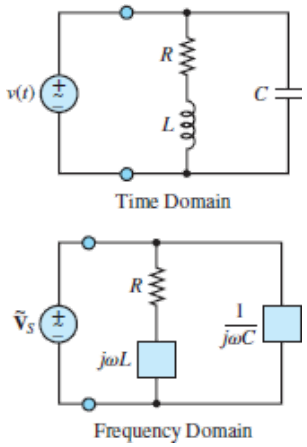


Figure 3.63

Assumptions: None.

Analysis: Compute the impedance of the load $\mathbf{Z}_o$ at the angular frequency ω = 377 rad/s:

$$\mathbf{Z}_o = (R + j\omega L) \| \frac{1}{j\omega C} = \frac{(R + j\omega L)/j\omega C}{R + j\omega L + 1/j\omega C}$$

$$= \frac{R + j\omega L}{1 - \omega^2 LC + j\omega CR} = 1.16 - j7.18$$

$$= 7.27\angle(-1.41)\ \Omega$$

Note that the equivalent load impedance at ω = 377 rad/s has a negative imaginary part, which is a feature of a *capacitive load*, as shown in Figure 3.64. The average power is

$$P_{avg} = \frac{|\tilde{\mathbf{V}}_s|^2}{|\mathbf{Z}_o|}\cos(\theta) = \frac{110^2}{7.27}\cos(-1.41) = 266\ \text{W}$$

Comment: At ω = 377 rad/s, the capacitance has a larger impact on the total equivalent impedance than the inductance. At lower frequencies, where the impedance of the capacitor is large compared to $R + j\omega L$, the parallel equivalent impedance will be inductive. It is instructive to determine the frequency when the parallel equivalent impedance has a zero imaginary part.

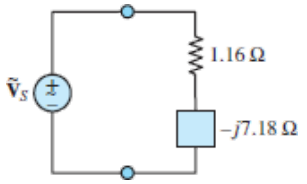


Figure 3.64

CHECK YOUR UNDERSTANDING

Consider the circuit shown in Figure 3.65. Find the impedance of the load "seen" by the voltage source and compute the average power dissipated by the load. The constant 155.6 multiplying the cosine function is the peak amplitude, not the rms amplitude.

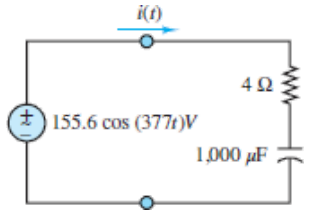


Figure 3.65

Answer: $\mathbf{Z} = 4.8e^{-j33.5^\circ}\Omega$; $P_{avg} = 2,103.4$ W

CHECK YOUR UNDERSTANDING

For Example 3.22, compute the average power dissipated by the internal source resistance R_S. Also compute the expression $P_{avg} = \tilde{I}_S^2 |\mathbf{Z}| \cos(\theta_Z)$.

Answer: 101.46 W; 595 W

3.8 APPARENT POWER AND THE POWER TRIANGLE

Average power was previously derived as

$$P = P_{avg} = \tilde{V}\tilde{I}\cos(\theta_V - \theta_I) = \tilde{V}\tilde{I}\cos(\theta_Z) = \tilde{V}\tilde{I}\ \text{pf} \tag{3.89}$$

Page 222

This notation can be simplified by defining the *apparent power* S and the *real power* $P = P_{avg}$ such that

$$S = \tilde{V}\tilde{I} \qquad \text{apparent power} \tag{3.90}$$

and

$$P = S\cos(\theta_Z) = S\ \text{pf} \qquad \text{real power} \tag{3.91}$$

The trigonometric relationship shown in Equation 3.91 suggests a right triangle with an interior angle θ_Z , a hypotenuse of length S and one leg of length P as shown in Figure 3.66. The third leg of the triangle is defined as the *reactive power* Q, which is due to the presence of capacitance and/or inductance.

$$Q = S\sin(\theta_Z) \qquad \text{reactive power} \tag{3.92}$$

Figure 3.66 shows a *power triangle*, which is a compact representation of the three elements of AC power and is a very useful visual aid in problem solving.

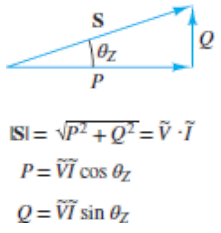


Figure 3.66 Power triangle

The apparent power can also be represented as the magnitude of a complex exponential $Se^{j\theta_Z}$ in the complex plane with a real part P and an imaginary part Q. This representation results in the concept of *complex power* **S** where

$$\mathbf{S} = Se^{j\theta_Z} = S\cos(\theta_Z) + jS\sin(\theta_Z) = P + jQ \qquad \text{complex power} \tag{3.93}$$

Recall from the generalized Ohm's law that $\theta_V = \theta_I + \theta_Z$ such that

$$\mathbf{S} = Se^{j\theta_Z} = \tilde{V}\tilde{I}\,e^{j(\theta_V - \theta_I)} = (\tilde{V}e^{j\theta_V})(\tilde{I}e^{-j\theta_I}) = \tilde{\mathbf{V}}\tilde{\mathbf{I}}^* \qquad \text{complex power} \tag{3.94}$$

where $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{I}}$ are the effective voltage and current phasors, respectively, and the asterisk indicates the complex conjugate.

The generalized Ohm's law can be used to represent the expressions for real and reactive power in Equations 3.91 and 3.92 in forms similar to those introduced in chapter 1 for the power dissipated by a resistor.

$$P = S\cos(\theta_Z) = \tilde{V}\tilde{I}\cos(\theta_Z) = \tilde{I}^2|\mathbf{Z}|\cos(\theta_Z) = \frac{\tilde{V}^2}{|\mathbf{Z}|}\cos(\theta_Z) \quad \text{real power} \tag{3.95}$$

Likewise

$$Q = S\sin(\theta_Z) = \tilde{V}\tilde{I}\sin(\theta_Z) = \tilde{I}^2|\mathbf{Z}|\sin(\theta_Z) = \frac{\tilde{V}^2}{|\mathbf{Z}|}\sin(\theta_Z) \quad \text{reactive power} \tag{3.96}$$

The practice of using a triangle to represent the elements of AC power can be extended to also represent the elements of impedance **Z** as an *impedance triangle* as shown in Figure 3.67.

$$\mathbf{Z} = |\mathbf{Z}|\cos(\theta_Z) + j|\mathbf{Z}|\sin\theta_Z = R + jX \tag{3.97}$$

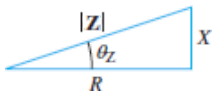


Figure 3.67 Impedance triangle

The resistance $R = |\mathbf{Z}|\cos(\theta_Z)$ and the reactance $X = |\mathbf{Z}|\sin(\theta_Z)$ were introduced earlier in this chapter.

It is important to realize that the impedance and power triangles are *similar*; that is, the two triangles have the same shape. This fact is often very helpful in problem solving.

The definitions of resistance and reactance can be used in the expressions for real and reactive power in Equations 3.95 and 3.96 to write:

$$P = \tilde{I}^2|\mathbf{Z}|\cos(\theta_Z) = \tilde{I}^2R \tag{3.98}$$

and

$$Q = \tilde{I}^2|\mathbf{Z}|\sin(\theta_Z) = \tilde{I}^2X \tag{3.99}$$

Table 3.6 summarizes these expressions for real and reactive power.

Table 3.6 Real and reactive power

Real power P	Reactive power Q
$\tilde{V}\tilde{I}\cos(\theta_Z)$	$\tilde{V}\tilde{I}\sin(\theta_Z)$
$\tilde{I}^2R$	$\tilde{I}^2X$

Similar expressions involving the effective voltage $\tilde{V}$ are found by substituting $\tilde{I} = \tilde{V}/|\mathbf{Z}|$ in Equations 3.98 and 3.99 to write:

$$P = \frac{\tilde{V}^2}{|\mathbf{Z}|^2}|\mathbf{Z}|\cos(\theta_Z) = \frac{\tilde{V}^2}{|\mathbf{Z}|^2}R \tag{3.100}$$

and

$$Q = \frac{\tilde{V}^2}{|\mathbf{Z}|^2}|\mathbf{Z}|\sin(\theta_Z) = \frac{\tilde{V}^2}{|\mathbf{Z}|^2}X \tag{3.101}$$

Finally, since $\tilde{V} = \tilde{I}|\mathbf{Z}|$ and $|\mathbf{Z}|^2 = \mathbf{Z}\mathbf{Z}^*$ the complex power can be expressed in a similar fashion.

$$\mathbf{S} = \tilde{I}^2\mathbf{Z} = \frac{\tilde{I}^2\mathbf{Z}\mathbf{Z}^*}{\mathbf{Z}^*} = \frac{\tilde{I}^2|\mathbf{Z}|^2}{\mathbf{Z}^*} = \frac{\tilde{V}^2}{\mathbf{Z}^*} \tag{3.102}$$

It is also important to keep in mind that capacitors and inductors (reactive loads) do not dissipate energy themselves; they are lossless elements. However, they do influence power dissipation in a circuit by affecting the voltage across and current through resistors, which do dissipate energy. It is worth noting that in purely resistive networks $Q = 0$, pf =1 and $P = S$, where P represents the real work done (per unit time) by a circuit. For example, P of an electric motor is the work done (per unit time) by the motor to perform a task. From the perspective of the utility company that provides the electric power for the motor and of the owner of the motor who has to pay the utility bill, it would be best if all of the apparent power S provided by the utility company was converted to real work. However, electric motors have inductance (e.g., coils of wire) such that $Q \neq 0$, pf < 1, and $P < S$. It is possible to *correct* the effect of a motor's inductance by adding capacitance in parallel with the motor so as to decrease Q and thereby decrease the apparent power S that must be provided for a given P required by the task.

FOCUS ON PROBLEM SOLVING

AC POWER COMPUTATION

1. Use AC circuit analysis methods to compute (if necessary) the voltage ac and current through the load. Convert peak amplitudes to effective (rms) valu

$$\tilde{\mathbf{V}} = \tilde{V}\angle\theta_V \qquad \text{and} \qquad \tilde{\mathbf{I}} = \tilde{I}\angle\theta_I$$

2. Compute $\theta_Z = \theta_V - \theta_I$ and the power factor pf = $\cos(\theta_Z)$. Draw the imped triangle, as shown in Figure 3.67.
3. Compute the apparent power $S = |\mathbf{S}| = \tilde{V}\tilde{I}$ such that $P = P_{\text{avg}} = S$ pf and Q $\sin(\theta_Z)$.
4. Draw the power triangle, as shown in Figure 3.66, and confirm that $S^2 = P^2$ and that $\tan(\theta_Z) = Q/P$.
5. If Q is negative, the load is capacitive and the power factor is *leading*; if positive, the load is inductive and the power factor is *lagging*.

Page 224

LO

EXAMPLE 3.24 AC Power Calculations

Problem

Compute the complex power for the load $\mathbf{Z}_o$ of Figure 3.68.

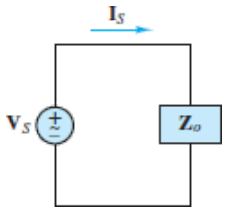


Figure 3.68

Solution

Known Quantities: Source, load voltage, and current.

Find: $\mathbf{S} = P + jQ$ for the complex load.

Schematics, Diagrams, Circuits, and Given Data: $v(t) = 100\cos(\omega t + 0.262)$ V; $i(t) = 2\cos(\omega t - 0.262)$ A; $\omega = 377$ rad/s.

Assumptions: All angles are given in units of radians unless indicated otherwise.

Analysis: The constants multiplying the cosine functions are always peak, not rms, values. These functions can be converted to phasors with rms amplitudes as follows:

$$\tilde{\mathbf{V}} = \frac{100}{\sqrt{2}}\angle 0.262\,\text{V} \qquad \tilde{\mathbf{I}} = \frac{2}{\sqrt{2}}\angle(-0.262)\,\text{A}$$

Compute the phase angle of the load, and the real and reactive power, using equations 3.95 and 3.96.

$$\theta_Z = \angle\tilde{\mathbf{V}} - \angle\tilde{\mathbf{I}} = 0.524 \text{ rad}$$
$$P = |\tilde{\mathbf{V}}||\tilde{\mathbf{I}}|\cos(\theta_Z) = \frac{200}{2}\cos(0.524) = 86.6\,\text{W}$$
$$Q = |\tilde{\mathbf{V}}||\tilde{\mathbf{I}}|\sin(\theta_Z) = \frac{200}{2}\sin(0.524) = 50\,\text{VAR}$$

Apply the definition of complex power (equation 3.94) to repeat the same calculation:

$$\mathbf{S} = \tilde{\mathbf{V}}\tilde{\mathbf{I}}^* = \frac{100}{\sqrt{2}}\angle 0.262 \times \frac{2}{\sqrt{2}}\angle -(-0.262) = 100\angle 0.524$$
$$= (86.6 + j50)\,\text{VA}$$

Therefore

$$P_{avg} = 86.6\text{ W} \qquad Q = 50\text{ VAR}$$

EXAMPLE 3.25 Real and Reactive Power Calculations

Problem

Compute the complex power for the load of Figure 3.69.

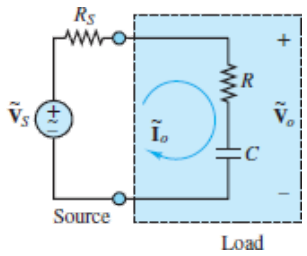


Figure 3.69

Solution

Known Quantities: Source voltage and resistance; load impedance.

Find: $\mathbf{S} = P + jQ$ for the complex load.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 110\angle 0°$ V; $R_S = 2\ \Omega$; $R = 5\ \Omega$; $C = 2{,}000\ \mu$F; $\omega = 377$ rad/s.

Page 225

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis: The load impedance is:

$$\mathbf{Z}_o = R + \frac{1}{j\omega C} = (5 - j1.326)\,\Omega = 5.173\angle(-0.259)\ \Omega$$

Apply voltage division and the generalized Ohm's law to compute the load voltage and current:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{5 - j1.326}{7 - j1.326} \times 110 = 79.86\angle(-0.072)\,\text{V}$$

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{79.86\angle(-0.072)}{5.173\angle(-0.259)} = 15.44\angle 0.187\,\text{A}$$

Finally, compute the complex power, as defined in equation 3.94:

$$\mathbf{S} = \tilde{\mathbf{V}}_o\tilde{\mathbf{I}}_o^* = 79.9\angle(-0.072) \times 15.44\angle(-0.187) = 1{,}233\angle(-0.259)$$
$$= (1{,}192 - j316)\text{VA}$$

Therefore:

$$P = 1{,}192\ \text{W} \qquad Q = -316\ \text{VAR}$$

Comment: Is the reactive power capacitive or inductive? Since $Q < 0$, the reactive power is capacitive!

EXAMPLE 3.26 Real Power Transfer for Complex Loads

Problem

Compute the complex power for the load between terminals *a* and *b* of Figure 3.70. Repeat the computation with the inductor removed from the load, and compare the real power for the two cases.

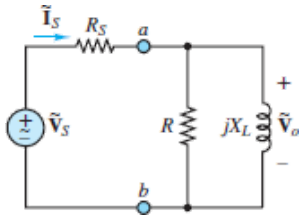


Figure 3.70

Solution

Known Quantities: Source voltage and resistance; load impedance.

Find:

1. $\mathbf{S}_1 = P_1 + jQ_1$ for the complex load.
2. $\mathbf{S}_2 = P_2 + jQ_2$ for the real load.
3. For each case, compute the ratio of the real power dissipated by the load to the overall real power dissipated by the circuit.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{V}_S = 110\angle 0°\,\text{V}$; $R_S = 4\ \Omega$; $R = 10\ \Omega$; $jX_L = j6\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. With the inductor included in the load, its impedance $\mathbf{Z}_o$ is:

$$\mathbf{Z}_o = R||j\omega L = \frac{10 \times j6}{10 + j6} = 5.145\angle 1.03\ \Omega$$

Page 226

Apply voltage division to compute the load voltage $\tilde{\mathbf{V}}_o$ and the generalized Ohm's law to compute the current $\tilde{\mathbf{I}}_o = \tilde{\mathbf{I}}_S$.

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{5.145\angle 1.03}{4 + 5.145\angle 1.03} \times 110 = 70.9\angle 0.444\,\text{V}$$

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{70.9\angle 0.444}{5.145\angle 1.03} = 13.8\angle(-0.586)\,\text{A}$$

Finally, compute the complex power, as defined in equation 3.94:

$$\mathbf{S}_1 = \tilde{\mathbf{V}}_o\tilde{\mathbf{I}}_o^* = 70.9\angle 0.444 \times 13.8\angle 0.586 = 978\angle 1.03\,\text{VA}$$
$$= (503 + j839)\,\text{VA}$$

Therefore:

$$P_1 = 503\,\text{W} \qquad Q_1 = 839\,\text{VAR}$$

2. With the inductor excluded from the load (Figure 3.71), its impedance is

$$\mathbf{Z}_o = R = 10\ \Omega$$

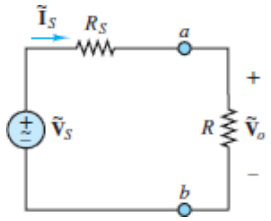


Figure 3.71

Compute the load voltage and current:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{10}{4 + 10} \times 110 = 78.57\angle 0\,\text{V}$$

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{78.57\angle 0}{10} = 7.857\angle 0\,\text{A}$$

Finally, compute the complex power, as defined in equation 3.94:

$$\mathbf{S}_2 = \tilde{\mathbf{V}}_o\tilde{\mathbf{I}}_o^* = 78.57\angle 0 \times 7.857\angle 0 = 617\angle 0 = (617 + j0)\,\text{VA}$$

Therefore:

$$P_2 = 617 \text{ W} \qquad Q_2 = 0 \text{ VAR}$$

3. To compute the overall real power P_{total} dissipated by the circuit, it is necessary to include the impact of the line resistance R_S and compute for each case:

$$\mathbf{S}_{\text{total}} = \tilde{\mathbf{V}}_S \tilde{\mathbf{I}}_S^* = P_{\text{total}} + jQ_{\text{total}}$$

For case 1:

$$\tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_{\text{total}}} = \frac{\tilde{\mathbf{V}}_S}{R_S + \mathbf{Z}_o} = \frac{110}{4 + 5.145\angle 1.03} = 13.8\angle(-0.586)\,\text{A}$$

$$\mathbf{S}_{1_{\text{total}}} = \tilde{\mathbf{V}}_S \tilde{\mathbf{I}}_S^* = 110 \times 13.8\angle(+0.586) = (1{,}264 + j838)\,\text{VA} = P_{1_{\text{total}}} + jQ_{1_{\text{total}}}$$

The percent real power transfer is

$$100 \times \frac{P_1}{P_{1_{\text{total}}}} = \frac{503}{1{,}264} = 39.8\%$$

For case 2:

$$\tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_{\text{total}}} = \frac{\tilde{\mathbf{V}}_S}{R_S + R} = \frac{110}{4 + 10} = 7.857\angle 0\,\text{A}$$

$$\mathbf{S}_{2_{\text{total}}} = \tilde{\mathbf{V}}_S \tilde{\mathbf{I}}_S^* = 110 \times 7.857 = (864 + j0)\,\text{VA} = P_{2_{\text{total}}} + jQ_{2_{\text{total}}}$$

Page 227

The percent real power transfer is

$$100 \times \frac{P_2}{P_{2_{\text{total}}}} = \frac{617}{864} = 71.4\%$$

Comments: If it were possible to eliminate the reactive part of the impedance, the percentage of real power transferred from the source to the load would be increased significantly. The procedure to accomplish this task is called *power factor correction*.

EXAMPLE 3.27 Complex Power and Power Triangle

Problem

Find the reactive and real power for the load of Figure 3.72. Draw the associated power triangle.

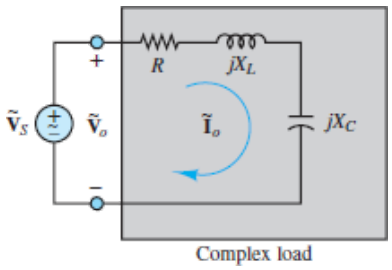


Figure 3.72

Solution

Known Quantities: Source voltage; load impedance.

Find: $\mathbf{S} = P + jQ$ for the complex load.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 60\angle 0\,\text{V}$; $R = 3\ \Omega$; $jX_L = j9\ \Omega$; $jX_C = -j5\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians.

Analysis: First, compute the load current:

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{60\angle 0}{3 + j9 - j5} = \frac{60\angle 0}{5\angle 0.9273} = 12\angle(-0.9273)\ \text{A}$$

Next, compute the complex power, as defined in equation 3.94:

$$\mathbf{S} = \tilde{\mathbf{V}}_o \tilde{\mathbf{I}}_o^* = 60\angle 0 \times 12\angle 0.9273 = 720\angle 0.9273 = (432 + j576)\ \text{VA}$$

Therefore:

$$P = 432\ \text{W} \qquad Q = 576\ \text{VAR}$$

The total reactive power must be the sum of the reactive powers in each of the elements, such that $Q = Q_C + Q_L$. Compute these two quantities as follows:

$$Q_C = |\bar{\mathbf{I}}_o|^2 \times X_C = (144)(-5) = -720 \text{ VAR}$$
$$Q_L = |\bar{\mathbf{I}}_o|^2 \times X_L = (144)(9) = 1{,}296 \text{ VAR}$$

and

$$Q = Q_L + Q_C = 576 \text{ VAR}$$

Comments: The power triangle corresponding to this circuit is drawn in Figure 3.73. The complex power **S** results from the vector addition of the three components P, Q_C, and Q_L.

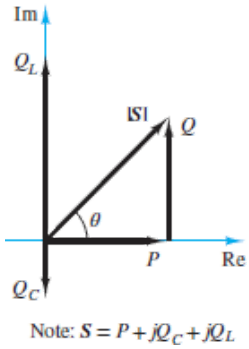


Figure 3.73

Page 228

CHECK YOUR UNDERSTANDING

Compute the real and reactive power for the load of Example 3.22.

Answer: $P_{avg} = 595$ W; $Q = -359$ VAR

CHECK YOUR UNDERSTANDING

Compute the real and reactive power for the load of Figure 3.65.

Answer: $P_{avg} = 2.1$ kW; $Q = 1.39$ kVAR

CHECK YOUR UNDERSTANDING

Refer to Example 3.26, and compute the percent of real power transfer for the case where the inductance of the load is one-half of the original value.

Answer: 29.3 %

CHECK YOUR UNDERSTANDING

Compute the power factor for the load of Example 3.27 with and without the inductor in the circuit.

Answer: pf = 0.6, lagging (with L in circuit); pf = 0.5145, leading (without L)

3.9 POWER FACTOR CORRECTION

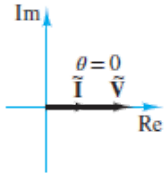
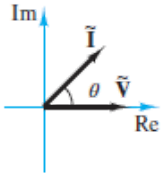
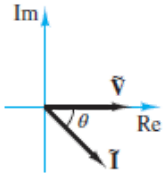
A power factor close to unity signifies an efficient transfer of energy from the AC source to the load while a small power factor corresponds to inefficient use of energy, as illustrated in Example 3.26. If a load requires a given real power P, the current required by the load will be minimized when the power factor is maximized, that is, when pf = $\cos(\theta_Z) \rightarrow 1$. When pf < 1, it is possible to increase it (i.e., *correct* it) by adding, as appropriate, reactance (e.g., capacitance) to the load. When pf is leading, inductance must be added; when pf is lagging, capacitance must be added.

If $\theta_Z > 0$, then $Q > 0$, the load is inductive, the load current *lags* the load voltage, and the power factor pf is lagging. Alternatively, if $\theta_Z < 0$, then $Q < 0$, the load is capacitive, the load current *leads* the load voltage, and the power factor pf is leading.

Table 3.7 illustrates and summarizes these concepts. For simplicity, the phase angle of the voltage phasor $\tilde{\mathbf{V}}$ shown in the table is zero and acts as a reference angle for the current phasor.

Page 229

Table 3.7 Important facts related to complex power

	Resistive load	**Capacitive load**	**Inductive load**
Ohm's law	$\tilde{\mathbf{V}} = \tilde{\mathbf{I}}\mathbf{Z}$	$\tilde{\mathbf{V}} = \tilde{\mathbf{I}}\mathbf{Z}$	$\tilde{\mathbf{V}} = \tilde{\mathbf{I}}\mathbf{Z}$
Complex impedance	$\mathbf{Z} = R$	$\mathbf{Z} = R + jX$ $X < 0$	$\mathbf{Z} = R + jX$ $X > 0$
Phase angle	$\theta = 0$	$\theta < 0$	$\theta > 0$
Complex plane sketch			
Explanation	The current is in phase with the voltage.	The current "leads" the voltage.	The current "lags" the voltage.
Power factor	Unity	Leading, < 1	Lagging, < 1
Reactive power	0	Negative	Positive

In practice, the load designed for a useful industrial task is often inductive because of the presence of electric motors. The power factor of an inductive load can be *corrected* by adding capacitance in parallel with the load. This procedure is called *power factor correction*.

The measurement and correction of the power factor for the load are an extremely important aspect of any industrial engineering application that requires the use of substantial quantities of electric power. In particular, industrial plants, construction sites, heavy machinery, and other heavy users of electric power must be

aware of the power factor that their loads present to the electric utility company. As was already observed, a low power factor results in greater current draw from the electric utility and greater line losses. Thus, computations related to the power factor of complex loads are of great utility to any practicing engineer.

FOCUS ON PROBLEM SOLVING

POWER FACTOR CORRECTION

1. Follow the steps outlined in the Focus on Problem Solving box "AC Po Computation" to find the initial phase angle of the load θZi, power factor real power P_i, and reactive power Q_i. An initial power triangle is helpful visualizing this information.
2. Assume an inductive load attached to a voltage source, such as that show Figures 3.68 and 3.74. Augment the load with a parallel capacitor such that:

$$\Delta Q = Q_C = \frac{\tilde{\mathbf{V}}^2}{|\mathbf{Z}_C|}\sin(\theta_Z) = -\omega C\tilde{\mathbf{V}}^2$$

3. Express the final reactive power Q_f as:

$$Q_f = Q_i + \Delta Q$$

Page 230

4. Since the voltage across the original load is unchanged by the addition of capacitor in parallel, the total real power of the augmented load is unchanged. Thus, $P_f = P_i$ and the final (corrected) phase angle of the augme load is:

$$\theta_{Z_f} = \tan^{-1}\left(\frac{Q_f}{P_f}\right)$$

   It is helpful to draw a final power triangle to visualize the effect of the par capacitor.
5. The final corrected power factor is:

$$\text{pf}_f = \cos(\theta_{Z_f})$$

Comments: Often, the objective is to calculate the capacitance needed to produ given power factor. In those cases, use the initial and final power factors to com the initial and final reactive powers. Use those results to calculate the chang reactive power ΔQ and then the capacitance C.

EXAMPLE 3.28 Power Factor Correction

Problem

Calculate the power factor for the load of Figure 3.74. Correct it to unity by adding a capacitor in parallel with the load.

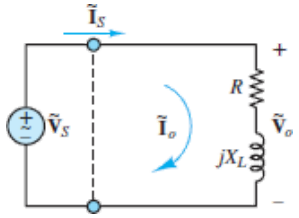


Figure 3.74

Solution

Known Quantities: Source voltage; load impedance.

Find:

1. $\mathbf{S} = P + jQ$ for the complex load.
2. Value of parallel capacitance that results in pf = 1.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 117\angle 0\,\text{V}$ rms; $R = 50\ \Omega$; $jX_L = j86.7\ \Omega$; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. First, compute the load impedance:

$$\mathbf{Z}_o = R + jX_L = 50 + j86.7 = 100\angle 1.05\ \Omega$$

Next, compute the load current $\tilde{\mathbf{I}}_o = \tilde{\mathbf{I}}_S$:

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{117\angle 0}{50 + j86.7} = \frac{117\angle 0}{100\angle 1.05} = 1.17\angle(-1.05)\ \text{A}$$

The complex power, as defined in equation 3.94, is

$$\mathbf{S} = \tilde{\mathbf{V}}_o \tilde{\mathbf{I}}_o^* = 117\angle 0 \times 1.17\angle 1.05 = 137\angle 1.05 = (68.4 + j118.5)\ \text{VA}$$

Page 231

Therefore:

$$P = 68.4\ \text{W} \qquad Q = 118.5\ \text{VAR}$$

The power triangle corresponding to this circuit is drawn in Figure 3.75. The vector diagram shows how the complex power **S** results from the vector addition of the two components P and Q.

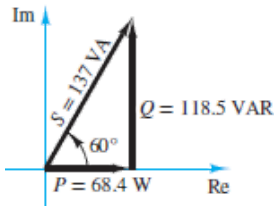


Figure 3.75

2. To correct the power factor to unity it is necessary to reduce the reactive power Q by adding in parallel a capacitor with $Q_C = -118.5$ VAR as shown in Figure 3.76. The required capacitance is found by:

$$X_C = \frac{|\tilde{\mathbf{V}}_o|^2}{Q_C} = -\frac{(117)^2}{118.5} = -115\ \Omega$$

The reactance X_C is related to the capacitance by:

$$jX_C = \frac{1}{j\omega C} = -\frac{j}{\omega C}$$

Thus, the result is

$$C = -\frac{1}{\omega X_C} = -\frac{1}{377(-115)} = 23.1\,\mu\text{F}$$

3. The total current required of the source is $\tilde{\mathbf{I}}_S = \tilde{\mathbf{I}}_o + \tilde{\mathbf{I}}_c$, where:

$$\tilde{\mathbf{I}}_c = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_c} = (j\omega C)(117\angle 0) = (377)(23.1\,\mu\text{F})(117)\angle(\pi/2) \approx 1.02\angle 90^\circ\ \text{A}$$

Notice that $|\tilde{\mathbf{I}}_c| = |\tilde{\mathbf{V}}_S|/|X_c| \approx 117/115 \approx 1.02$ A. The total current is computed by phasor addition to be:

$$\tilde{\mathbf{I}}_S \approx 1.17\angle(-1.05) + 1.02\angle(\pi/2) \approx 0.585\angle 0\ \text{A}$$

The corrected power factor pf = 1 implies that the impedance of the load is now purely real; that is, $\theta_Z = 0$. Thus, the source current must now be *in phase* with the source voltage; and it is.

Comments: Notice that the magnitude of the source current is reduced by increasing the power factor. The power factor correction, which is a very common procedure in electric power systems, is illustrated in Figure 3.76.

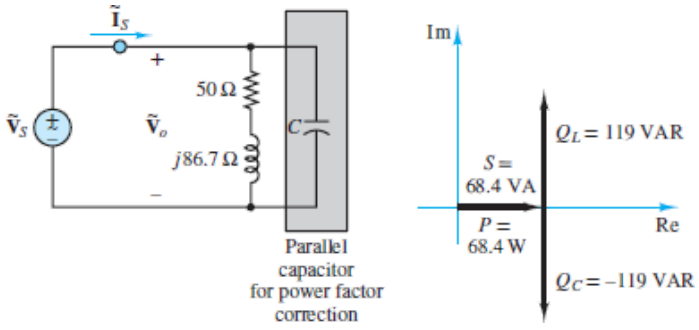


Figure 3.76 Power factor correction

EXAMPLE 3.29 Can a Series Capacitor Be Used for Power Factor Correction?

Problem

The circuit of Figure 3.77 suggests the use of a series capacitor for power factor correction. Why is this approach *not* a feasible alternative to the parallel capacitor approach demonstrated in Example 3.28?

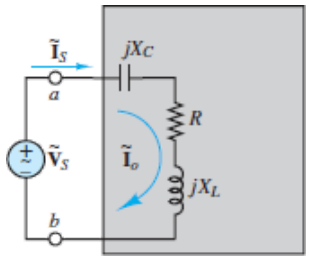


Figure 3.77

Solution

Known Quantities: Source voltage; load impedance.

Find: Load (source) current.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 117\angle 0$ V; $R = 50\ \Omega$; $jX_L = j86.7\ \Omega$; $jX_C = -j86.7\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis: First, compute the impedance of the load between terminals *a* and *b:*

$$\mathbf{Z}_o = R + jX_L + jX_C = 50 + j86.7 - j86.7 = 50\ \Omega$$

Notice that the reactance of the capacitor was chosen so as to make the total load purely resistive. Thus, $\theta_Z = 0$ and the corrected power factor is pf = 1. So far, so good.

Next, compute the current through the series load:

$$\tilde{\mathbf{I}}_o = \tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_o} = \frac{117\angle 0}{50} = 2.34\ \text{A}$$

The corrected power factor pf = 1 implies that the impedance of the load is now purely real; that is, $\theta_Z = 0$ and the source current is *in phase* with the source voltage.

The problem with this approach to power factor correction is revealed by computing the initial current through the load, prior to the addition of the capacitor.

$$(\tilde{\mathbf{I}}_o)_{\text{initial}} = \frac{\tilde{\mathbf{V}}_S}{R + jX_L} = \frac{117\angle 0}{50 + j86.7} \approx 1.17\angle(-\pi/3)\ \text{A}$$

Comments: Notice the increase in the source current as a result of the additional capacitor in series. Consequently, the power required of the source increased as well. In practice, adding capacitance in parallel can be accomplished relatively easily with one large bank located somewhere on an industrial site and away from the production motors themselves. Electric utilities motivate industries to raise power factors by offering discounted rates ($/kWh).

EXAMPLE 3.30 Power Factor Correction

Problem

A capacitor is used to correct the power factor of the 100 kW and lagging pf = 0.7 load of Figure 3.78. Determine the reactive power of the load alone, and compute the capacitance required for a corrected power factor pf = 1.
Page 233

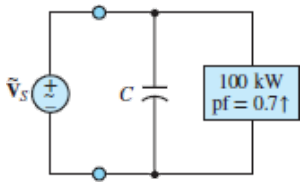


Figure 3.78

Solution

Known Quantities: Source voltage; load power and power factor.

Find:

1. The reactive power Q of the load alone.
2. The capacitance C required for a corrected power factor pf = 1.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 480\angle 0$ V rms; $P = 10^5$ W; pf = 0.7 lagging for the load; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. For the load alone, pf = 0.7 lagging or $\cos(\theta_Z) = 7/10$, and the power triangle has the *shape* shown in Figure 3.79. The real power is given as $P = 100$ kW, so the reactive power of the load can be computed using trigonometry.

$$Q = P\tan(\theta_Z) = (100\text{ kW})(\sqrt{51}/7) = 102\text{ kVAR}$$

   Since the power factor is lagging, the reactive power is positive as indicated in Table 3.7 and shown in the power triangle of Figure 3.80.

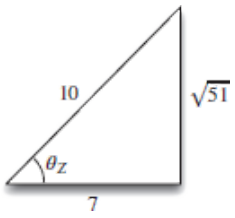


Figure 3.79 Relative dimensions of power triangle

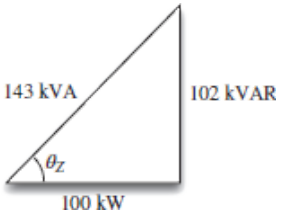


Figure 3.80 Power triangle

2. To set the corrected power factor to pf = 1 the capacitance must contribute –102 kVAR of reactive power.

$$Q_C = \Delta Q = Q_{\text{final}} - Q_{\text{initial}} = 0 - 102\text{ kVAR} = -102\text{ kVAR}$$

   Since the voltage across capacitor $\tilde{\mathbf{V}}_C$ equals the source voltage $\tilde{\mathbf{V}}_S$, the reactive power of the capacitor is

$$Q_C = \frac{|\tilde{\mathbf{V}}_C|^2}{|X_C|}\sin(-90^\circ) = -(\omega C)|\tilde{\mathbf{V}}_S|^2 = -(377)(480^2)C$$

   Thus, to correct the power factor to pf = 1 (zero total reactive power), the capacitor must satisfy:

$$Q_C = -(377)(480^2)C = -102\,\text{kVAR}$$

or

$$C = \frac{102\,\text{kVAR}}{(377)(480^2)} = 1{,}175\,\mu\text{F}$$

Use trigonometry and/or the Pythagorean theorem to show that the apparent power $|S| = 143$ kVA, as indicated in Figure 3.80.

Comments: Note that it is not necessary to know the load impedance explicitly to perform power factor correction; however, it is a useful exercise to compute the equivalent impedance seen by $\tilde{\mathbf{V}}_s$ and check that $\cos(\theta_Z) = 0.7$.

LO

EXAMPLE 3.31 Power Factor Correction

Problem

Figure 3.81 shows a second load added to the circuit of Figure 3.78. Determine the capacitance required for an overall corrected power factor pf = 1. Draw the phasor diagram showing the relationship between $\tilde{\mathbf{I}}_C$, $\tilde{\mathbf{I}}_1$, and $\tilde{\mathbf{I}}_2$.

Page 234

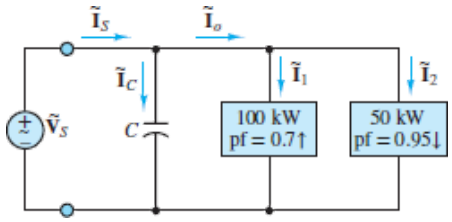


Figure 3.81

Solution

Known Quantities: Source voltage; load power and power factor.

Find:

1. The total reactive power of loads 1 and 2.
2. The capacitance C required for an overall power factor pf = 1.
3. $\tilde{\mathbf{I}}_C$, $\tilde{\mathbf{I}}_1$, and $\tilde{\mathbf{I}}_2$, and construct a phasor diagram of these currents.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 480\angle 0$ V rms; $P_1 = 100$ kW; $\text{pf}_1 = 0.7$ lagging; $P_2 = 50$ kW; $\text{pf}_2 = 0.95$ leading; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. Compute $\tilde{\mathbf{I}}_1$ and $\tilde{\mathbf{I}}_2$ using the relation $P = |\tilde{\mathbf{V}}||\tilde{\mathbf{I}}|\text{pf}$.

$$P_1 = |\tilde{\mathbf{V}}_S||\tilde{\mathbf{I}}_1|\cos(\theta_1) \quad \rightarrow \quad |\tilde{\mathbf{I}}_1| = \frac{P_1}{|\tilde{\mathbf{V}}_S|\cos(\theta_1)} \approx 298 \text{ A}$$

and

$$\angle\tilde{\mathbf{V}}_S = \angle\tilde{\mathbf{I}}_1 + \theta_{Z_1} \quad \rightarrow \quad \angle\tilde{\mathbf{I}}_1 = \angle\tilde{\mathbf{V}}_S - \theta_{Z_1} = 0 - \cos^{-1}(0.7) \approx -0.795 \text{ rad}$$

It is important to keep in mind that although inverse trigonometric functions are double-valued [e.g., $\cos^{-1}(0.7) \approx \pm 0.795$ rad], the power factor for load 1 is lagging such that $\theta_{Z_1} = +0.795$ rad is the correct choice.

Similarly, for load 2:

$$P_2 = |\tilde{\mathbf{V}}_S||\tilde{\mathbf{I}}_2|\cos(\theta_2) \quad \rightarrow \quad |\tilde{\mathbf{I}}_2| = \frac{P_2}{|\tilde{\mathbf{V}}_S|\cos(\theta_2)} \approx 110 \text{ A}$$

and

$$\angle\tilde{\mathbf{V}}_S = \angle\tilde{\mathbf{I}}_2 + \theta_{Z_2} \quad \rightarrow \quad \angle\tilde{\mathbf{I}}_1 = \angle\tilde{\mathbf{V}}_S - \theta_{Z_2} = 0 - \cos^{-1}(0.95) \approx +0.318 \text{ rad}$$

The power factor for load 2 is leading such that $\theta Z2 = -0.318$ rad is the correct choice.

Now use the given data and the relation $Q = P\tan(\theta_Z)$ to compute the reactive power for each load.

$$Q_1 = P_1 \tan(+0.795 \text{ rad}) \approx +102 \text{ kVAR}$$

and

$$Q_2 = P_2 \tan(-0.318 \text{ rad}) \approx -16.4 \text{ kVAR}$$

The power triangles for the two loads are shown in Figures 3.82 and 3.83. The total reactive power is therefore $Q = Q_1 + Q_2 \approx 85.6$ kVAR.

Page 235

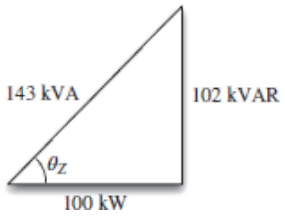


Figure 3.82 Power triangle for load 1

Figure 3.83 Power triangle for load 2

2. To set the corrected power factor to pf = 1 the capacitance must contribute –85.6 kVAR of reactive power.

$$Q_C = \Delta Q = Q_{\text{final}} - Q_{\text{initial}} = 0 - 85.6\,\text{kVAR} = -85.6\,\text{kVAR}$$

For a capacitor alone its reactive power is

$$Q_C = \frac{|\tilde{\mathbf{V}}_C|^2}{X_C} = -(\omega C)|\tilde{\mathbf{V}}_S|^2 = -(377)(480^2)C$$

Thus, to correct the power factor to pf = 1 (zero total reactive power), the capacitor must satisfy:

$$Q_C = -(377)(480^2)C = -85.6\,\text{kVAR}$$

or

$$C = \frac{85.6\,\text{kVAR}}{(377)(480^2)} \approx 985\,\mu\text{F}$$

3. To compute the capacitor current it is not possible to use $P = |\tilde{\mathbf{V}}||\tilde{\mathbf{I}}|\text{pf}$ because $P = 0$ and pf = 0 for a capacitor. Instead, the generalized Ohm's law provides an alternative approach.

$$\tilde{\mathbf{V}}_C = \tilde{\mathbf{I}}_C \mathbf{Z}_C \quad \rightarrow \quad |\tilde{\mathbf{I}}_C| = \frac{|\tilde{\mathbf{V}}_C|}{|\mathbf{Z}_C|} = \omega C |\tilde{\mathbf{V}}_C| \approx 178.2\,\text{A}$$

where $\tilde{\mathbf{V}}_C = \tilde{\mathbf{V}}_S$. The phase angle of $\tilde{\mathbf{I}}_C$ is

$$\angle\tilde{\mathbf{I}}_C = \angle\tilde{\mathbf{V}}_C - \theta_{Z_C} = 0 - (-\pi/2) = +\pi/2\,\text{rad}$$

The current phasor diagram can now be drawn as shown in Figure 3.84.

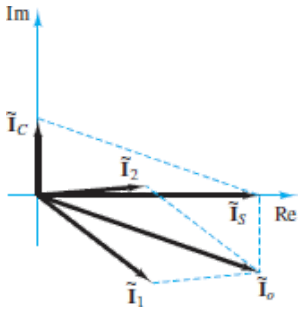


Figure 3.84

Comment: The power triangle suggests that the capacitor current can also be calculated using the relation $Q_C = |\tilde{\mathbf{V}}_C||\tilde{\mathbf{I}}_C| \sin(\theta_C)$, where $\theta_C = -\pi/2$ and $Q_C = |\tilde{\mathbf{V}}_C|^2/X_C = -(\omega C)|\tilde{\mathbf{V}}_C|^2$. Try it!

CHECK YOUR UNDERSTANDING

Compute the magnitude of $\tilde{\mathbf{I}}_s$ after the power factor correction in Example 3.28.

Answer: 0.585 A

CHECK YOUR UNDERSTANDING

Two cases of the voltage across and the current through a load are given below. Determine the power factor of the load, and whether it is leading or lagging, for each case.

a. $v(t) = 540\cos(\omega t + 15°)$ V, $i(t) = 2\cos(\omega t + 47°)$ A

b. $v(t) = 155\cos(\omega t - 15°)$ V, $i(t) = 2\cos(\omega t - 22°)$ A

Answer: a. 0.848, leading; b. 0.9925, lagging

Page 236

CHECK YOUR UNDERSTANDING

Determine if a load is capacitive or inductive, given the following facts:

a. pf = 0.87, leading
b. pf = 0.42, leading
c. $v(t) = 42\cos(\omega t)$ V, $i(t) = 4.2\sin(\omega t)$ A [*Hint:* $\sin(\omega t)$ lags $\cos(\omega t)$.]
d. $v(t) = 10.4\cos(\omega t - 22°)$ V, $i(t) = 0.4\cos(\omega t - 22°)$ A

Answer: a. Capacitive; b. capacitive; c. inductive; d. neither (resistive)

CHECK YOUR UNDERSTANDING

Compute the power factor for an inductive load with L = 100 mH in series with R = 0.4 Ω. Assume ω = 377 rad/s.

Answer: pf = 0.0106, lagging

FOCUS ON MEASUREMENTS

The Wattmeter

The instrument used to measure power is called a wattmeter. The external part of a wattmeter consists of four connections and a metering mechanism that displays the amount of real power dissipated by a circuit. The external and internal appearance of a wattmeter is depicted in Figure 3.85. Inside the wattmeter are two coils: a currentsensing coil and a voltage-sensing coil. In this example, assume for simplicity that the impedance of the current-sensing coil $\mathbf{Z}_I$ is zero and that the impedance of the voltagesensing coil $\mathbf{Z}_V$ is infinite. In practice, this will not necessarily be true; some correction mechanism will be required to account for the impedance of the sensing coils.

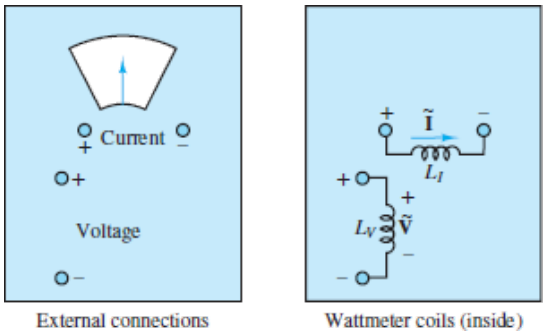


Figure 3.85

A wattmeter should be connected as shown in Figure 3.86 to provide both current and voltage measurements. The current-sensing coil is placed in series with the Page 237load and that the voltage-sensing coil is placed in parallel with the load. In this manner, the wattmeter is seeing the current through and the voltage across the load.

Remember that the power dissipated by a circuit element is related to these two quantities. The wattmeter, then, is constructed to provide a readout of the real power absorbed by the load: $P = \text{Re}(\mathbf{S}) = \text{Re}(\tilde{\mathbf{V}}\tilde{\mathbf{I}}^*)$.

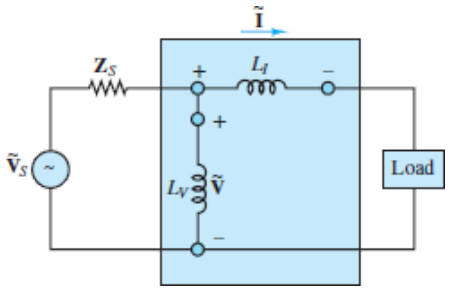


Figure 3.86

Problem

1. For the circuit shown in Figure 3.87, show the connections of a wattmeter between the ideal voltage source and the load and find the power dissipated by the load.
2. Show the connections that will determine the power dissipated by R_2. What should the meter read?

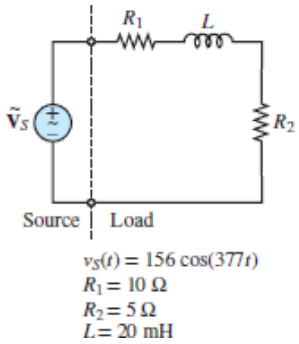


Figure 3.87

Solution

1. To measure the power dissipated by the load, the current through and the voltage across the entire load circuit must be measured. This means that the wattmeter must be connected as shown in Figure 3.88. The wattmeter should read

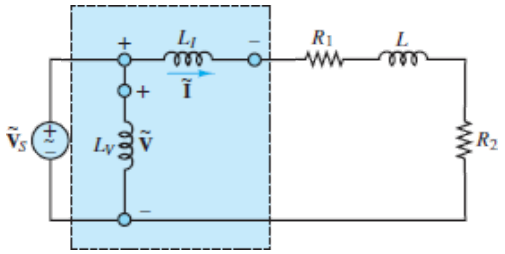


Figure 3.88

Page 238

$$P = \text{Re}\left(\tilde{\mathbf{V}}_S \tilde{\mathbf{I}}^*\right) = \text{Re}\left[\left(\frac{156}{\sqrt{2}}\angle 0\right)\left(\frac{(156/\sqrt{2})\angle 0}{R_1 + R_2 + j\omega L}\right)^*\right]$$

$$= \text{Re}\left[110\angle 0^\circ \left(\frac{110\angle 0}{15 + j7.54}\right)^*\right]$$

$$= \text{Re}\left[110\angle 0^\circ \left(\frac{110\angle 0}{16.79\angle 0.466}\right)^*\right] = \text{Re}\frac{110^2}{16.79\angle(-0.466)}$$

$$= \text{Re}(720.67\angle 0.466)$$

$$= 643.88\,\text{W}$$

2. To measure the power dissipated by R_2 alone, measure the current through R_2 and the voltage across R_2 *alone*. The connection is shown in Figure 3.89. The meter will read

$$P = |\tilde{\mathbf{I}}^2| R_2 = \left[\frac{110}{(15^2 + 7.54^2)^{1/2}}\right]^2 \times 5 = \frac{110^2}{15^2 + 7.54^2} \times 5$$

$$= 215\,\text{W}$$

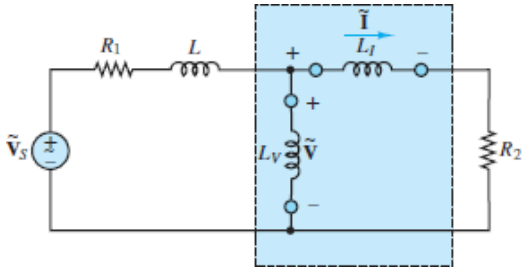


Figure 3.89

FOCUS ON MEASUREMENTS

Power Factor

Problem:

A capacitor is being used to correct the power factor of a load to unity, as shown in Figure 3.90. The capacitor value is varied, and measurements of the total current are taken. Explain how it is possible to zero in on the capacitance value necessary to bring the power factor to unity just by monitoring the current $\tilde{\mathbf{I}}_s$.

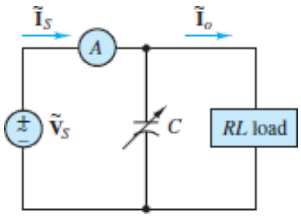


Figure 3.90

Page 239

Solution:

The current through the load is

$$\tilde{\mathbf{I}}_o = \frac{\tilde{V}_s \angle 0^\circ}{R + j\omega L} = \frac{\tilde{V}_s}{R^2 + \omega^2 L^2}(R - j\omega L)$$

$$= \frac{\tilde{V}_s R}{R^2 + \omega^2 L^2} - j\frac{\tilde{V}_s \omega L}{R^2 + \omega^2 L^2}$$

The current through the capacitor is

$$\bar{\mathbf{I}}_C = \frac{\tilde{V}_S \angle 0^\circ}{1/j\omega C} = j\tilde{V}_S \omega C$$

The source current to be measured is

$$\bar{\mathbf{I}}_S = \bar{\mathbf{I}}_o + \bar{\mathbf{I}}_C = \frac{\tilde{V}_S R}{R^2 + \omega^2 L^2} + j\left(\tilde{V}_S \omega C - \frac{\tilde{V}_S \omega L}{R^2 + \omega^2 L^2}\right)$$

The magnitude of the source current is

$$\tilde{I}_S = \sqrt{\left(\frac{\tilde{V}_S R}{R^2 + \omega^2 L^2}\right)^2 + \left(\tilde{V}_S \omega C - \frac{\tilde{V}_S \omega L}{R^2 + \omega^2 L^2}\right)^2}$$

When the load is a pure resistance, then the current and voltage are in phase, the power factor is 1, and all the power delivered by the source is dissipated by the load as real power. This corresponds to equating the imaginary part of the expression for the source current to zero or, equivalently, to the following expression:

$$\frac{\tilde{V}_S \omega L}{R^2 + \omega^2 L^2} = \tilde{V}_S \omega C$$

in the expression for $|\tilde{I}_S|$. Thus, the magnitude of the source current is actually a minimum when the power factor is unity! It is therefore possible to "tune" a load to a unity pf by observing the readout of the ammeter while changing the value of the capacitor and selecting the capacitor value that corresponds to the lowest source current value.

Conclusion

This chapter introduced concepts and tools useful in the analysis of AC circuits. The importance of AC circuit analysis cannot be overemphasized, for a number of reasons. First, circuits made up of resistors, inductors, and capacitors constitute reasonable models for more complex devices, such as transformers, electric motors and electronic amplifiers. Second, sinusoidal signals are ever-present in the analysis of many physical systems, not just circuits. Third, AC power is essential to all industrial activities and to the conveniences of residential life. Virtually all engineers will be exposed to AC power systems in their careers and the material presented in this chapter provides the tools needed to understand them. Upon completion of this chapter a student will have learned to:

1. *Compute currents, voltages and energy stored in capacitors and inductors*.

2. *Calculate the average and root-mean-square value of an arbitrary (periodic) signal.*

Page 240

3. *Convert time-domain sinusoidal voltages and currents to phasor notation, and vice versa, and represent circuits using impedances.*
4. *Apply the circuit analysis methods of Chapter 2 to AC circuits in phasor form.*
5. *Compute average AC power and the power factor of a complex load.*
6. *Compute apparent, real and reactive power for complex loads and draw a power triangle.*
7. *Compute the capacitance required to perform power factor correction on a load.*

HOMEWORK PROBLEMS

Section 3.2: Capacitors and Inductors

3.1 The current through a 0.8-H inductor is given by $i_L = \sin(100t + \frac{\pi}{4})$. Write the expression for the voltage across the inductor.

3.2 For each case shown below, derive the expression for the current through a 200-μF capacitor. $v_C(t)$ is the voltage across the capacitor.

a. $v_C(t) = 22\cos(20t - \frac{\pi}{3})$ V

b. $v_C(t) = -40\cos(90t + \frac{\pi}{2})$ V

c. $v_C(t) = 28\cos(15t + \frac{\pi}{8})$ V

d. $v_C(t) = 45\cos(120t + \frac{\pi}{4})$ V

3.3 Derive the expression for the voltage across a 200-mH inductor when its current is:

a. $i_L = -2\sin 10t$ A

b. $i_L = 2\cos 3t$ A

c. $i_L = -10\sin(50t - \frac{\pi}{4})$ A

d. $i_L = 7\cos(10t + \frac{\pi}{4})$ A

3.4 In the circuit shown in Figure P3.4, assume R = 1 Ω and L = 2 H. Also, let:

$$i(t) = \begin{cases} 0\,\text{A} & -\infty < t < 0 \\ t\,\text{A} & 0 \le t < 10\,\text{s} \\ 10\,\text{A} & 10\,\text{s} \le t < \infty \end{cases}$$

Find the energy stored in the inductor for all time.

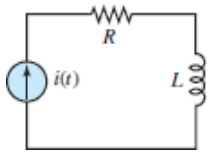


Figure P3.4

3.5 Refer to Problem 3.4 and find the energy delivered by the source for all time.

3.6 In the circuit shown in Figure P3.4, assume $R = 2\ \Omega$ and $L = 4$ H. Also, let:

$$i(t) = \begin{cases} 0\,\text{A} & -\infty < t < 0 \\ 10 + 2(t-5)\,\text{A} & 0 \le t < 5\,\text{s} \\ 2 - 2(t-9)\,\text{A} & 5 \le t < 9\,\text{s} \\ 2\,\text{A} & 9\,\text{s} \le t < \infty \end{cases}$$

Find:

a. The energy stored in the inductor for all time.

b. The energy delivered by the source for all time.

3.7 In the circuit shown in Figure P3.7, assume $R = 2\ \Omega$ and $C = 0.1$ F. Also, let:

$$v(t) = \begin{cases} 0\,\text{V} & \text{for } -\infty < t < 0 \\ t\,\text{V} & \text{for } 0 \le t < 10\,\text{s} \\ 10\,\text{V} & \text{for } 10\,\text{s} \le t < \infty \end{cases}$$

Find the energy stored in the capacitor for all time.

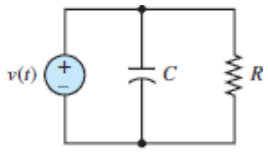


Figure P3.7

3.8 Refer to Problem 3.7 and find the energy delivered by the source for all time.

3.9 In the circuit shown in Figure P3.7, assume $R = 4\ \Omega$ and $C = 0.2$ F. Also, let:

$$v(t) = \begin{cases} 0\,\text{V} & -\infty < t < 0 \\ 4 + (t-4)\,\text{V} & 0 \le t < 4\,\text{s} \\ 1 - 0.5(t-10)\,\text{V} & 4 \le t < 10\,\text{s} \\ 1\,\text{V} & t > 10\,\text{s} \end{cases}$$

Find:

a. The energy stored in the capacitor for all time.

b. The energy delivered by the source for all time.

Page 241

3.10 The voltage waveform shown in Figure P3.10 is piecewise linear and continuous across a 20-mH inductor. Determine the current $i_L(t)$ through the inductor, assuming $i_L(0) = 50$ mA.

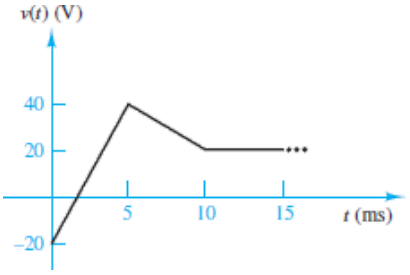


Figure P3.10

3.11 The voltage waveform shown in Figure P3.10 is piecewise linear and continuous across a 100-μF capacitor. Determine the current $i_C(t)$ through the capacitor. Explain the concept of current through a capacitor even when the space between the capacitor plates is a perfect insulator. How is current through a capacitor different from leakage current?

3.12 The voltage across a 0.5-mH inductor, plotted as a function of time, is shown in Figure P3.12. Determine the current through the inductor at $t = 6$ ms. Assume $i_L(0) = 0$A.

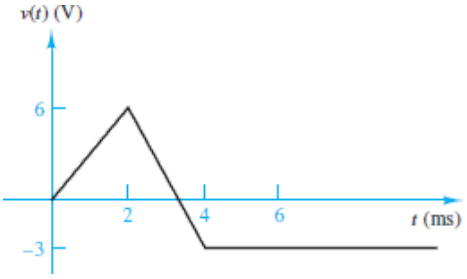


Figure P3.12

3.13 Figure P3.13 shows the voltage across a capacitor plotted as a function of time where:

$$v_{PK} = 20\text{ V} \qquad T = 40\ \mu\text{s} \qquad C = 680\text{ nF}$$

Determine and plot the waveform for the current through the capacitor as a function of time. How is the current affected by the discontinuities in slope in the voltage waveform?

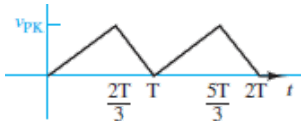


Figure P3.13

3.14 The current through a 16-μH inductor is zero at $t = 0$, and the voltage across the inductor (shown in Figure P3.14) is:

$$v(t) = \begin{cases} 0\,\text{V} & t \le 0 \\ 3t^2\,\text{V} & 0 \le t \le 20\,\mu\text{s} \\ 1.2\,\text{nV} & t \ge 20\,\mu\text{s} \end{cases}$$

Determine the current through the inductor at $t = 30\ \mu s$.

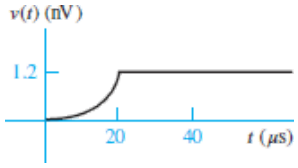


Figure P3.14

3.15 The voltage across a generic element *X* has the waveform shown in Figure P3.15. For $0 < t < 10$ ms, determine and plot the current through *X* when it is a:

a. 7-Ω resistor.

b. 0.5-μF capacitor.

c. 7-mH inductor. Assume $iL\ (0) = 0$A.

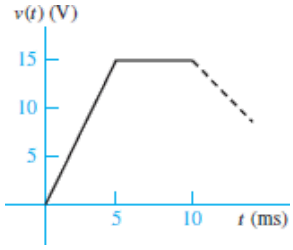


Figure P3.15

3.16 The plots shown in Figure P3.16 are the voltage across and the current through an ideal capacitor. Determine its capacitance.

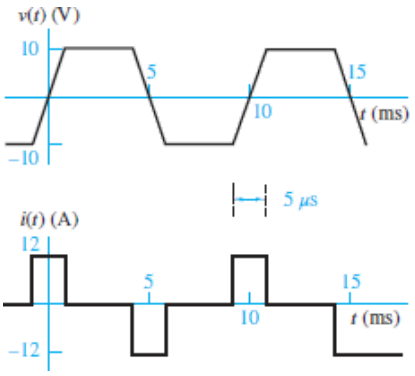


Figure P3.16

Page 242

3.17 The plots shown in Figure P3.17 are the voltage across and the current through an ideal inductor. Determine its inductance.

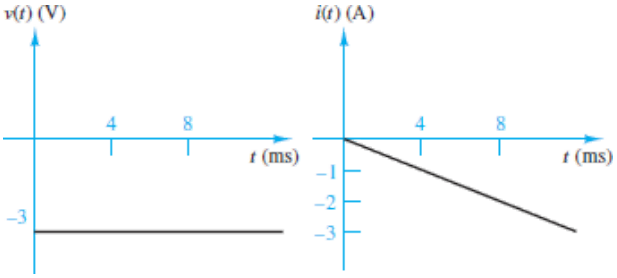


Figure P3.17

3.18 The plots shown in Figure P3.18 are the voltage across and the current through an ideal capacitor. Determine its capacitance.

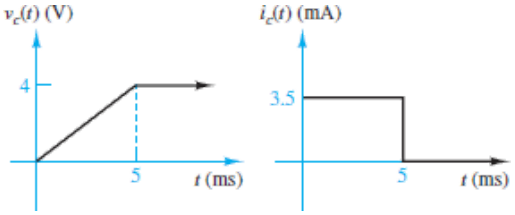


Figure P3.18

3.19 The plots shown in Figure P3.19 are the voltage across and the current through an ideal capacitor. Determine its capacitance.

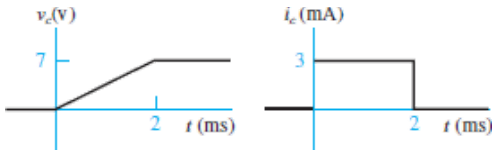


Figure P3.19

3.20 The voltage $v_L(t)$ across a 10-mH inductor is shown in Figure P3.20. Find the current $i_L(t)$ through the inductor. Assume $i_L(0) = 0$ A.

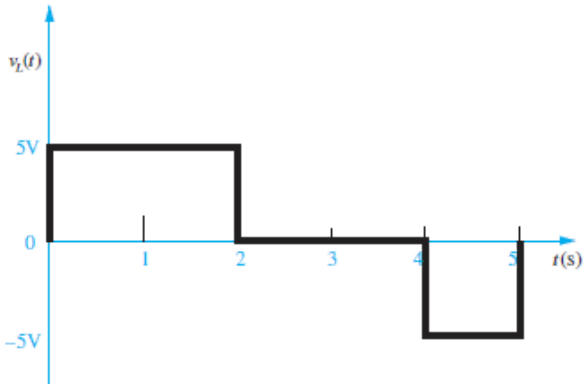


Figure P3.20

3.21 The current through a 2-H inductor is plotted in Figure P3.21. Plot the inductor voltage $v_L(t)$. (Assume $i_L(0) = 0$A.).

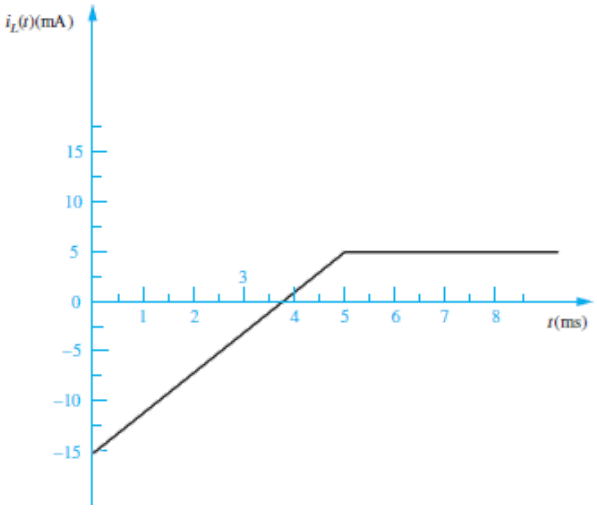


Figure P3.21

3.22 The voltage across a 100-mH inductor and a 500-μF capacitor is shown in Figure P3.22. Plot the inductor and capacitor currents, $i_L(t)$ and $i_C(t)$, for $0 < t < 6$ s, assuming $i_L(0) = 0$ A.

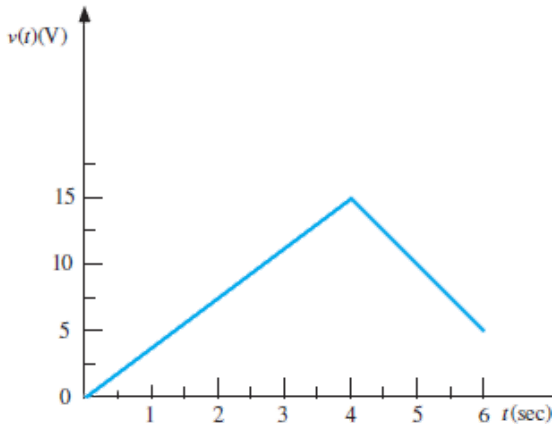


Figure P3.22

3.23 In the circuit shown in Figure P3.4, assume $R = 1\ \Omega$ and $L = 2$ H. Also, let:

$$i(t) = \begin{cases} 0\,\text{A} & -\infty < t < 0 \\ t\,\text{A} & 0 \le t < 1\,\text{s} \\ -(t-2)\,\text{A} & 0 \le t < 2\,\text{s} \\ 0\,\text{A} & 2\,\text{s} \le t < \infty \end{cases}$$

Find the energy stored in the inductor for all time.

Page 243

3.24 In the circuit shown in Figure P3.7, assume $R = 2\ \Omega$ and $C = 0.1$ F. Also, let:

$$v(t) = \begin{cases} 0\,\text{V} & -\infty < t \le 0 \\ 2t\,\text{V} & 0 \le t \le 1\,\text{s} \\ -2(t-2)\,\text{V} & 1 \le t \le 2\,\text{s} \\ 0\,\text{V} & 2\,\text{s} \le t < \infty \end{cases}$$

Find the energy stored in the capacitor for all time.

3.25 The voltage $v_C(t)$ across a capacitor is shown in Figure P3.25. Determine and sketch the current $i_C(t)$ through the capacitor.

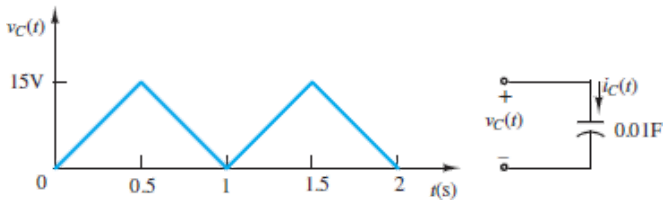


Figure P3.25

3.26 The voltage $v_L(t)$ across an inductor is shown in Figure P3.26. Determine and sketch the current $i_L(t)$ through the inductor. Assume $i_L(0) = 0$ A.

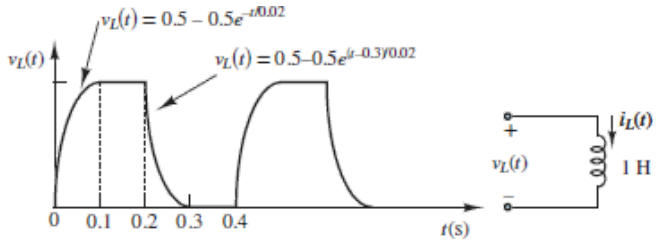


Figure P3.26

3.27 Assume dc steady-state conditions and find the energy stored in each capacitor and inductor shown in Figure P3.27.

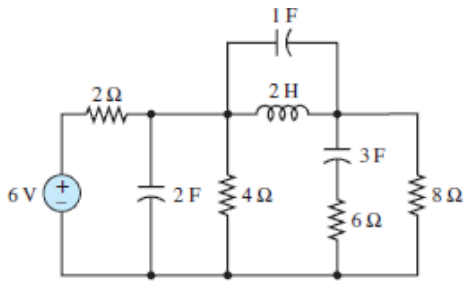


Figure P3.27

3.28 Assume dc steady-state conditions and find the energy stored in each capacitor and inductor shown in Figure P3.28.

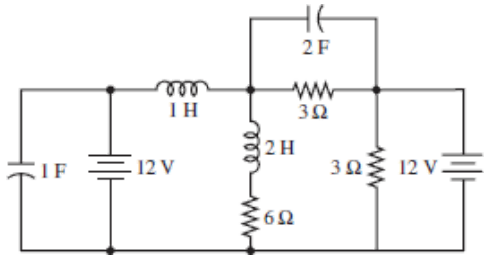


Figure P3.28

Section 3.3: Time-Dependent Waveforms

3.29 Find the average and rms values of $x(t)$ when:

$$x(t) = 3\cos(7\omega t) + 4$$

3.30 The output voltage waveform of a controlled rectifier is shown in Figure P3.30. The input voltage waveform was a sinusoid of amplitude 110 V rms. Find the average and rms voltages of the output waveform in terms of the firing angle θ.

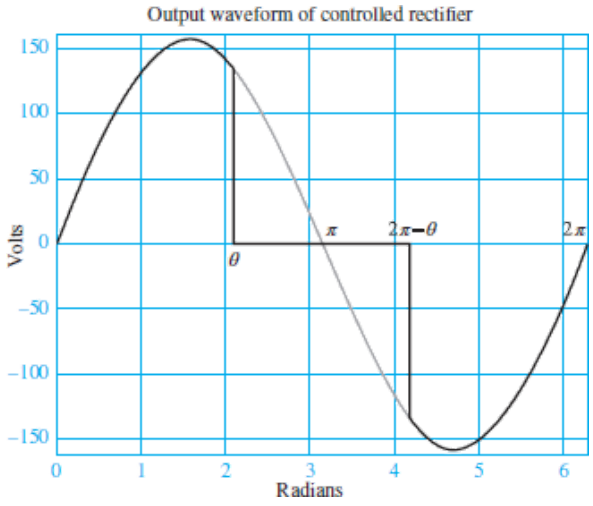


Figure P3.30

3.31 Refer to Problem 3.30 and find the angle θ that would cause the rectified waveform to deliver to a resistive load exactly one-half of the total power delivered to the same load by the input waveform.

3.32 Find the ratio between the average and rms value of the waveform shown in Figure P3.32.

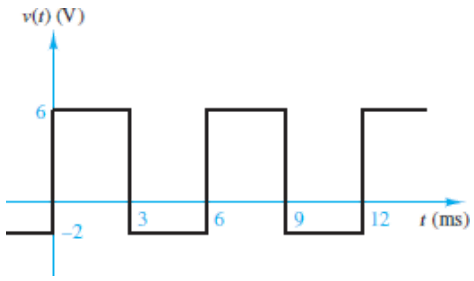


Figure P3.32

3.33 The current through a 1-Ω resistor is shown in Figure P3.33. Find the power dissipated by the resistor.

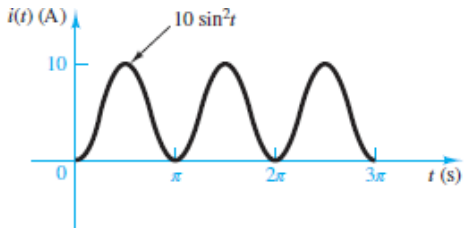


Figure P3.33

3.34 Derive the ratio between the average and r\ms value of the voltage waveform of Figure P3.34.

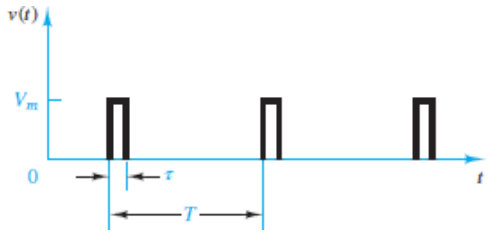


Figure P3.34

3.35 Find the rms value of the current waveform shown in Figure P3.35.

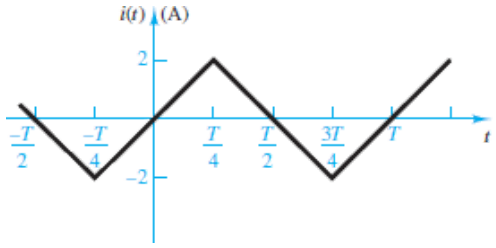


Figure P3.35

3.36 Determine the rms (or effective) value of $v(t) = V_{DC} + v_{ac} = 35 + 63 \sin(215t)$ V

Section 3.4: Phasor Solution of Circuits With Sinusoidal Sources

3.37 Find the phasor form of the following functions:

a. $v(t) = 155\cos(377t - 25°)$ V

b. $v(t) = 5\sin(1{,}000t - 40°)$ V

c. $i(t) = 10\cos(10t + 63°) + 15\cos(10t - 42°)$ A

d. $i(t) = 460\cos(500\pi t - 25°) - 220\sin(500\pi t + 15°)$ A

3.38 Convert the following complex numbers to polar form:

a. $7 + j9$

b. $-2 + j7$

c. $j\frac{2}{3} + 4 - j\frac{1}{3} + 3$

3.39 Convert the rectangular forms to polar form and compute the product. Also compute the product directly using the rectangular forms. Compare the results.

a. $(50 + j10)(4 + j8)$

b. $(j2 - 2)(4 + j5)(2 + j7)$

3.40 Complete the following exercises in complex arithmetic.

a. Find the complex conjugate of $(4 + j4)$, $(2 - j8)$, $(-5 + j2)$.

b. Multiply the numerator and denominator of each ratio by the complex conjugate of the denominator. Use the result to express each ratio in polar form.

$$\frac{1 + j7}{4 + j4} \qquad \frac{j4}{2 - j8} \qquad \frac{1}{-5 + j2}$$

c. Convert the numerator and denominator of each ratio in part b to polar form. Use the result to express each ratio in polar form.

3.41 Convert the following expressions to rectangular form:

$$j^j \qquad e^{-j\pi} \qquad e^{j2\pi}$$

3.42 Find $v(t) = v_1(t) + v_2(t)$ where

$$v_1(t) = 10\cos(\omega t + 30°)$$
$$v_2(t) = 20\cos(\omega t + 60°)$$

using:

Page 245

a. Trigonometric identities.

b. Phasors.

3.43 The current through and the voltage across a circuit element are, respectively,

$$i(t) = 8\cos\left(\omega t + \frac{\pi}{4}\right) \quad \text{A}$$

$$v(t) = 2\cos\left(\omega t - \frac{\pi}{4}\right) \quad \text{V}$$

where $\omega = 600$ rad/s. Determine:

a. Whether the element is a resistor, capacitor, or inductor.

b. The value of the element in ohms, farads, or henrys.

3.44 Express the sinusoidal waveform shown in Figure P3.44 using time-dependent and phasor notation.

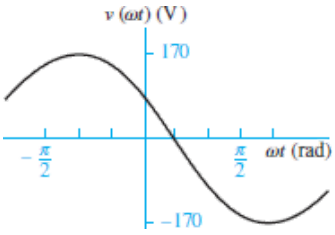


Figure P3.44

3.45 Express the sinusoidal waveform shown in Figure P3.45 using time-dependent and phasor notation.

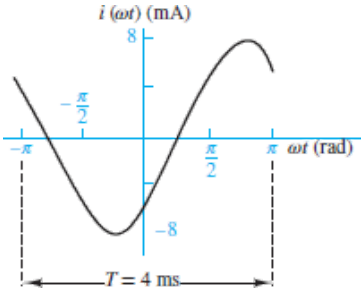


Figure P3.45

Section 3.5: Impedance

3.46 Convert the following pairs of voltage and current waveforms to phasor form. Each pair of waveforms corresponds to an unknown element. Determine whether each element is a resistor, a capacitor, or an inductor, and compute the value of the corresponding parameter **R**, **C**, or **L**.

a. $v(t) = 20\cos(400t + 1.2),\ i(t) = 4\sin(400t + 1.2)$

b. $v(t) = 9\cos(900t - \frac{\pi}{3}),\ i(t) = 4\sin\left(900t + \frac{2}{3}\pi\right)$

c. $v(t) = 13\cos\left(250t + \frac{\pi}{3}\right),\ i(t) = 7\sin\left(250t + \frac{5}{6}\pi\right)$

3.47 Determine the equivalent impedance seen by the voltage source v_S in Figure P3.47 when:

$$v_s(t) = 10\cos(4{,}000t + 60°)\ \text{V}$$
$$R_1 = 800\ \Omega \qquad R_2 = 500\ \Omega$$
$$L = 200\ \text{mH} \qquad C = 70\ \text{nF}$$

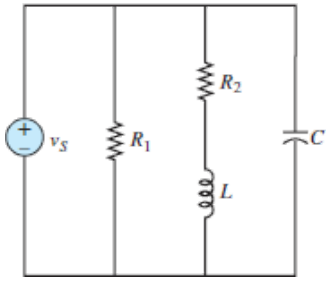


Figure P3.47

3.48 Determine the equivalent impedance seen by the voltage source v_S in Figure P3.47 when:

$$v_s(t) = 5\cos(1{,}000t + 30°)\ \text{V}$$
$$R_1 = 300\ \Omega \qquad R_2 = 300\ \Omega$$
$$L = 100\ \text{mH} \qquad C = 50\ \text{nF}.$$

3.49 The generalized version of Ohm's law for impedance elements is

$$\mathbf{V} = \mathbf{IZ}$$

Assume the current through a 0.5-μF capacitor is given by:

$$i_C(t) = I_o\cos\left(\omega t + \frac{\pi}{6}\right)$$
$$I_o = 13\ \text{mA} \qquad \omega = 1{,}000\ \text{rad/s}$$

*I*o = 13 mA ω = 1,000 rad/s

a. Express the current in phasor notation.

b. Determine the impedance of the capacitor.

c. Determine the voltage across the capacitor, in phasor notation.

3.50 Determine $i_2(t)$ in the circuit shown in Figure P3.50. Assume:

$$i_1(t) = 100\cos(\omega t + 4)\,\text{mA}$$
$$i_3(t) = 80\sin(\omega t - 1.2)\,\text{mA}$$
$$i_4(t) = 150\sin(\omega t + 2)\,\text{mA}$$
$$\omega = 377\,\text{rad/s}$$

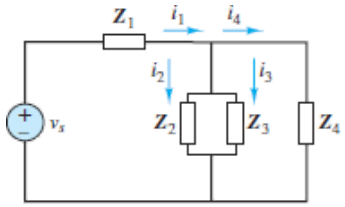


Figure P3.50

3.51 Use phasor techniques to solve for the current $i(t)$ shown in Figure P3.51.

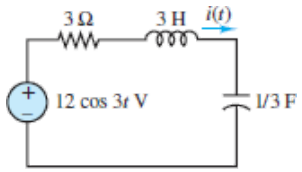


Figure P3.51

3.52 Use phasor techniques to solve for the voltage $v(t)$ shown in Figure P3.52.

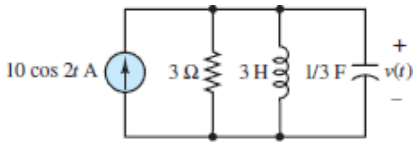


Figure P3.52

3.53 With reference to Problem 3.52, find the value of ω for which the current through the resistor is maximum.

3.54 Find $v_{out}(t)$ shown in Figure P3.54.

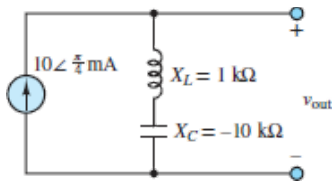


Figure P3.54

3.55 Find the impedance **Z** shown in Figure P3.55, assuming $\omega = 2$ rad/s, $R_1 = R_2 = 2$ Ω, $C = 0.25$ F, and $L = 1$ H.

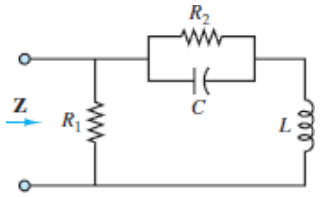


Figure P3.55

3.56 Find the sinusoidal steady-state output $v_{out}(t)$ for each circuit shown in Figure P3.56.

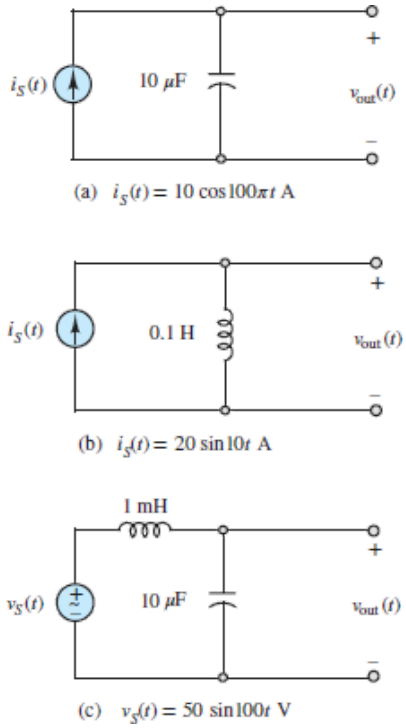


Figure P3.56

3.57 Find the frequency that causes the equivalent impedance $\mathbf{Z}_{eq}$ in Figure P3.57 to be purely resistive.

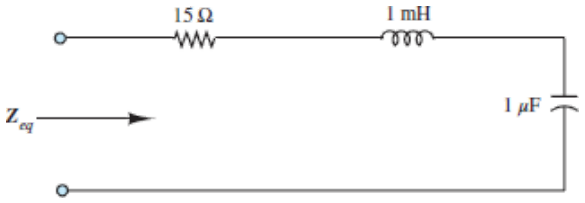


Figure P3.57

Page 247

3.58 a. Find the equivalent impedance $\mathbf{Z}_o$ seen by the voltage source in Figure P3.58(a). Assume the frequency is 377 rad/s.

b. What capacitance should be placed between terminals *a* and b, as shown in Figure P3.58(b), to make the equivalent impedance $\mathbf{Z}_o$ purely resistive? [*Hint:* Find *C* so that the phase angle of $\mathbf{Z}_o$ is zero.]

c. What is the amplitude of $\mathbf{Z}_o$ when the capacitor is included?

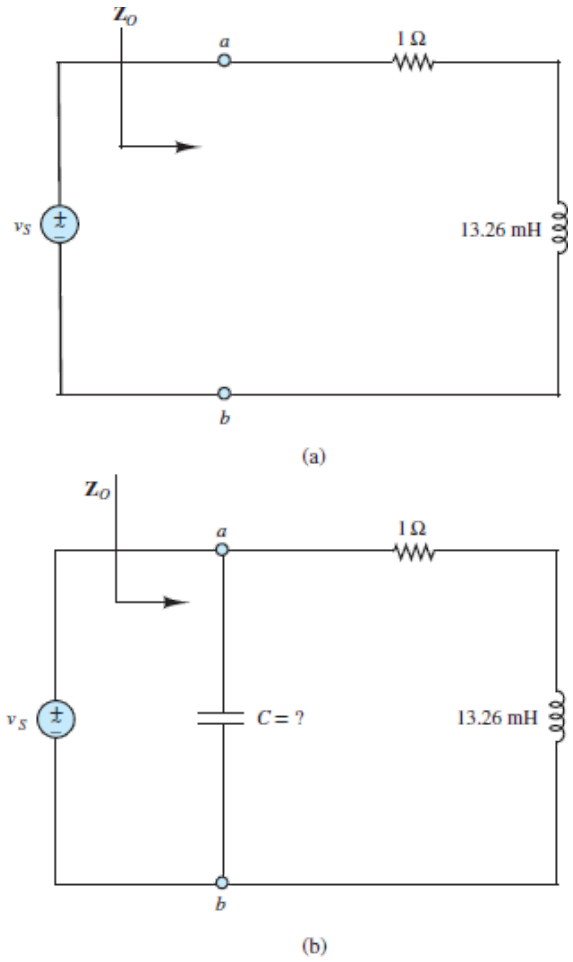


Figure P3.58

3.59 A common model for a practical capacitor has a "leakage" resistance, R_C, in parallel with an ideal capacitor, as shown in Figure P3.59. The effects of lead wires are also represented by resistances R_1 and R_2 and inductances L_1 and L_2.

a. Assume $C = 1\ \mu$F, $R_C = 100\ \text{M}\Omega$, $R_1 = R_2 = 1\ \mu\Omega$, and $L_1 = L_2 = 0.1\ \mu$H, and find the equivalent impedance $\mathbf{Z}_{ab}$ seen across terminals a and b as a function of frequency ω.

b. Find the range of frequencies for which $\mathbf{Z}_{ab}$ is capacitive.

[*Hint:* Assume that R_C is much greater than $1/\omega C$ such that R_C can be ignored in part b.]

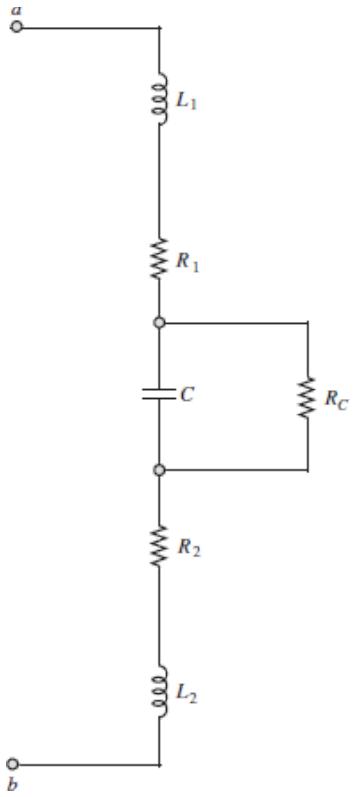


Figure P3.59

Section 3.6: AC Circuit Analysis

3.60 Determine the voltage $v_2(t)$ across R_2 in the circuit of Figure P3.60.

$i(t) = 20\cos(533.33t)$ A
$R_1 = 8\ \Omega$ $R_2 = 16\ \Omega$
$L = 15$ mH $C = 117\ \mu$F

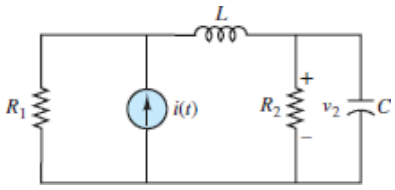


Figure P3.60

Page 248

3.61 Determine the frequency so that the current $\mathbf{I}_i$ and the voltage $\mathbf{V}_o$ in Figure P3.61 are in phase.

$$\mathbf{Z}_s = 13{,}000 + j\omega 3 \quad \Omega$$

$$R = 120\ \Omega$$

$$L = 19\ \text{mH} \qquad C = 220\ \text{pF}$$

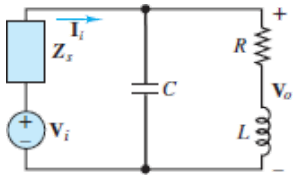


Figure P3.61

3.62 A common model for a practical inductor is a coil resistance in series with an inductance L. The coil resistance accounts for the internal losses of an inductor. Figure P3.62 shows an ideal capacitor in parallel with a practical inductor. Determine the current supplied by the voltage source v_S. Assume:

$$v_S(t) = V_o \cos(\omega t + 0)$$

$$V_o = 10\ \text{V} \qquad \omega = 6\ \text{Mrad/s} \qquad R_S = 50\ \Omega$$

$$R_C = 40\ \Omega \qquad L = 20\ \mu\text{H} \qquad C = 1.25\ \text{nF}$$

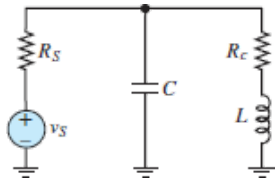


Figure P3.62

3.63 Solve for $\mathbf{I}_1$ in the circuit shown in Figure P3.63.

$$\mathbf{I} = 20\angle -\tfrac{\pi}{4}\ \text{A} \qquad R = 3\ \Omega$$

$$\mathbf{Z}_1 = -j3\ \Omega \qquad \mathbf{Z}_2 = -j7\ \Omega$$

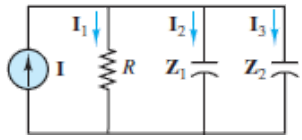


Figure P3.63

3.64 Solve for $\mathbf{V}_R$ shown in Figure P3.64. Assume:

$\omega = 3$ rad/s $\quad \mathbf{V}_S = 13\angle 0$ V

$R = 15\,\Omega \quad L_1 = 7$ H $\quad L_2 = 2$ H

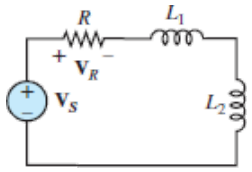


Figure P3.64

3.65 Find the current $i_R(t)$ through the resistor shown in Figure P3.65.

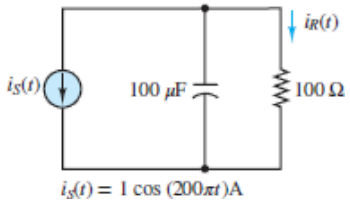


Figure P3.65

3.66 Determine the voltage $v_L(t)$ across the inductor shown in Figure P3.66.

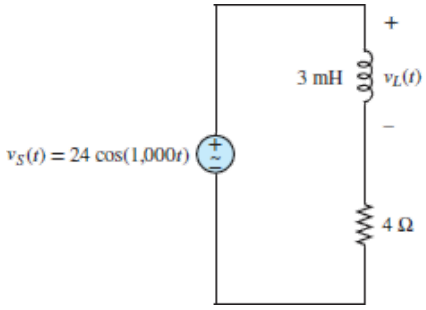


Figure P3.66

3.67 Determine the current $i_R(t)$ through the resistor shown in Figure P3.67. Assume $i_S(t)$ is given in amperes.

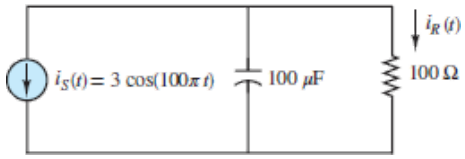


Figure P3.67

3.68 Using phasor techniques, solve for v_{R2} shown in Figure P3.68.

$$i(t) = 3\cos(200t)\text{ A}$$
$$R_1 = 3\,\Omega \qquad R_2 = 5\,\Omega$$
$$L = 18\text{ mH} \qquad C = 170\,\mu\text{F}.$$

Page 249

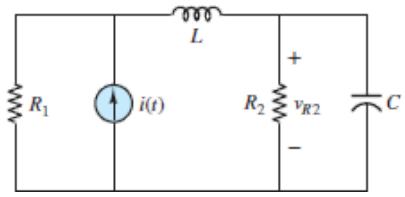


Figure P3.68

3.69 Use phasor techniques to solve for i_L in the circuit shown in Figure P3.69.

$$i_1(t) = 5\cos(500t)\text{ A}$$
$$i_2(t) = 5\cos(500t)\text{ A}$$
$$R = 5\,\Omega \qquad C = 2\text{ mF} \qquad L = 2\text{ mH}$$

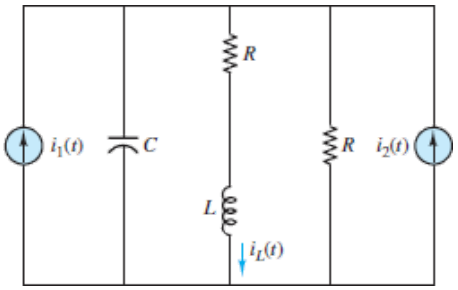


Figure P3.69

3.70 Determine the Thévenin equivalent network seen by the load R_o in Figure P3.70. Assume:

$$R_S = R_o = 500\,\Omega \qquad L = 10\text{ mH} \qquad R = 1\text{ k}\Omega$$

and:

a. $v_S(t) = 10\cos(1{,}000t)$

b. $v_S(t) = 10\cos(1{,}000{,}000t)$

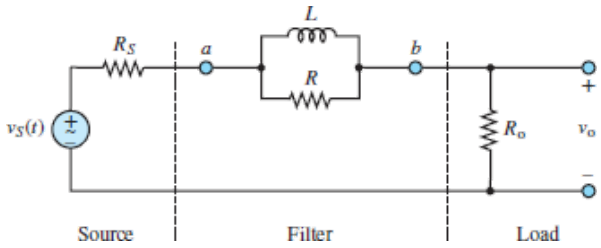


Figure P3.70

3.71 Determine the Norton equivalent network seen by the capacitor in Figure P3.71. Use the result and current division to find $i_C(t)$. Assume:

$$i(t) = 0.5\cos(300t)\text{ A}$$
$$R_1 = R_2 = 40\ \Omega$$
$$L_1 = L_2 = 200\text{ mH}$$
$$C = 15\ \mu\text{F}$$

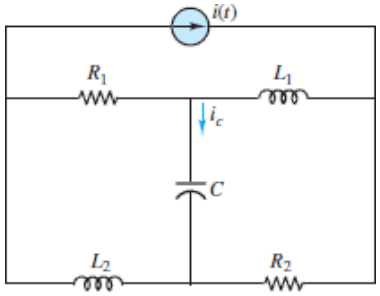


Figure P3.71

3.72 Use phasor techniques to solve for $i_L(t)$ in Figure P3.72. Assume $v_S(t) = 2\cos 2t$ V, $R_1 = R_2 = 4\ \Omega$, $L = 2$ H, and $C = 0.25$ F.

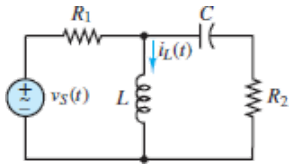


Figure P3.72

3.73 Use the mesh current method to determine the currents $i_1(t)$ and $i_2(t)$ in Figure P3.73. Assume $\mathbf{V}_1 = 10e^{-j40}$ V, $\mathbf{V}_2 = 12e^{j40}$ V, $R_1 = 8\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 6\ \Omega$, $X_L = 10\ \Omega$, $X_C = -14\ \Omega$.

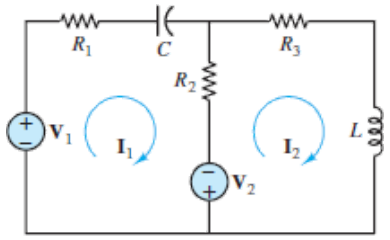


Figure P3.73

Page 250

3.74 Use the node voltage method to determine the node voltages $v_a(t)$ and $v_b(t)$ shown in Figure P3.74. Assume:

$$i(t) = 2\cos(300t)\ \text{A}$$

$$v(t) = 7\cos(300t + \pi/4)\ \text{V}$$

$$R_1 = 4\ \Omega \qquad R_2 = 3\ \Omega \qquad R_3 = 5\ \Omega$$

$$L = 300\ \text{mH} \qquad C = 300\ \mu\text{F}$$

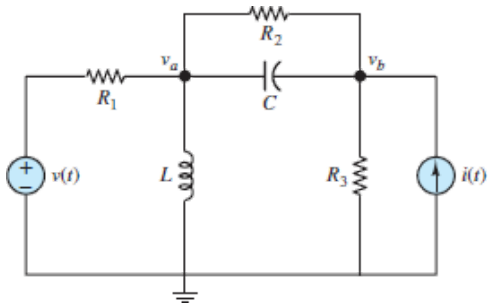


Figure P3.74

3.75 The circuit shown in Figure P3.75 is a Wheatstone bridge, which can be used to determine the reactance X_4 of an inductor or capacitor. R_1 and R_2 are adjusted until v_{ab} is zero.

a. Assume a balanced bridge ($v_{ab} = 0$) and determine X_4 in terms of the other circuit elements.

b. Assume a balanced bridge and let $C_3 = 4.7\ \mu\text{F}$, $L_3 = 0.098$ H, $R_1 = 100\ \Omega$, $R_2 = 1\ \Omega$, and $v_S(t) = 24\sin(2{,}000t)$. What is the reactance of the unknown circuit element? Is it a capacitor or an inductor? What is its value?

c. What frequency should be avoided in this circuit, and why?

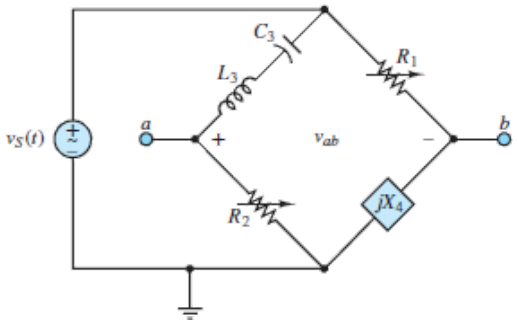


Figure P3.75

3.76 Find the Thévenin equivalent network seen by the capacitor C in Figure P3.76. Use the result and voltage division to determine $v_C(t)$. Assume:

$$v(t) = \cos(300t)\ \text{V}$$
$$i(t) = 2\cos(300t)\ \text{A}$$
$$R_1 = 8\ \Omega \quad R_2 = 8\ \Omega$$
$$L = 3\ \mu\text{H} \quad C = 5\ \mu\text{F}$$

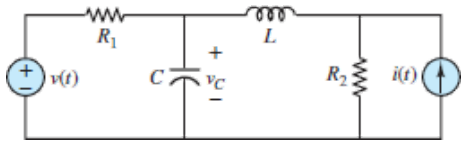


Figure P3.76

3.77 Determine the Thévenin equivalent network seen by the load $\mathbf{Z}_o$ shown in Figure P3.77. Assume: $\mathbf{V}_S = 10\angle 0°$ V, $R_S = 40\ \Omega$, $X_L = 40\ \Omega$, and $X_C = -2{,}000\ \Omega$.

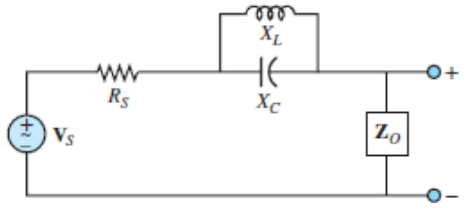


Figure P3.77

3.78 Find the Thévenin equivalent network seen across terminals a and b in Figure P3.78.

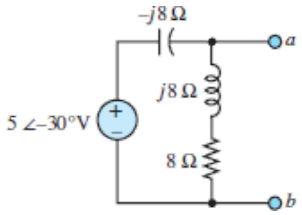


Figure P3.78

3.79 Determine the Norton equivalent network seen by the capacitor in Figure P3.79. Use the result and current division to find $i_C(t)$. Assume:

$$v_S(t) = 4\cos(100t)\,\text{V}$$
$$R_1 = 7\,\Omega \qquad R_2 = 8\,\Omega$$
$$L = 30\,\text{mH} \qquad C = 10\,\text{mF}$$

Page 251

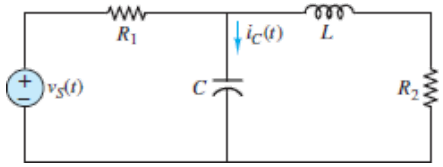


Figure P3.79

3.80 Find the Thévenin equivalent network seen by R_2 in Figure P3.80. Use the result and voltage division to determine the voltage $v_2(t)$ across R_2. Assume:

$$v(t) = 70\cos(275t)\,\text{V}$$
$$R_1 = R_2 = 42\,\Omega$$
$$L = 1\,\text{mH} \qquad C = 12\,\mu\text{F}.$$

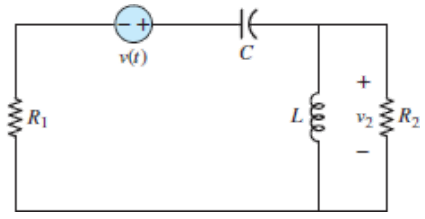


Figure P3.80

3.81 Use the mesh current method to find the phasor mesh current equations of Figure P3.81.

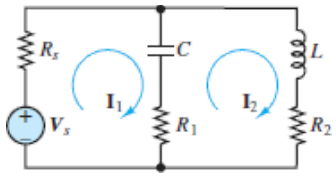


Figure P3.81

3.82 Write the node equations required to solve for all voltages and currents in the circuit of Figure P3.81. Assume all impedances and the source voltage are known.

3.83 Determine $\mathbf{V}_o$ in the circuit of Figure P3.83. Assume:

$$\mathbf{V}_i = 4\angle\frac{\pi}{6}\ \text{V} \qquad \omega = 1{,}000\ \text{rad/s}$$
$$L = 60\ \text{mH} \qquad C = 12.5\ \mu\text{F}$$
$$R_o = 120\ \Omega$$

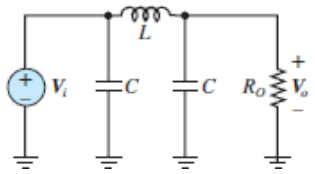


Figure P3.83

Section 3.7: Instantaneous and Average Power

3.84 The heating element in a soldering iron has a resistance of 20 Ω. Find the average power dissipated in the soldering iron if it is connected to a voltage source of 90 V rms.

3.85 A coffeemaker has a rated power of 1,000 W at 240 V rms. Find the resistance of the heating element.

3.86 A current source $i(t)$ is connected to a 50-Ω resistor. Find the average power delivered to the resistor, given that $i(t)$ is

a. $7\cos 100t$ A

b. $7\cos(100t - 30°)$ A

c. $7\cos 100t - 3\cos(100t - 60°)$ A

d. $7\cos 100t - 3$ A

3.87 Find the rms value of each of the following periodic currents:

a. $\cos 200t + 3\cos 200t$

b. $\cos 10t + 2 \sin 10t$

c. $\cos 50t + 1$

d. $\cos 30t + \cos(30t + \pi/6)$

3.88 A current of 2.5 A through a neon light advertisement is supplied by a 115 V rms voltage source. The current lags the voltage by 30°. Find the impedance of the light, the real power dissipated by it, and its power factor.

3.89 Compute the average power dissipated by the load seen by the voltage source in Figure P3.89. Let ω = 377 rad/s, $\tilde{\mathbf{V}}_s = 50\angle 0$ V, R = 10 Ω, L = 0.08 H, and C = 200 μF.

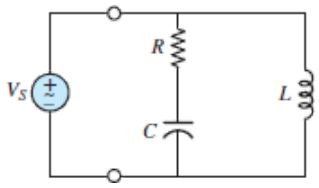


Figure P3.89

Page 252

3.90 A drilling machine is driven by a single-phase induction machine connected to a 110 V rms supply. Assume that the machining operation requires 1 kW, that the tool machine has 90 percent efficiency, and that the supply current is 14 A rms with a power factor of 0.8. Find the AC machine efficiency.

3.91 Given the waveform of a voltage source shown in Figure P3.91, find:

a. The steady DC voltage that would cause the same heating effect across a resistance.

b. The average current supplied to a 10-Ω resistor connected across the voltage source.

c. The average power supplied to a 1-Ω resistor connected across the voltage source.

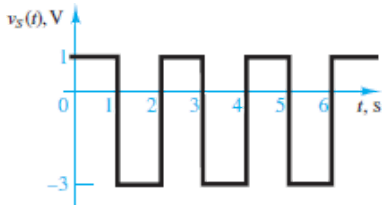


Figure P3.91

3.92 A current source $i(t)$ is connected to a 100-Ω resistor. Find the average power delivered to the resistor, given that $i(t)$ is:

a. $4\cos(100t)$ A

b. $4\cos(100t - 50°)$ A

c. $4\cos(100t - 3)\cos(100t - 50°)$ A

d. $4\cos(100t - 3)$ A

3.93 Find the rms value of each of the following periodic currents:

$\cos(2t) + \sin(2t)$ A

$\cos(377t) + 1$ A

$\cos(2t) + \cos(2t + 135°)$ A

$\cos(2t) + \cos(3t)$ A

Section 3.8: Apparent Power and the Power Triangle

3.94 A current of 10 A rms results when a single-phase circuit is placed across a 220 V rms source. The current lags the voltage by 60°. Find the power dissipated by the circuit and the power factor.

3.95 An inductive network is supplied by a 120 V rms, 60-Hz voltage source. An ammeter and a wattmeter indicate that 12 A rms is drawn from the source and 800 W are consumed by the network. Determine:

a. The network power factor.

b. The network phase angle.

c. The network impedance.

d. The equivalent resistance and reactance of the network.

3.96 For the following numeric values, determine the average power, P, the reactive power, Q, and the complex power, S, of the circuit shown in Figure P3.96. *Note:* Phasor quantities are rms.

a. $v_S(t) = 650\cos(377t)$ V

$i_o(t) = 20\cos(377t - 10°)$ A

b. $\tilde{\mathbf{V}}_S = 460\angle 0°$ V rms

$\tilde{\mathbf{I}}_o = 14.14\angle -45°$ A rms

c. $\tilde{\mathbf{V}}_S = 100\angle 0°$ V rms

$\tilde{\mathbf{I}}_o = 8.6\angle -86°$ A rms

d. $\tilde{\mathbf{V}}_S = 208\angle -30°$ V rms

$\tilde{\mathbf{I}}_o = 2.3\angle -63°$ A rms

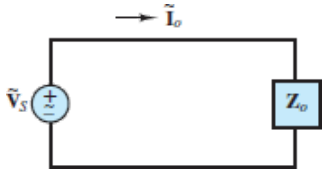


Figure P3.96

3.97 For the circuit of Figure P3.96, determine the power factor for the load $\boldsymbol{Z}_o$ and determine whether it is leading or lagging for the following conditions:

a. $v_S(t) = 679\cos(\omega t + 15°)$ V

$i_o(t) = 20\cos(\omega t + 47°)$ A

b. $v_S(t) = 163\cos(\omega t + 15°)$ V

$i_o(t) = 20\cos(\omega t - 22°)$ A

c. $v_S(t) = 294\cos(\omega t)$ V

$i_o(t) = 1.7\cos(\omega t + 175°)$

d. $\mathbf{Z}_o = (48 + j16)\ \Omega$

3.98 For the circuit of Figure P3.96, determine whether the load is capacitive or inductive, assuming:

a. $pf = 0.87$ (leading)

b. $pf = 0.42$ (leading)

Page 253

c. $v_S(t) = 42\cos(\omega t)$ V

$i_L(t) = 4.2\sin(\omega t)$ A

d. $v_S(t) = 10.4\cos(\omega t - 12°)$ V

$i_L(t) = 0.4\cos(\omega t - 12°)$ A

3.99 For the circuit shown in Figure P3.99, assume $C = 265\ \mu$F, $L = 25.55$ mH, and $R = 10\ \Omega$. Find the real and reactive power if:

a. $v_S(t) = 120\cos(377t)$ V (i.e., the frequency is 60 Hz)

b. $v_S(t) = 650 \cos(314t)$ V (i.e., the frequency is 50 Hz)

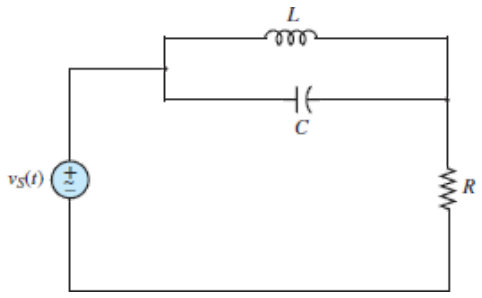


Figure P3.99

3.100 A load impedance, $\mathbf{Z}_o = 10 + j3\ \Omega$, is connected to a source with line resistance equal to 1 Ω, as shown in Figure P3.100. Calculate the following values:

a. The average power delivered to the load.

b. The average power absorbed by the line.

c. The apparent power supplied by the generator.

d. The power factor of the load.

e. The power factor of line plus load.

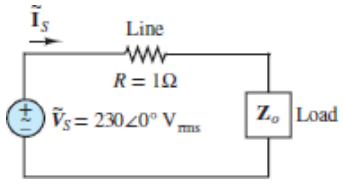


Figure P3.100

3.101 For the circuit shown in Figure P3.101, find:

a. The Thévenin equivalent network seen by the load.

b. The power dissipated by the load resistor.

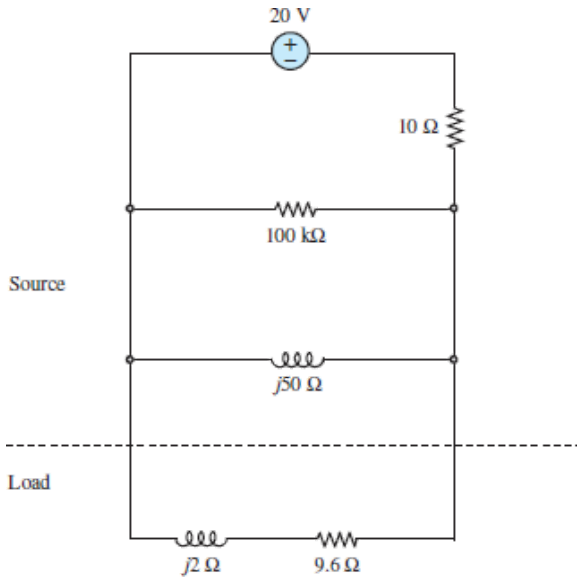


Figure P3.101

3.102 For the circuit of Figure P3.96, determine the power factor of the load for each case listed below. Is it leading or lagging?

a. $v_s(t) = 50\cos(\omega t)$ V

$i_o(t) = 20\sin(\omega t + 1.2)$ A

b. $v_s(t) = 110\cos(\omega t + 0.1)$ V

$i_o(t) = 10\cos(\omega t - 0.1)$ A

c. $\mathbf{Z}_o = (20 + j5)\ \Omega$

d. $\mathbf{Z}_o = (20 - j5)\ \Omega$

3.103 For the circuit of Figure P3.96, determine whether the load $\mathbf{Z}_o$ is capacitive or inductive, if:

a. its power factor is pf = 0.76 lagging.

b. its power factor is pf = 0.5 (leading).

c. $v_s(t) = 10\cos(\omega t)$ V, $i_o(t) = \cos(\omega t)$ A.

d. $v_s(t) = 100\cos(\omega t)$ V, $i_o(t) = 12\cos(\omega t + \pi/4)$ A.

3.104 Find the real and reactive power supplied by the voltage source shown in Figure P3.104 for $\omega = 5$ rad/s and $\omega = 15$ rad/s. Let $v_S = 15\cos(\omega t)$ V, $R = 5\ \Omega$, $C = 0.1$ F, $L_1 = 1$ H, $L_2 = 2$ H.

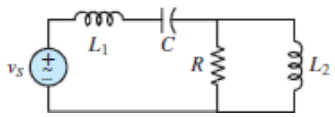


Figure P3.104

Page 254

3.105 In Figure P3.105, assume $\tilde{\mathbf{V}}_{S1} = 10\angle -\pi/4$ V rms, $\tilde{\mathbf{V}}_{S2} = 12\angle 0.8$ V rms, $R_1 = 2\,\Omega$, $R_2 = 3\,\Omega$, $X_L = 4\,\Omega$, and $X_C = -4\,\Omega$. Find:

a. The amplitude of the current supplied by each source.

b. The total real power supplied by each source.

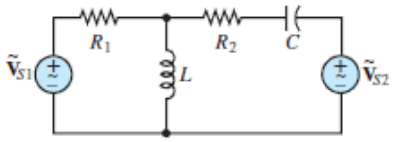


Figure P3.105

3.106 The load $\mathbf{Z}_o$ shown in Figure P3.106 consists of a 20-Ω resistor in series with a 0.01-H inductor. Assuming $f = 60$ Hz, $R = 0.5\,\Omega$, $\tilde{\mathbf{V}}_S = 100\angle 0$ V rms. Calculate:

a. The apparent power supplied by the voltage source.

b. The apparent power delivered to the load.

c. The power factor of the load.

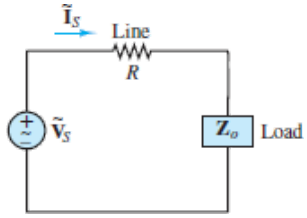


Figure P3.106

3.107 Calculate the real and reactive power of the load between terminals *a* and *b* in Figure P3.107. Assume $f = 60$ Hz, $\tilde{\mathbf{V}}_S = 70\angle 0$ V rms, $R_S = 2\,\Omega$, $R_o = 18\,\Omega$, and $X_L = 5\,\Omega$.

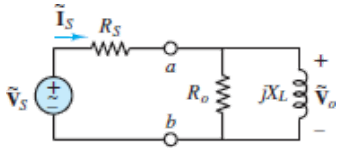


Figure P3.107

3.108 Calculate the apparent power, real power, and reactive power supplied by the voltage source shown in Figure P3.108. Draw the power triangle. Assume $f = 60$ Hz, $\tilde{\mathbf{V}}_S = 70\angle 0$ V rms, $R = 18\ \Omega$, $C = 50\ \mu$F, and $L = 0.001$ H.

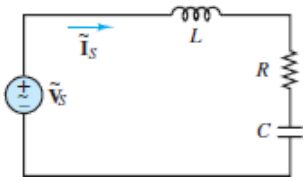


Figure P3.108

3.109 Suppose that the electricity in your home has gone out on a hot, humid summer day and the power company will not be able to fix the problem for several days. The freezer in the basement contains $300 worth of food that you cannot afford to let spoil. You would also like to keep one window air conditioner running, as well as run the refrigerator in your kitchen. When these appliances are on, they draw the following currents (all values are rms):

Air conditioner:	9.6 A rms @ 120 V rms
	pf = 0.90 lagging
Freezer:	4.2 A rms @ 120 V rms
	pf = 0.87 lagging
Refrigerator:	3.5 A rms @ 120 V rms
	pf = 0.80 lagging

In the worst-case scenario, how much power must an emergency generator supply?

3.110 The French TGV high-speed train absorbs 11 MW at 300 km/h (186 mi/h). The power supply module shown in Figure P3.110 consists of two 25-kV rms single-phase AC power stations connected at the same overhead line, one at each end of the module. For the return circuits, the rail is used. The train is also designed to operate at a low speed with 1.5-kV DC in railway stations or under the old electrification lines. The natural (average) power factor in the AC operation is 0.8. Assume that the equivalent specific resistance of the overhead line is 0.2 Ω/km and that the rail resistance can be neglected. Find:

a. A simple circuit model for the system.

b. The locomotive's current in the condition of a 10 percent voltage drop.

c. The reactive power supplied by the power stations.

d. The supplied real power, overhead line losses, and maximum distance between two power stations supplied in the condition of a 10 percent voltage drop when the train is located at the half-distance between the stations.

e. Overhead line losses in the condition of a 10 percent voltage drop when the train is located at the half-distance between the stations, assuming pf = 1. (The French TGV is designed with a state-of-the-art power compensation system.)

Page 255

f. The maximum distance between the two power stations supplied in the condition of a 10 percent voltage drop when the train is located at the half-distance between the stations, assuming the DC (1.5-kV) operation at one-quarter power.

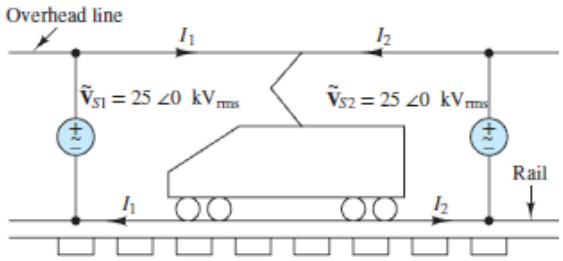


Figure P3.110

3.111 The voltage and current supplied by a source to a load are

$$\tilde{\mathbf{V}}_S = 7\angle 0.873 \text{ V rms} \qquad \tilde{\mathbf{I}}_S = 13\angle(-0.349) \text{ A rms}$$

Determine:

a. The real power consumed as work and dissipated as heat in the load.

b. The reactive power stored in the load.

c. The impedance angle of the load and its power factor.

3.112 Determine the real power dissipated and the reactive power stored in each of the impedances shown in Figure P3.112. Assume:

$$\tilde{\mathbf{V}}_{s1} = \frac{170}{\sqrt{2}}\angle 0 \text{ V rms}$$

$$\tilde{\mathbf{V}}_{s2} = \frac{170}{\sqrt{2}}\angle \pi/2 \text{ V rms}$$

$$\omega = 377 \text{ rad/s}$$

$$\mathbf{Z}_1 = 0.7\angle\frac{\pi}{6}\ \Omega$$

$$\mathbf{Z}_2 = 1.5\angle 0.105\ \Omega$$

$$\mathbf{Z}_3 = 0.3 + j0.4\ \Omega$$

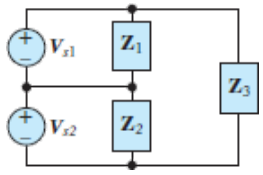


Figure P3.112

3.113 The following are supplied by a source to a load:

$$\tilde{\mathbf{V}}_s = 170\angle(-0.157) \text{ V rms} \qquad \tilde{\mathbf{I}}_s = 13\angle 0.28 \text{ A rms}$$

Determine:

a. The real power consumed as work and dissipated as heat in the load.

b. The reactive power stored in the load.

c. The impedance angle of the load and its power factor.

Section 3.9: Power Factor Correction

3.114 A single-phase motor draws 220 W at a power factor of 0.8 lagging when connected across a 240 V rms, 60-Hz source. A capacitor is connected in parallel with the load to produce a unity power factor. Determine the required capacitance.

3.115 The networks seen by the voltage sources in Figure P3.115 have unity power factor. Determine C_P and C_S.

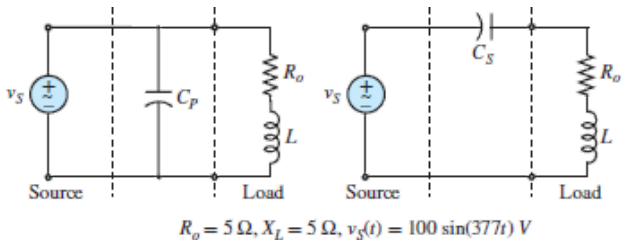


Figure P3.115

3.116 A 1,000-W electric motor is connected to a 120 V_{rms}, 60-Hz source. The power factor seen by the source is 0.8, lagging. To correct the pf to 0.95 lagging, a capacitor is placed in parallel with the motor. Calculate the current drawn from the source with and without the capacitor connected. Determine the value of the capacitor required to make the correction.

3.117 The motor inside a blender can be modeled as a resistance in series with an inductance, as shown in Figure P3.117. The wall socket source is modeled as an ideal 120 V rms voltage source in series with a 2-Ω output resistance. Assume the source frequency is $\omega = 377$ rad/s.

a. What is the power factor of the motor?

b. What is the power factor seen by the voltage source?

c. What is the average power, P_{AV}, consumed by the motor?

d. What value of capacitor when placed in parallel with the motor will change the power factor seen by the voltage source to 0.9 lagging?

Page 256

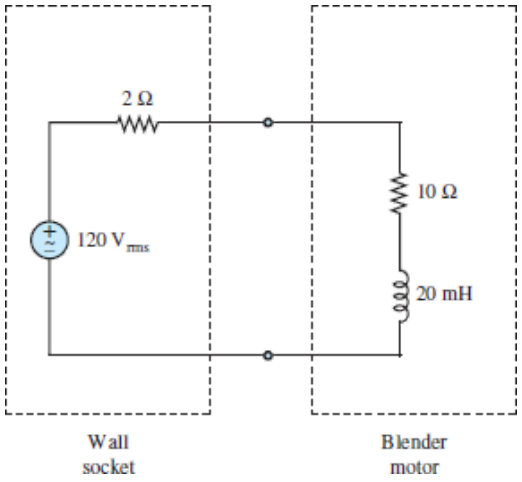


Figure P3.117

3.118 For the following numerical values, determine the capacitance to be placed in parallel with the load $\mathbf{Z}_o$ shown in Figure P3.96 that will result in a unity power factor seen by the voltage source. Assume $\omega = 377$ rad/s.

a. $\tilde{\mathbf{V}}_s = 300\angle 0\,\text{V rms}, \tilde{\mathbf{I}}_o = 80\angle(-0.15\pi)\,\text{A rms}$

b. $\tilde{\mathbf{V}}_s = 100\angle 0\,\text{V rms}, \tilde{\mathbf{I}}_o = 30\angle(-\pi/4)\,\text{A rms}$

c. $\tilde{\mathbf{V}}_s = 12\angle(-\pi/4)\,\text{V rms}, \tilde{\mathbf{I}}_o = 3\angle(-\pi/2)\,\text{A rms}$

3.119 For the circuit shown in Figure P3.119, assume $f = 60$ Hz, $\tilde{\mathbf{V}}_S = 90\angle 0$ V rms, $R = 25\ \Omega$, $X_L = 70\ \Omega$, and $X_C = -8\ \Omega$. Calculate:

a. The capacitance C and the inductance L.

b. The power factor seen by the voltage source.

c. The new capacitance required to correct that power factor to unity.

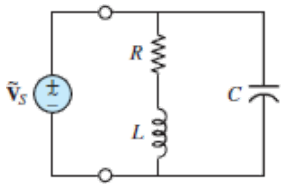


Figure P3.119

3.120 Refer to problem 3.108 and determine the capacitance needed in parallel with the voltage source to correct the power factor seen by the source to 0.95. Draw the power triangle.

3.221 A single-phase motor is modeled as a resistor R in series with an inductor L as shown in Figure P3.121. The capacitor corrects the power factor between terminals a and b to unity. Assume the meters shown are ideal and $f = 50$ Hz, $V = 220$ V rms, $I = 20$ A rms, and $I_1 = 25$ A rms. Find the capacitor value.

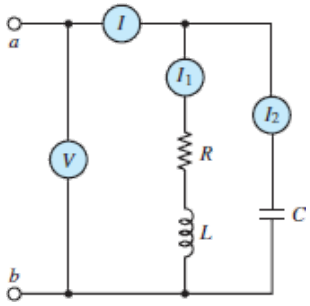


Figure P3.121

3.122 An industrial assembly hall is continuously lit by one hundred 40-W mercury vapor lamps in parallel and supplied by a 120 V rms, 60-Hz source. The power factor seen by the source is 0.65, which is so low that a 25 percent penalty is applied at billing. If the average price of 1 kWh is \$0.05 and the average cost of a capacitor is \$50 per mF, compute how long it will take before the billing penalty equals the cost of the capacitor needed to correct the power factor to 0.85.

3.123 Refer to Problem 3.122 and assume that each lamp is now available with a compensating capacitor in parallel with the original lamp. Find:

a. The compensating capacitor value for unity power factor seen by the source.

b. The maximum number of additional lamps that can be installed without exceeding the original current supplied by the source when using uncompensated lamps.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]A dielectric material is a material that is not an electrical conductor but contains a large number of electric dipoles, which become polarized in the presence of an electric field.

C H A P T E R

4

TRANSIENT ANALYSIS

Chapter 4 focuses on the *transient* portion of the complete response of a time-dependent circuit. Recall from Chapter 3 that the complete response is composed of two parts: (1) the transient response and (2) the *steady-state* response. (These parts can also be arranged as natural and forced responses, respectively.) Chapter 3 explored the latter part for circuits with AC sources; Chapter 4 explores the former for circuits that experience a *transient event,* such as the throwing of a switch. The general qualities of a transient response are independent of the type of event.

The fundamental quality of any transient response is that it eventually vanishes to zero. Once this occurs, only a steady-state solution remains. The role of the transient solution is to provide a *transition over time* from one state (i.e., an "old" or "initial" steady state) to another (i.e., a "new" or "final" steady state). The Latin root of the adjective *transient* is *trans,* meaning "across." Literally, the transient solution is a bridge across time from one steady state to another. In most of the examples presented in this chapter, both the "old" and the "new" are, for simplicity, DC steady-states. However, transient analysis is applicable to a transition between two AC steady states or any other pair of states, which need not be steady.

When a switch opens or closes in an electric circuit, the voltages and currents in that circuit will, in general, transition to a new state. The throwing of a switch is a transient event because it causes a short-circuit (a closed switch) to be replaced Page 258by an open-circuit (an open switch), or vice versa. These two switch positions

produce two distinctly different circuits. The abrupt change from one to the other provokes a transient response.

The transition from the "old" state to the "new" state does not happen instantaneously because capacitors and inductors store energy. Some finite time is required to charge and discharge the energy storage elements to reach the "new" steady state. The transition may take place quickly, but it cannot take place instantaneously. The energy stored in capacitors and inductors is a function of the capacitor voltage and inductor current, respectively. Thus, those two quantities are known as *state variables*.

The objectives of transient analysis can be expressed by the following questions:

1. What are the *initial conditions* on the *state variables* at the moment of the transient event?
2. How are the initial conditions on the state variables related to the initial conditions on other variables?
3. What is the manner of the transition from the initial conditions to the final steady state of any variable?
4. How fast or slow is that transition?
5. What is the final steady state of any variable?

Two types of circuits are examined in this chapter: first-order *RC* and *RL* circuits, which contain a single storage element, and second-order circuits, which contain two irreducible storage elements. The simplest of the second-order circuits to analyze are the series *LC* and parallel *LC* circuits. All of the fundamental aspects of transient responses are revealed in these circuits, which is why they are the focus of this chapter. However, other more complicated circuits are also explored and analyzed.

A first-order circuit contains a single storage element. A second-order circuit contains two irreducible storage elements.

Throughout this chapter, practical applications of first- and second-order circuits are introduced. Numerous analogies are presented to emphasize the general nature of the solution methods and their applicability to a wide range of physical systems, including hydraulics, mechanical systems, and thermal systems.

Learning Objectives

Students will learn to...

1. Write differential equations in standard form for circuits containing inductors and capacitors. *Section 4.2*
2. Determine the steady state of DC circuits containing inductors and capacitors *Section 4.2*
3. Determine the complete solution of first-order circuits excited by switched DC sources. *Section 4.3*
4. Determine the complete solution of second-order circuits excited by switched DC sources. *Section 4.4*

4.1 TRANSIENT ANALYSIS

Figure 4.1 shows two typical results due to a transient event at $t = 0.2$ s in a DC circuit [Figure 4.1(a)] and an AC circuit [Figure 4.1(b)], respectively. Each waveform has three parts:

- The *initial steady state* for $0 \leq t \leq 0.2$ s.
- The *transient response* for $0.2 \leq t \leq 1.8$ s (approximately).
- The *final steady state* for $t > 1.8$ s.

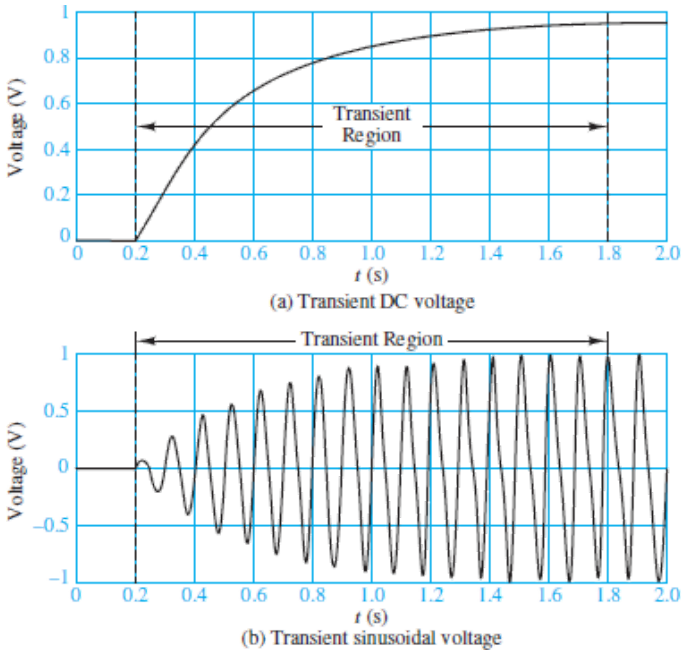


Figure 4.1 First- and second-order transient responses

The objective of **transient analysis** is to determine the manner and speed with which voltages and currents transition from one steady state to another.

Figure 4.2 shows a typical parallel *LC* circuit used to explore transient responses. The single-pole, single-throw (SPST) switch connects the battery to the *RLC* network suddenly at $t = 0$ initiating a transient response. The complexity of transient analysis increases with the number of irreducible energy storage elements in the circuit. Luckily, first- and second-order circuits exhibit all of the fundamental aspects of transient behavior.

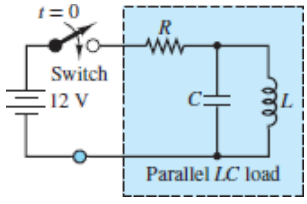


Figure 4.2 Circuit with switched DC excitation

The discussion and analysis in this chapter is focused on circuits that conform to the general circuit models shown in Figure 4.3, where the network in the box acts as the load and consists of either one or two *storage elements* and possibly various resistors. In Figure 4.3(a), R_T is the Thévenin equivalent resistance seen by the load and V_T is the open-circuit voltage across terminals *a* and *b*. In Figure 4.3(b), R_N is the

Norton equivalent resistance R_N seen by the load and I_N is the short-circuit current from terminal *a* to terminal *b*.

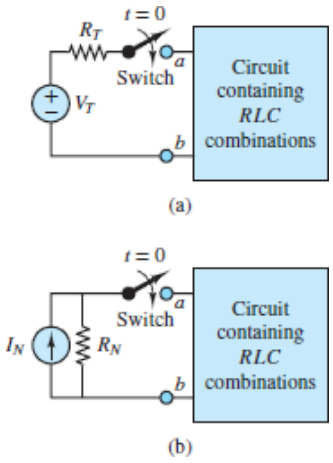


Figure 4.3 General models of the transient analysis problem. The load may contain *RLC* combinations while the source is either a (a) Thévenin or (b) Norton equivalent network.

Page 260

When the load is first order, containing either an inductor or capacitor, the transient response will be either a **rising** or **falling exponential** waveform, such as those shown in Figure 4.4. Both of these waveforms *decay* over time; that is, the transient response goes to zero leaving only the new steady-state response.

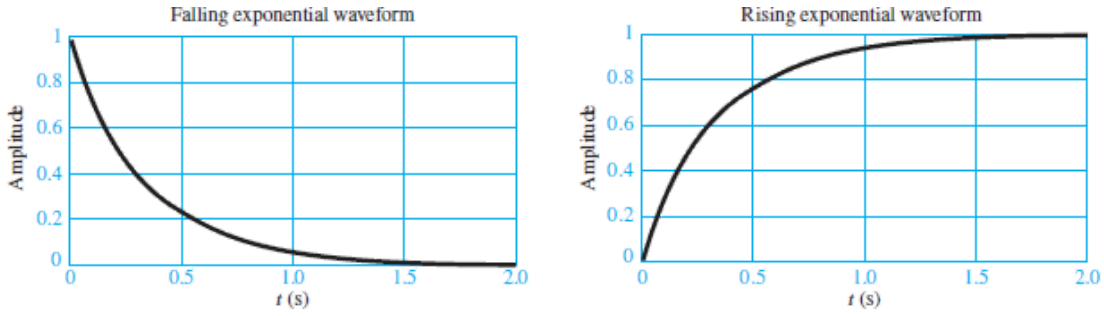


Figure 4.4 Falling and rising exponential responses

In the case of two storage elements, series and parallel *LC* networks are considered in detail although a method for solving more complicated arrangements is

also presented. The analysis of second-order circuits is complicated because there are three distinctly different transient responses possible, depending upon the magnitude of a **dimensionless damping ratio** ζ. When $\zeta > 1$, the transient response is *overdamped* and is represented by the sum of two exponentially decaying waveforms, either rising or falling. When $\zeta < 1$, the transient response is *underdamped* and is represented by a *decaying sinusoid*. When $\zeta = 1$, the transient response is *critically damped* and is represented by a waveform that has aspects of both the overdamped and underdamped waveforms. The impact of ζ ("zeta") is exemplified in the transient response to the sudden switching of a DC source, as shown in Figure 4.5.

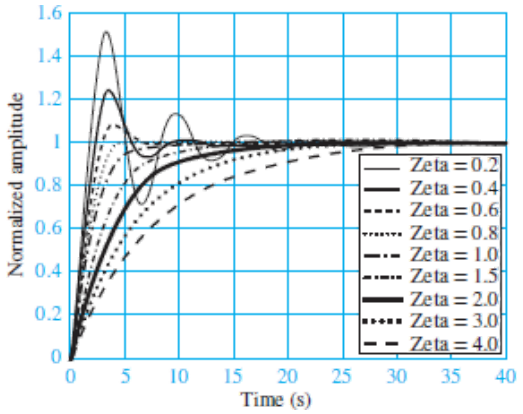


Figure 4.5 Typical second-order transient responses for various values of the dimensionless damping ratio ζ (zeta)

Page 261

4.2 ELEMENTS OF TRANSIENT PROBLEM SOLVING

The key elements involved in the solution of a first- or second-order transient problem are outlined below. The discussion in this section and the remainder of the chapter is limited to circuits containing DC sources only. The mathematics for circuits containing AC sources is somewhat more complicated; however, the fundamental ideas are the same.

Time Intervals

The moment of a transient event is defined as $t = 0$. The moments immediately before and after the event are denoted as $t = 0^-$ and $t = 0^+$, respectively. The initial steady state is determined by the behavior of the circuit for the time interval $t < 0$. The final steady state is the behavior of the circuit as $t \to \infty$, which should be understood to mean "t gets very large." In between the initial and final steady states is the transient response.

In practice, the final steady state is reached when $t \leq t_\infty$, where t_∞ marks the *effective* end of the transient response. The most common choice for t_∞ is 5τ, where τ is a *time constant* associated with the circuit.

Initial Steady State ($t < 0$)

For simplicity, the circuits examined in this chapter often assume a DC steady-state prior to the transient event. The implication is that capacitors and inductors act as open- and closed-circuits, respectively, and the DC circuit analysis methods discussed in Chapters 1 and 2 can be applied.

In a DC steady-state, a capacitor acts as an open-circuit and an inductor acts as a short-circuit.

State Variables

The state variables in electric circuits are the currents through inductors and the voltages across capacitors. The number of state variables equals the number of irreducible storage elements. Thus, first- and second-order circuits have one and two state variables, respectively. It is usually best to first solve for the transient response of the state variables and then solve for other variables through their relationships to the state variables. Regardless of the solution method employed, it is always necessary to know the values of the state variables at $t = 0^-$.

Initial Conditions

The initial conditions on the transient response of a circuit are determined by its stored energy at the instant of the transient event. Recall that energy is stored in capacitors, as expressed by their voltages, and in inductors, as expressed by their currents. Since the energy stored in a capacitor or inductor cannot change instantaneously, the voltage across a capacitor and the current through an inductor

also cannot change instantaneously. In other words, the state variables are continuous functions of time.

Page 262

The continuity requirement on the state variables is evident in the v-i relationships for capacitors and inductors.

$$i_C = C\frac{dv_C}{dt} \quad \text{and} \quad v_L = L\frac{di_L}{dt} \tag{4.1}$$

A discontinuity in v_C or i_L would require i_C or v_L, respectively, to be infinite. Since it is not physically possible to achieve an infinite current or voltage, v_C and i_L must always be continuous.

The same is not guaranteed for other nonstate variables in a circuit. The current through a resistor or capacitor, and the voltage across a resistor or inductor, may be discontinuous. An important implication of these results is that only the state variables are guaranteed to be continuous across a transient event.

Only the current through an inductor and the voltage across a capacitor are guaranteed to be continuous. Consequently, these two state variables are also continuous across a transient event. In mathematical terms:

$$v_C(0^+) = v_C(0^-) \tag{4.2}$$

$$i_L(0^+) = i_L(0^-) \tag{4.3}$$

Other variables may or may not be continuous across a transient event. Only state variables should be used to express initial conditions due to a transient event.

Energy and the Transient Response

During a transient response, energy is, in general, continually stored and released, supplied and dissipated within a circuit until a new steady state is reached. Independent voltage and current sources, if present, will supply energy; storage elements will store and/or release energy; and resistors will dissipate energy. These

processes will continue until a new steady state is reached, in which the energy supplied continually equals the energy dissipated.

Consider the circuit shown in Figure 4.6. For $t < 0$, assume that the capacitor has been connected to the battery for a long time so that the capacitor voltage v_C equals the battery voltage V_B. Notice that the current through each resistor is zero for $t < 0$.

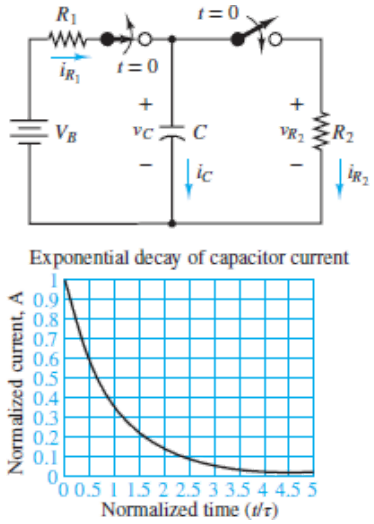


Figure 4.6 Energy stored in a capacitor is dissipated by a resistor.

At $t = 0$ the two switches are thrown such that the capacitor is disconnected from the battery loop but simultaneously connected to R_2 in a simple series loop. Since the voltage across the capacitor must be continuous with time, $v_C = V_B$ at $t = 0^+$. At the same moment, the voltage across R_2 has changed from zero to v_C and, therefore, the current through R_2 has also changed from zero to some finite nonzero value. Since KCL requires $i_C + i_{R_2} = 0$ for the series loop, the capacitor current is:

$$i_C = -i_{R_2} = -\frac{v_{R_2}}{R_2} = -\frac{v_C}{R_2} \tag{4.4}$$

Page 263

where the expression for i_{R_2} is simply Ohm's law. Use equation 4.1 to substitute for i_C to find:

$$C\frac{dv_C}{dt} = -\frac{v_C}{R_2} \tag{4.5}$$

Divide both sides of the equation by C to find:

$$\frac{dv_C}{dt} = -\frac{1}{R_2 C} v_C \tag{4.6}$$

Equation 4.6 indicates that the rate of change of the voltage across the capacitor is proportional to the voltage across the capacitor itself. That is, at $t = 0^+$ the capacitor is *discharging* at its maximum rate because v_C itself and, thus, i_{R_2} are both maximums at that moment. As the capacitor continues to discharge, v_C and i_{R_2} continue to decrease such that the *rate* of decrease in v_C decreases as well.

MAKE THE CONNECTION

Thermal Capacitance

A hydraulic capacitor can store energy in its fluid in much the same way an electric capacitor stores energy in its charge. (See the Make the Connection sidebar, "Fluid Capacitance" in Chapter 3.) The thermal capacitance C_t of an object is related to two physical properties: mass and specific heat:

$$C_t = mc; m = \text{mass [kg]}$$
$$c = \text{specific heat [J/°C-kg]}$$

Physically, thermal capacitance is related to the ability of a mass to store heat and describes how much the temperature of the mass will rise for a given addition of heat. If we add heat at the rate q for time Δt and the resulting temperature rise is ΔT, then we can define the thermal capacitance to be

$$C_t = \frac{\text{heat added}}{\text{temperature rise}} = \frac{q\Delta t}{\Delta T}$$

If the temperature rises from value T_0 at time t_0 to T_1 at time t_1, then we can write

$$T_1 - T_0 = \frac{1}{C_t}\int_{t_0}^{t_1} q(t)dt$$

or, in differential form,

$$C_t \frac{dT(t)}{dt} = q(t)$$

The graph of Figure 4.6 shows the normalized transient response of i_{R_2}. One can easily check that the slope at any point on the curve is proportional to the value at the same point. This type of relationship wherein the rate of change of a variable is proportional to the value of the variable itself is the fundamental quality of the exponential function. Thus, the transient response of the R_2C series loop shown in Figure 4.6 is characterized by:

$$\frac{dv_C}{dt} \propto v_C(t) \propto e^{-t/\tau} \qquad \text{where} \qquad \tau = R_2C \tag{4.7}$$

The parameter τ is known as a *time constant.* Such decaying exponentials, whether rising or falling, are ubiquitous in the mathematical representations of transient responses of physical systems.

Notice that the normalized transient response of i_{R_2} is shown in Figure 4.6 up to $t = 5\tau$, at which time v_C and i_{R_2} have undergone over 99 percent of their change from the old to the new steady states. For most practical purposes, the capacitor can be considered fully discharged for $t \leq 5\tau$.

Now consider what happens if the switches in the circuit of Figure 4.6 are returned to their original positions at some moment after $t = 5\tau$. The capacitor will then be disconnected from R_2 and reconnected in a series loop with the battery V_B and the resistor R_1. At that moment, the capacitor is now *charging* at its maximum rate because the voltage $(V_B - v_C)$ across R_1 and, thus, i_{R_1} are maximums. As the capacitor continues to charge, $(V_B - v_C)$ and i_{R_1} continue to decrease such that the *rate* of increase in v_C decreases as well. The result is another decaying, but rising, exponential, such as that shown on the right in Figure 4.4. The time constant τ for the V_BR_1C series loop is R_1C.

These fundamental behaviors also occur for first-order circuits containing an inductor and one or more resistors and independent sources.

It is also worth noting that the currents i_{R_1} and i_{R_2} in this illustrative example were discontinuous across the transient events. As emphasized earlier in this section, only

state variables (e.g., v_C) are guaranteed to be continuous across a transient event.

Finally, for circuits with two storage elements it is possible that those elements will exchange energy back and forth with each other during the transient response. When this phenomenon occurs, the result is oscillating voltages and currents in the circuit even as the magnitudes of the oscillations decay exponentially over time.

Page 264

Time Constants

First-order circuits have one time constant τ, which is a measure of the speed of response of the circuit to a transient event. A small or large time constant indicates a fast or slow response, respectively. The time constant τ of a first-order circuit is either:

$$R_T C \quad \text{or} \quad \frac{L}{R_N} \tag{4.8}$$

depending upon whether the storage element is a capacitor or an inductor. Here, R_T and R_N are the Thévenin and Norton equivalent resistances seen by the capacitor and inductor, respectively.

Figure 4.7 shows a typical first-order decaying exponential. The time constant τ can be found graphically by two methods. The simplest and most common method is to determine τ as the time required for the exponential curve to decay $(e - 1)/e$ (or approximately 63 percent) of the difference between its initial value $x(0)$ and its long-term steady state $x(\infty)$. An alternate method is to determine τ as the time marked by the intersection of the tangent to the exponential curve at $t = 0$ and the horizontal asymptote $x(\infty)$.

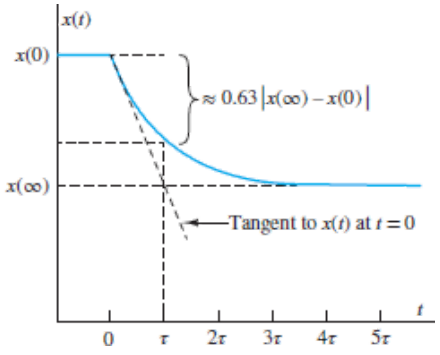


Figure 4.7 Generic first-order response $x(t)$ suggesting two graphical methods for finding a time constant

Second-order circuits essentially have two time constants, which are commonly related to and expressed as two parameters known as the *dimensionless damping ratio* ζ and the *natural frequency* ω_n.

MAKE THE CONNECTION

Thermal System Dynamics

To describe the dynamics of a thermal system, we write a differential equation based on energy balance. The difference between the heat added to the mass by an external source and the heat leaving the same mass (by convection or conduction) must be equal to the heat stored in the mass:

$$q_{\text{in}} - q_{\text{out}} = q_{\text{stored}}$$

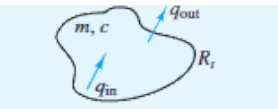
An object is internally heated at the rate q_{in} in ambient temperature $T = T_a$; the thermal capacitance and thermal resistance are C_t and R_t. From energy balance:

$$q_{\text{in}}(t) - \frac{T(t) - T_a}{R_t} = C_t \frac{dT(t)}{dt}$$

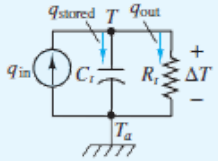
$$R_t C_t \frac{dT(t)}{dt} + T(t) = R_t q_{\text{in}}(t) + T_a$$

$$\tau_t = R_t C_t \quad K_{St} = R_t$$

This first-order system is identical in its form to an electric *RC* circuit, as shown below.



Thermal system



Equivalent electric circuit

Long-Term Steady State

The long-term steady state is that which remains after the transient response has decayed completely. For the first-order decaying exponential shown in Figure 4.7 the long-term steady state is $x(\infty)$. The long-term steady state typically depends upon the independent sources present in the $t > 0$ circuit. If all of those sources are DC the long term steady-state will also be DC, where capacitors act as open-circuits and inductors act as short-circuits.

Complete Response

The complete response is simply the sum of the transient response and the long-term steady state. In general, the transient response will contain one unknown constant for each state variable in the circuit. Thus, the complete Page 265response will also contain the same number of unknown constants. The values of these unknown constants are determined by the initial conditions on the circuit at $t = 0^+$.

A common mistake when learning to solve transient circuit problems is to apply the initial conditions to the transient response alone rather than to the complete solution. Forewarned, forearmed; don't make this mistake!

Natural and Forced Responses

Often, it is useful to express the complete response as the sum of *natural* and *forced* responses instead of the sum of a transient response and long-term steady state. Either way the complete response is unchanged. The natural response is that part of the complete system response due to the initial energy stored in the system at $t = 0$. The forced response is that part due to independent sources present in the $t > 0$ circuit.

Equation 4.9 expresses the complete response $x(t)$ of an arbitrary first-order circuit variable as the sum of a transient response, with its characteristic exponential decay, and a long-term steady state $x(\infty)$.

$$x(t) = [x(0^+) - x(\infty)]e^{-t/\tau} + x(\infty) \tag{4.9}$$

The transient response portion includes the difference between the initial condition $x(0^+)$ and the long-term steady state. This expression can be reconstructed as:

$$x(t) = x_N(t) + x_F(t) = x(0^+)e^{-t/\tau} + x(\infty)(1 - e^{-t/\tau}) \tag{4.10}$$

The first and second terms in equation 4.10 are known as the *natural* and *forced* responses, $x_N(t)$ and $x_F(t)$, respectively. A similar construction can be made for the complete response of a second-order circuit.

EXAMPLE 4.1 Initial Conditions

Problem

For the circuit shown in Figure 4.8(a), determine the current through the inductor just before the switch is opened.

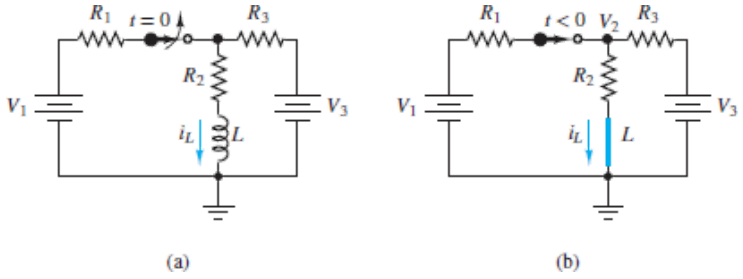


Figure 4.8 (a) Circuit for Example 4.1; (b) the same circuit just before the switch is opened

Page 266

Solution

Known Quantities: $R_1 = 1$ kΩ; $R_2 = 5$ kΩ; $R_3 = 3.33$ kΩ; $L = 0.1$ H; $V_1 = 12$ V; $V_3 = 4$ V.

Find: The current i_L through the inductor.

Assumptions: Assume the switch has been closed for a long time prior to $t = 0$ such that the circuit is in a DC steady-state.

Analysis: For $t < 0$, the circuit is in a DC steady-state condition, and the inductor acts as a short-circuit, as shown in Figure 4.8(b). The current i_L through the inductor can be found quickly by applying KCL at node V_2:

$$\frac{V_2 - V_1}{R_1} + \frac{V_2 - 0}{R_2} + \frac{V_2 - V_3}{R_3} = 0$$

Collect the coefficients of V_1, V_2, and V_3 to find:

$$\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) V_2 - \frac{V_1}{R_1} - \frac{V_3}{R_3} = 0$$

Finally, rearrange the terms to find:

$$V_2 = \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)^{-1} \left(\frac{V_1}{R_1} + \frac{V_3}{R_3}\right) = 8.80 \text{ V}$$

To determine the current through the inductor, observe that

$$i_L(0) = \frac{V_2}{R_2} = \frac{8.80}{5{,}000} = 1.76 \text{ mA}$$

Comments: The current $i_L(0)$ is *the* initial condition for the $t > 0$ circuit behavior. Only the state variables (i.e., the current through an inductor and the voltage across a capacitor) are guaranteed to be continuous across a transient event, such as the opening or closing of a switch.

EXAMPLE 4.2 Continuity of Inductor Current and Capacitor Voltage Problem

Find the initial conditions at $t = 0$ on the current through the inductor and the voltage across the capacitor in the circuit in Figure 4.9.

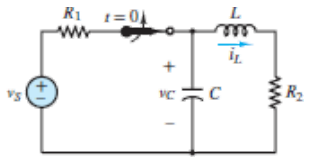


Figure 4.9

Solution

Known Quantities: v_S; R_1; R_2; L; C

Find: The current through the inductor and the voltage across the capacitor at $t = 0^+$.

Assumptions: The switch has been closed for a very long time prior to $t = 0$.

Analysis: In a DC steady-state, the inductor acts as a short-circuit and the capacitor acts as an open-circuit. Then, the circuit is effectively a single loop with a current i equal to the inductor short-circuit current and given by:

$$i = i_L = \frac{v_S}{R_1 + R_2} \qquad t < 0$$

Page 267

The voltage across the capacitor open-circuit is given by voltage division.

$$v_C = v_S \frac{R_2}{R_1 + R_2} \qquad t < 0$$

Since neither the current through an inductor nor the voltage across a capacitor can change instantaneously, the initial conditions on the inductor current and capacitor voltage are

$$i_L(t = 0^+) = i_L(t = 0^-) = \frac{v_S}{R_1 + R_2}$$

$$v_C(t = 0^+) = v_C(t = 0^-) = v_S \frac{R_2}{R_1 + R_2}$$

LO

EXAMPLE 4.3 Continuity of Inductor Current

Problem

Find the initial condition and final value of the inductor current in the circuit in Figure 4.10.

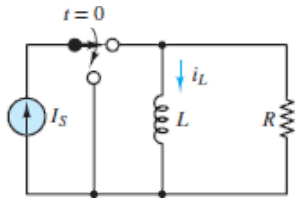


Figure 4.10

Solution

Known Quantities: Source current I_S; inductor and resistor values.

Find: Inductor current at $t = 0^+$ and as $t \to \infty$.

Schematics, Diagrams, Circuits, and Given Data: $I_S = 10$ mA.

Assumptions: The current source has been connected to the circuit for a very long time.

Analysis: For $t < 0$, the inductor acts as a short-circuit. Thus, the voltage across and the current through resistor R are zero such that all of the current I_S is passing through the inductor. At $t = 0^+$, the switch opens and since the inductor current must be continuous

$$i_L(0^+) = i_L(0^-) = I_S$$

For $t > 0$, the current source is in its own isolated loop, cut off from the inductor and resistor. The inductor and resistor are in series in a separate isolated loop. Since this loop has no source, the loop current will eventually decay to zero (the long-term steady state) due to the energy dissipation of the resistor. A qualitative sketch of the current as a function of time is shown in Figure 4.11.

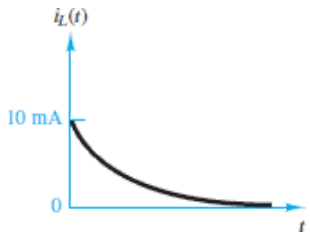


Figure 4.11

Comments: The direction of the current through R for $t > 0$ is determined by the initial condition on the inductor current.

EXAMPLE 4.4 Long-Term DC Steady-State

Problem

Determine the capacitor voltage in the circuit in Figure 4.12(a) a long time after the switch has been closed.

Page 268

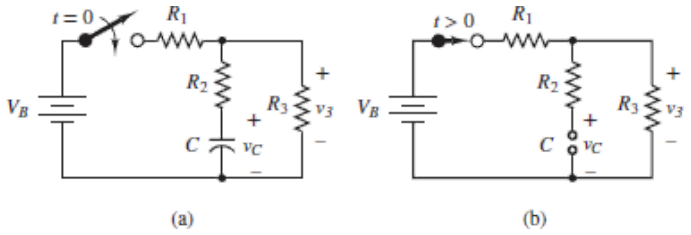


Figure 4.12 (a) Circuit for Example 4.4; (b) same circuit a long time after the switch is closed

Solution

Known Quantities: The values of the circuit elements are $R_1 = 100\ \Omega$; $R_2 = 75\ \Omega$; $R_3 = 250\ \Omega$; $C = 1\ \mu$F; $V_B = 12$ V.

Analysis: After the switch has been closed for a long time ($t \to \infty$), any transient response has decayed away and the circuit has reached a new DC steady-state. In a DC state the capacitor acts as an open-circuit, as shown in Figure 4.12(b). As a result, no current is through resistor R_2, and so resistors R_1 and R_3 are in a *virtual* series connection. Apply voltage division to find:

$$v_3(\infty) = \frac{R_3}{R_1 + R_3} V_B = \frac{250}{350}(12) = 8.57 \text{ V}$$

Since the current through R_2 is zero, the voltage across R_2 is also zero. Then, v_C equals the voltage drop from the upper right node to the bottom node, which is, of course, also equal to v_3. Thus:

$$v_C(\infty) = v_3(\infty) = 8.57 \text{ V}$$

Comments: The voltage $v_C(\infty)$ is the DC long-term steady-state voltage across the capacitor.

EXAMPLE 4.5 Writing the Differential Equation of an *RC* Circuit

Problem

Derive a differential equation for the voltage across the capacitor shown in Figure 4.13.

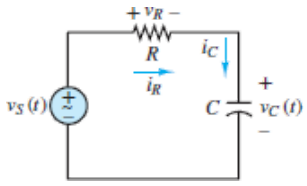


Figure 4.13

Solution

Known Quantities: R; C; $v_S(t)$.

Find: The differential equations in $v_C(t)$ and $i(t)$.

Assumptions: None.

Analysis: Apply KVL around the loop to obtain:

$$v_S - i_R R - v_C = 0$$

Use KCL and the *i-v* relationship for a capacitor

$$i_R = i_C = C\frac{dv_C}{dt}$$

Page 269

to substitute for i_R.

$$v_S - RC\frac{dv_C}{dt} - v_C = 0$$

Rearranging terms, the result is

$$RC\frac{dv_C}{dt} + v_C = v_S$$

Since all of the terms in the sum must have the same dimensions, we can infer that the dimension of RC is time! It is worth mentioning that for more complicated circuits the R in RC is the Thévenin equivalent resistance "seen" by the capacitor.

A differential equation in the current i_R can also be found by differentiating both sides of the KVL equation above to obtain:

$$\frac{dv_S}{dt} - R\frac{di_R}{dt} - \frac{dv_C}{dt} = 0$$

Again, use the i-v relationship for a capacitor to substitute for the derivative of v_C to obtain:

$$\frac{dv_S}{dt} - R\frac{di_R}{dt} - \frac{i_C}{C} = 0$$

Recall that $i_R = i_C$ and multiply both sides of the equation by C and rearrange to obtain:

$$RC\frac{di_R}{dt} + i_R = C\frac{dv_S}{dt}$$

Keep in mind that both the current i_R and the voltage v_C are functions of time as they transition from the old to the new steady state.

Notice that the left hand sides of the differential equations for v_C and i_R are identical. In general, the left hand side of the differential equation for any variable in a circuit is characteristic of every variable in that circuit.

Comments: First-order RC circuits have one state variable, v_C, the voltage across the capacitor.

EXAMPLE 4.6 Writing the Differential Equation of an *RL* Circuit

Problem

Derive differential equations from the circuit shown in Figure 4.14.

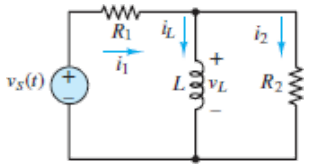


Figure 4.14

Solution

Known Quantities: $R_1 = 10\ \Omega$; $R_2 = 5\ \Omega$; $L = 0.4$ H.

Find: The differential equations for i_L and v_L.

Assumptions: None.

Analysis: Apply KCL at the top right node to obtain:

$$i_1 - i_L - i_2 = 0$$

Page 270

Apply KVL around the left mesh to obtain:

$$v_S - i_1 R_1 - v_L = 0 \quad \text{or} \quad i_1 = \frac{v_S - v_L}{R_1}$$

Apply Ohm's law to R_2 to obtain:

$$v_L = i_2 R_2 \quad \text{or} \quad i_2 = \frac{v_L}{R_2}$$

Use these expressions to substitute for i_1 and i_2 in the KCL equation to find:

$$\frac{v_S - v_L}{R_1} - i_L - \frac{v_L}{R_2} = 0$$

Finally, use the differential *i*-*v* relationship for an inductor

$$v_L = L\frac{di_L}{dt}$$

to substitute for v_L. The result is

$$\frac{1}{R_1}\left[v_S - L\frac{di_L}{dt}\right] - i_L - \frac{L}{R_2}\frac{di_L}{dt} = 0$$

Collect terms to find:

$$L\frac{R_1 + R_2}{R_1 R_2}\frac{di_L}{dt} + i_L = \frac{v_S}{R_1}$$

Notice that the coefficient of the first derivative term is

$$\frac{L}{R_N} \quad \text{where} \quad R_N = \frac{R_1 R_2}{R_1 + R_2}$$

such that

$$\frac{L}{R_N}\frac{di_L}{dt} + i_L = \frac{v_S}{R_1}$$

where R_N is the Norton equivalent resistance seen by the inductor. Notice that the first term in the differential equation has dimensions of current per time. Since all of the terms in the sum must have the same dimensions, we can infer that the dimension of *L/RN* is time!

Substitute numerical values to obtain:

$$0.12\frac{di_L}{dt} + i_L = 0.1\,v_S$$

To obtain a differential equation for v_L use the KCL equation $i_L = i_1 - i_2$ and the expressions for i_1 and i_2 to substitute for i_L in the differential *i*-*v* relationship to obtain:

$$v_L = L\frac{d}{dt}\left[\frac{v_S - v_L}{R_1} - \frac{v_L}{R_2}\right]$$

Rearrange terms to find:

$$\frac{L}{R_N}\frac{dv_L}{dt} + v_L = \frac{L}{R_1}\frac{dv_S}{dt}$$

Notice that the left side of the differential equation is the same as that for i_L. This result is true for every variable in the circuit. The left side of the differential equation is characteristic of the entire circuit rather than just any one variable.

Comments: First-order *RL* circuits have one state variable, i_L, the current through the inductor.

CHECK YOUR UNDERSTANDING

The single-pole, single-throw (SPST) switch in part (a) of Example 4.1 is opened at $t = 0$. What is the inductor current after a long time has passed?

Answer: $i_L(\infty) = \frac{V_2}{R_2 + R_3} = 0.48 \text{ mA}$

CHECK YOUR UNDERSTANDING

Use the principle of superposition to find the initial condition $i_L(t = 0^+)$ in Example 4.1.

Answer: $i_L(t_0^+) = i_L(t_0^-) = \frac{V_1}{R_2}\frac{R_2 \| R_3}{R_2 R_2 \| R_3 + R_1} + \frac{V_2}{R_2}\frac{R_1 \| R_2}{R_2 R_1 \| R_2 + R_3} = 1.76 \text{ mA}$

CHECK YOUR UNDERSTANDING

The single-pole, double-throw (SPDT) switch in the circuit of Example 4.3 is thrown at $t = 0$. Suppose that after a long time $t = t_\infty$ the switch is thrown again, back to its original position. What is the initial current through the inductor at $t = t_\infty$? What is the eventual long-term steady state current through the inductor for $t > t_\infty$?

Answer: $i_L(t_\infty) = 0$; i_L at $t_\infty = I_S = 10$ mA

CHECK YOUR UNDERSTANDING

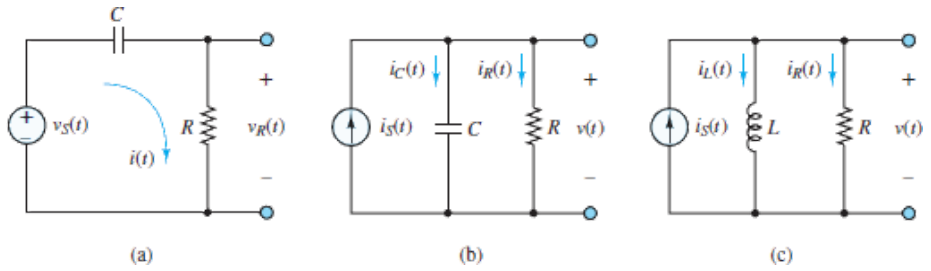
Suppose that the single-pole, single-throw (SPST) switch in part (b) of Example 4.4 is eventually opened again. What is the capacitor voltage after an additional long time has passed?

Answer: $v_C(t \to \infty) = 0$ V. The capacitor will discharge through R_2 and R_3.

Page 272

CHECK YOUR UNDERSTANDING

Use the differential *i-v* relations for capacitors and inductors along with KVL or KCL to write the differential equation for each of the circuits shown below.



Answer: (a) $RC\dfrac{dv_C(t)}{dt} + v_C(t) = v_S(t)$; (b) $RC\dfrac{dv(t)}{dt} + v(t) = Ri_S(t)$;

(c) $\dfrac{L}{R}\dfrac{di_L(t)}{dt} + i_L(t) = i_S(t)$

CHECK YOUR UNDERSTANDING

Apply KVL twice to derive a differential equation for v_C for $t > 0$ in the circuit of Example 4.5.

Answer: $\frac{d^2v_C}{dt^2} + \frac{R_2}{L}\frac{dv_C}{dt} + \frac{1}{LC}v_C = 0$

MAKE THE CONNECTION

First-Order Thermal System

An automotive transmission generates heat, when engaged, at the rate $q_{in} = 2{,}125$ W. The thermal capacitance of the transmission is $C_t = mc = 12$ kJ/°C. The effective convection resistance through which heat is dissipated is $R_t = 0.04$°C/W.

1. What is the steady-state temperature the transmission will reach when the initial (ambient) temperature is 5°C?

With reference to the Make the Connection sidebar "Thermal Capacitance," we write the differential equation based on energy balance:

$$R_t C_t \frac{dT}{dt} + T = R_t q_{in}$$

At steady state, the rate of change of temperature is zero; hence, $T(\infty) = R_t q_{in}$. Using the numbers given, $T(\infty) = 0.04 \times 2{,}125 = 85$°C

2. How long will it take the transmission to reach 90 percent of the final temperature?

The general form of the solution is

$$\begin{aligned} T(t) &= [T(0) - T(\infty)]e^{-t/\tau} + T(\infty) \\ &= T(0)\,e^{-t/\tau} + T(\infty) \\ &\quad \times (1 - e^{-t/\tau}) \\ &= 5\,e^{-t/\tau} + 85(1 - e^{-t/\tau}) \end{aligned}$$

Thus, the transmission temperature starts out at 5°C and increases to its final value of 85°C, as shown in the plot in Figure 4.18.

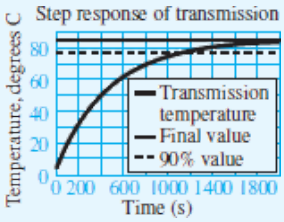


Figure 4.18 Temperature response of automotive transmission

Given the final value of 85°C, we calculate 90 percent of the final temperature to be 76.5°C. To determine the time required to reach this temperature, we solve the following equation for the argument t:

$$\begin{aligned} T(t_{90\%}) &= 76.5 \\ &= (5 - 85)\,e^{-t_{90\%}/\tau} + 85 \\ \frac{-8.5}{-80} &= e^{-t_{90\%}/\tau} \\ 0.10625 &= e^{-t_{90\%}/\tau} \\ t_{90\%} &= 2.24\tau = 1{,}076 \text{ s} \\ &= 17.9 \text{ min} \end{aligned}$$

4.3 FIRST-ORDER TRANSIENT ANALYSIS

First-order systems are important in all engineering disciplines and occur frequently in nature. Such systems are characterized by a single energy storage element and one associated state variable, where the energy of the state variable is dissipated such that the rate of change of the state variable is proportional to the state variable itself. The fundamental result is that the transient response of a first-order system is a *decaying exponential* function of time.

Ideal first-order electrical systems possess either capacitance or inductance (but not both) along with resistance and (perhaps) energy sources. Ideal first-order

mechanical systems possess mass and damping (e.g., sliding or viscous friction) but no elasticity or compliance. An ideal first-order fluid system possesses fluid capacitance and viscous dissipation, such as a hydraulic system with a liquid-filled tank and a variable orifice. Many conductive and convective thermal systems also exhibit first-order behavior.

In general, when solving transient circuit problems it is necessary to determine three elements: (1) the Page 273steady-state response prior to a transient event, (2) the transient response immediately following the transient event, and (3) the long-term steady-state response remaining after the transient response has decayed away. The steps involved in computing the complete response of a first-order circuit with *constant sources* are outlined below.

Circuit Simplification for $t > 0$

The first step to solve for the response after the transient event ($t > 0$) is to partition the circuit into a source network and load, with the energy storage element as the load, as shown in Figure 4.15. If the source network is linear, it can be replaced by its Thévenin or Norton equivalent network.

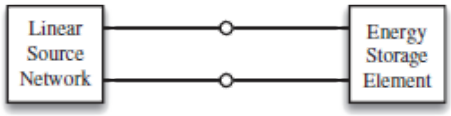


Figure 4.15 Generalized first-order circuit seen as a source network attached to an energy storage element as the load

Consider the case when the load is a capacitor and the source network is replaced by its Thévenin equivalent network, as shown in Figure 4.16. KVL can be applied around the loop to yield:

$$V_T - iR_T - v_c = 0$$

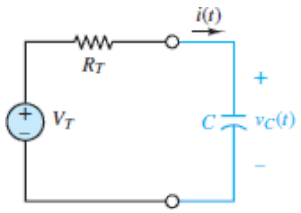


Figure 4.16 Generalized firstorder circuit with a capacitor load and a Thévenin source

Of course, $i = i_C$ and for a capacitor $i_C = C\, dv_C/dt$. After substituting and rearranging the terms, the result is

$$R_T C \frac{dv_C}{dt} + v_C = V_T \qquad \text{Capacitor load with Thévenin source} \tag{4.11}$$

For a DC source network, the long-term steady-state solution is simply $v_C = V_T$.

Likewise, consider the case when the load is an inductor and the source network is replaced by its Norton equivalent network, as shown in Figure 4.17. KCL can be applied at either node to yield:

$$I_N - \frac{v}{R_N} - i_L = 0$$

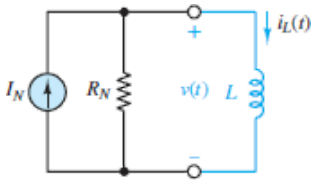


Figure 4.17 Generalized firstorder circuit with an inductor load and a Norton source

Page 274

Of course, $v = v_L$ and for an inductor $v_L = L\, di_L/dt$. After substituting and rearranging the terms, the result is:

$$\frac{L}{R_N}\frac{di_L}{dt} + i_L = I_N \qquad \text{Inductor load with Norton source} \tag{4.12}$$

For a DC source network, the long-term steady-state solution is simply $i_L = I_N$.

It is important to keep in mind that these equations are for $t > 0$, that is, the transient response. In general, the equivalent source network seen by the load after the transient event will be different from that seen by the load before the event. Equivalent network methods can be used for both domains but *do not assume* that the equivalent network seen by the load is unchanged by the event.

First-Order Differential Equation

Both equations 4.11 and 4.12 have the same general form:

$$\tau\frac{dx(t)}{dt} + x(t) = K_S f(t) \qquad \text{First-order system equation} \tag{4.13}$$

where the constants τ and K_S are the **time constant** and the **DC gain,** respectively. In this chapter, $f(t)$ is assumed equal to a constant F, which represents the contribution of one or more DC sources. With that assumption in mind, the general first-order differential equation is

$$\tau\frac{dx(t)}{dt} + x(t) = K_S F \qquad t \geq 0 \tag{4.14}$$

The solution for $x(t)$ has two parts: the *transient* response and the *long-term steady-state* response. These two parts can also be rearranged in terms of **natural** and **forced** responses. Either way, the sum of both parts is known as the **complete response**. One initial condition $x(0^+)$ is needed to specify the complete response.

MAKE THE CONNECTION

Hydraulic Tank

The analogy between electric and hydraulic circuits illustrated in earlier chapters can be applied to the hydraulic tank shown in Figure 4.20. The tank is cylindrical with cross-sectional area A, and the liquid contained in the tank exits the tank through a valve, which is modeled by a fluid resistance R. Initially, the level, or head, of the liquid is h_0. The principle of conservation of mass can be applied to the liquid in the tank to determine the rate at which the tank will empty. For mass to be conserved, the following equation must apply:

$$q_{in} - q_{out} = q_{stored}$$

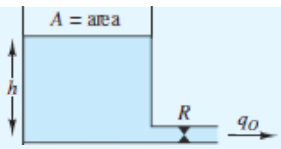


Figure 4.20 Analogy between electrical and fluid capacitance

In this equation, the variable q represents a volumetric flow rate in cubic meters per second. The flow rate into the tank is zero in this particular case, and the flow rate out is given by the pressure difference across the valve, divided by the resistance:

$$q_{out} = \frac{\Delta p}{R} = \frac{\rho g h}{R}$$

The expression $\Delta p = \rho g h$ is obtained from basic fluid mechanics: $\rho g h$ is the static pressure at the bottom of the tank, where ρ is the density of the liquid, g is the acceleration of gravity, and h is the (changing) liquid level.

The flow rate stored is related to the rate of change of the fluid volume contained in the tank (the tank stores energy in the mass of the fluid):

$$q_{stored} = A\frac{dh}{dt}$$

Thus, we can describe the emptying of the tank by means of the first-order linear ordinary differential equation

$$\begin{aligned} 0 - q_{out} &= q_{stored} \\ \Rightarrow \quad -\frac{\rho g h}{R} &= A\frac{dh}{dt} \\ \frac{RA}{\rho g}\frac{dh}{dt} + h &= 0 \\ \Rightarrow \quad \tau\frac{dh}{dt} + h &= 0 \\ \tau &= \frac{RA}{\rho g} \end{aligned}$$

We know from the content of the present section that the solution of the first-order equation with zero input and initial condition h_0 is

$$h(t) = h_0 e^{-t/\tau}$$

Thus, the tank will empty exponentially, with the time constant determined by the fluid properties, that is, by the resistance of the valve and by the area of the tank.

First-Order Transient Response

The transient response x_{tr} is found by setting $F = 0$ in equation 4.14 such that:

$$\tau \frac{dx_{tr}(t)}{dt} + x_{tr}(t) = 0 \tag{4.15}$$

The solution for x is found by assuming a solution of the form:

$$x_{tr}(t) = \alpha\, e^{st} \tag{4.16}$$

Substitution of this assumed solution into equation 4.15 results in a characteristic equation.

$$\tau s + 1 = 0 \qquad \text{Characteristic equation} \tag{4.17}$$

The solution for s is simply:

$$s = \frac{-1}{\tau} \tag{4.18}$$

Page 275

which is known as the root of the characteristic equation. Plugging in for s in equation 4.16 yields a decaying exponential.

$$x_{tr}(t) = \alpha e^{-t/\tau} \qquad \text{Transient response} \tag{4.19}$$

The constant α in equation 4.19 cannot be evaluated until the complete response has been found.

The amplitude of $x_{tr}(t)$ at $t = n\,\tau$ for $n = 0, 1, \ldots, 5$ is shown in Figure 4.19. The data show that x_{tr} has decayed by roughly 95 percent at three time constants and by over 99 percent at five time constants.

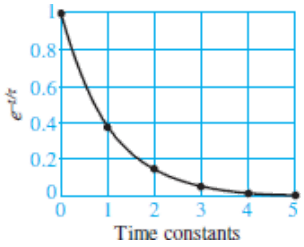


Figure 4.19 Normalized firstorder exponential decay

Long-Term Steady-State Response

Still assuming that the first-order circuit contains only DC sources, such that $f(t)$ is a constant F, the long-term steady-state response of a first-order system is the solution to:

$$\tau \frac{dx_{ss}(t)}{dt} + x_{ss}(t) = K_S F \qquad t \geq 0 \tag{4.20}$$

For constant F, $x_{ss} = K_S F$ is the solution. Thus:

$$\boxed{x_{ss}(t) \equiv x(\infty) = K_S F \qquad F = \text{constant}} \tag{4.21}$$

Complete First-Order Response

The complete response is the sum of the transient and long-term steady-state responses:

$$x(t) = x_{tr}(t) + x_{ss}(t) = \alpha e^{-t/\tau} + x(\infty) = \alpha e^{-t/\tau} + K_S F \qquad t \geq 0 \tag{4.22}$$

Apply the one initial condition $x(0^+)$ to solve for the unknown constant α:

$$\begin{aligned} x(0^+) &= \alpha + x(\infty) \\ \alpha &= x(0^+) - x(\infty) \end{aligned} \tag{4.23}$$

Substitute for α in equation 4.22 to find the complete response:

$$\boxed{x(t) = \left[x(0^+) - x(\infty)\right] e^{-t/\tau} + x(\infty) \qquad t \geq 0} \tag{4.24}$$

FOCUS ON PROBLEM SOLVING

FIRST-ORDER TRANSIENT CIRCUIT ANALYSIS

1. Find the value of the state variable just before the transient event at $t = 0^-$. is, find $v_C(0^-)$ or $i_L(0^-)$.
2. Set the value of the state variable just after the transient event equal to the v just before it. That is, set $v_C(0^+) = v_C(0^-)$ or $i_L(0^+) = i_L(0^-)$ as the ir condition on the transient response.

   Note: Only the state variable is guaranteed to be continuous across the trans event. The initial condition on an arbitrary variable $x(t)$ *must* be found from initial condition on the state variable.
3. For $t > 0$, treat the storage element as the load and simplify the remaining so network. Assuming the source network is linear, when the storage element is
   - a capacitor, replace the source network with its Thévenin equivalent (V_T R_T), as shown in Figure 4.16.
   - an inductor, replace the source network with its Norton equivalent (I_N R_N), as shown in Figure 4.17.
4. For $t > 0$, find the governing differential equation for the state variable.
   - When the load is a capacitor, apply KVL and KCL to find:

$$\tau\frac{dv_C}{dt} + v_C = V_T \qquad \text{where} \qquad \tau = R_T C$$

   - When the load is an inductor, apply KCL and KVL to find:

$$\tau\frac{di_L}{dt} + i_L = I_N \qquad \text{where} \qquad \tau = \frac{L}{R_N}$$

5. For $t > 0$, the complete solution for the state variable is found by solving governing differential equation and applying its initial condition.
   - When the load is a capacitor, the complete solution for the state variable is

$$v_C(t) = \left[v_C(0^+) - V_T\right] e^{-t/\tau} + V_T \qquad \text{where} \qquad \tau = R_T C$$

     For an arbitrary variable $x(t)$ the complete solution is

$$x(t) = \left[x(0^+) - x(\infty)\right] e^{-t/\tau} + x(\infty)$$

   - When the load is an inductor, the complete solution for the state variable is

$$i_L(t) = \left[i_L(0^+) - I_N\right] e^{-t/\tau} + I_N \qquad \text{where} \qquad \tau = L/R_N$$

     For an arbitrary variable $x(t)$ the complete solution is

$$x(t) = [x(0^+) - x(\infty)]\, e^{-t/\tau} + x(\infty)$$

Note: The left side of the governing differential equation for an arbitrary variable is the same as that for the state variable. The right side of the governing differe equation for an arbitrary variable $x(t)$ is simply the long-term DC steady-state v for $x(t)$. It is important to observe that the time constant is the same for all varia in a network; that is, it is a network characteristic.

EXAMPLE 4.7 Charging a Camera Flash—Capacitor Energy and Time Constants

Problem

A capacitor is used to store energy in a camera flash light. The camera operates on a 6-V battery as depicted in Figure 4.21. Determine the time required for the energy stored to reach 90 percent of its maximum. Compute the time in seconds and as a multiple of the time constant.

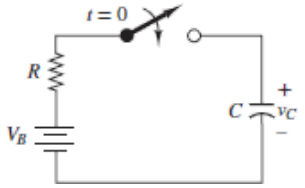


Figure 4.21 Equivalent circuit of camera flash charging circuit

Solution

Known Quantities: V_B; R; C.

Find: Time required to reach 90 percent of the total energy storage.

Schematics, Diagrams, Circuits, and Given Data: Figure 4.21; $V_B = 6$ V; $C = 1{,}000$ μF; $R = 1$ kΩ.

Assumptions: The capacitor is completely discharged prior to $t = 0$.

Analysis: In the long-term steady state ($t \to \infty$) the voltage across the capacitor equals V and the maximum possible energy stored in the capacitor is

$$E_{\text{total}} = \tfrac{1}{2} C v_C^2 = \tfrac{1}{2} C V_B^2 = 18 \times 10^{-3} \text{ J}$$

Thus, 90 percent of the maximum possible energy is

$$E_{90\%} = 0.9 \times 18 \times 10^{-3} = 16.2 \times 10^{-3} \text{ J}.$$

The corresponding capacitor voltage is calculated as follows:

$$\tfrac{1}{2} C v_C^2 = 16.2 \times 10^{-3}$$

$$v_C = \sqrt{\frac{2 \times 16.2 \times 10^{-3}}{C}} = 5.692 \text{ V}$$

The Thévenin equivalent resistance seen by the capacitor for $t > 0$ is simply R, and, thus, the time constant of the circuit is $\tau = R_T C = 10^3 \times 10^{-3} = 1$ s. The Thévenin (open-circuit)voltage is $V_T = V_B = 6$ V and the initial condition on the capacitor voltage is $v_C(0^+) = 0$ V. Refer to the Focus on Problem Solving box "First-Order Transient Circuit Analysis" to find that the complete solution for v_C is

$$v_C = 6(1 - e^{-t/\tau}) = 6(1 - e^{-t})$$

The time required to reach 90 percent of the energy is found by solving for time t when $v_C = 5.692$ V. Thus,

$$5.692 = 6(1 - e^{-t})$$

$$0.949 = 1 - e^{-t}$$

$$0.051 = e^{-t}$$

$$t = -\ln 0.051 = 2.97 \text{ s}$$

which is approximately 3τ.

Comments: The fact that the capacitor charges to 90 percent of its total energy in a period of roughly 3τ is not limited to this example. All first-order systems have the same functional form and, therefore, have the same result. What percentage of the voltage change has occurred in this same 3τ period? How many time constants are required for the voltage to reach 99 percent of its ultimate value? (See Figure 4.19.) Answers: 95 percent and 4.6τ

EXAMPLE 4.8 Simplifying a First-Order Transient Circuit

Problem

Determine a symbolic solution for the first-order circuit shown in Figure 4.22.

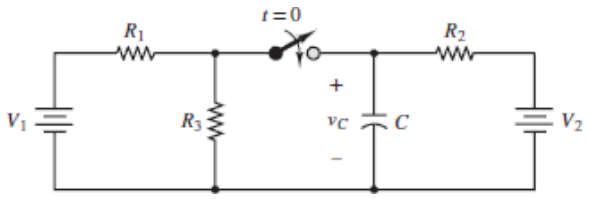


Figure 4.22

Solution

Known Quantities: V_1; V_2; R_1; R_2; R_3; C.

Find: The state variable $v_C(t)$ as a function of time for all t.

Schematics, Diagrams, Circuits, and Given Data: Figure 4.22.

Assumptions: Assume the switch was open for a very long time prior to closing, such that the circuit is in a DC steady-state prior to the transient event at $t = 0$.

Analysis:

Step 1: Find v_C for $t < 0$. For $t < 0$, the switch is open and the circuit is in a DC steady-state such that the capacitor acts as an open-circuit. Thus, there is no current through R_2 and its voltage drop is zero. Consequently, the voltage across the capacitor is V_2, as required by KVL.

$$v_C(t) = V_2 \qquad t < 0$$

Remember that it is always necessary to solve for the value of the state variable prior to the transient event even if the state variable is not the variable of ultimate interest.

Step 2: Find the initial condition on v_C. Since the voltage across a capacitor is always continuous, the initial condition on v_C at $t = 0$ is V_2.

$$v_C(0^+) = v_C(0^-) = V_2 \qquad \text{Continuity of capacitor voltage}$$

Step 3: Simplify the circuit for $t > 0$. After the switch is closed, the resulting circuit is as shown in Figure 4.23, which was redrawn to emphasize the two Thévenin sources (V_1, R_1) and (V_2, R_2) present. The approach is to select the capacitor as the load and simplify the rest of the network to its Thévenin equivalent network.

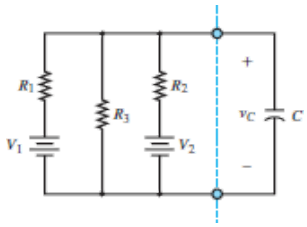


Figure 4.23 The circuit in Figure 4.22 for $t > 0$

Each Thévenin source in Figure 4.23 can be transformed to its equivalent Norton source as shown in Figure 4.24. The result is a network of resistors and independent current sources all in parallel. The current sources are combined (summed) to a single equivalent current source, and the resistors are replaced by a single equivalent resistance R_T. The resulting Norton source is then transformed to a Thévenin source. The final result is shown in Figure 4.25, where

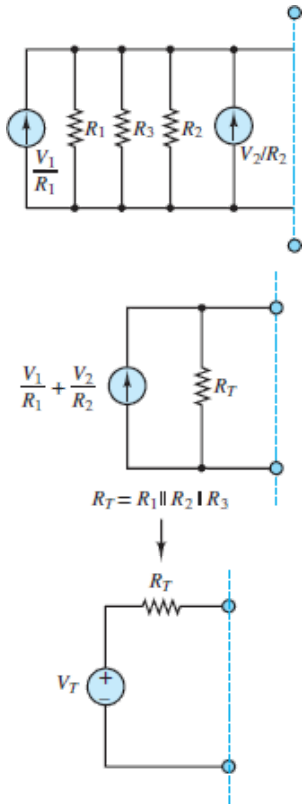


Figure 4.24 Simplification of the source network in Figure 4.23 to its Thévenin equivalent

$$R_T = R_1 \| R_2 \| R_3$$

$$V_T = \left(\frac{V_1}{R_1} + \frac{V_2}{R_2}\right) R_T$$

Page 279

Step 4: Find the differential equation. Apply KVL around the loop in Figure 4.25 to yield the differential equation for $t > 0$:

$$V_T - iR_T - v_C = V_T - R_T C\frac{dv_C}{dt} - v_C = 0 \qquad t > 0$$

$$R_T C\frac{dv_C}{dt} + v_C = V_T \qquad t > 0$$

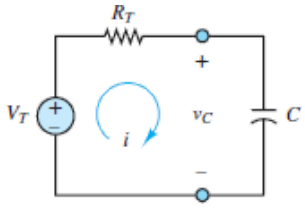


Figure 4.25 The circuit in Figure 4.23 simplified using Thévenin's theorem for $t > 0$

Step 5: Find the transient solution. The transient solution is found by setting the right side of the differential equation to zero and solving for v_C. The solution is always

$$(v_C)_{tr} = \alpha\, e^{-t/\tau}$$

The time constant associated with this first-order differential equation is $\tau = R_T C$. It is important to note that the unknown constant α is found by applying the initial condition to the *complete* solution, not to the transient solution alone.

Step 6: Find the long-term steady-state solution. The long-term DC steady-state solution for v_C is found after the switch has been closed for a very long time (practically $t \leq 5\tau$). The capacitor acts like an open-circuit such that $(v_C)_{ss} \equiv v_C(\infty) = V_T$.

Step 7: Complete solution. The complete solution is the sum of the transient and long-term steady-state solutions.

$$v_C(t) = (v_C)_{tr} + (v_C)_{ss} = \alpha\, e^{-t/\tau} + V_T$$

The unknown constant α is found by applying the initial condition $v_C(0^+) = V_2$. The result is

$$V_2 = v_C(0^+) = \alpha + V_T \qquad \text{or} \qquad \alpha = V_2 - V_T$$

Finally, the complete solution is

$$v_C(t) = (V_2 - V_T)e^{-t/R_TC} + V_T$$

EXAMPLE 4.9 Starting Transient of DC Motor

Problem

A DC motor can be modeled approximately as an equivalent first-order series *RL* circuit, as shown in Figure 4.26. Find the complete solution for i_L.

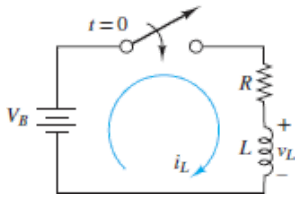


Figure 4.26 Circuit for Example 4.9

Solution

Known Quantities: Battery voltage V_B: resistance R: and inductance L.

Find: The state variable $i_L(t)$.

Schematics, Diagrams, Circuits, and Given Data: $R = 4\ \Omega$; $L = 0.1$ H; $V_B = 50$ V. Figure 4.26.

Assumptions: None.

Analysis:

Step 1: Find v_C for $t < 0$. The current through the inductor prior to the closing of the switch must be zero because the switch is open; thus,

$$i_L(t) = 0 \text{ A} \qquad t < 0$$

Page 280

Page 281

Step 2: Find the initial condition on i_L. Since the current through an inductor is always continuous, the initial condition on i_L at $t = 0$ is 0.

$$i_L(0^+) = i_L(0^-) = 0 \qquad \text{Continuity of inductor current}$$

Step 3: Simplify the circuit for $t > 0$. For $t > 0$, the network attached to the inductor is already in the form of a Thévenin source, so no further simplification is possible.

Step 4: Find the differential equation. Apply KVL around the loop in Figure 4.26 to find the differential equation for $t > 0$:

$$V_B - i_L R - v_L = 0 \qquad t > 0$$

$$V_B - i_L R - L\frac{di_L}{dt} = 0$$

Divide both sides of the equation by R to find:

$$\frac{L}{R}\frac{di_L(t)}{dt} + i_L(t) = \frac{1}{R}V_B \qquad t > 0$$

The time constant τ is the coefficient of the first-derivative term:

$$\tau = \frac{L}{R} = 0.025 \text{ s}$$

Step 5: Find the transient solution. The transient solution is found by setting the right side of the differential equation to zero and solving for i_L. The solution is always of the form:

$$(i_L)_{tr} = \alpha\, e^{-t/\tau} = \alpha\, e^{-t/0.025} = \alpha\, e^{-40t}$$

It is important to note that the unknown constant α is found by applying the initial condition to the *complete* solution, not to the transient solution alone.

Step 6: Find the long-term steady-state solution. The long-term DC steady-state solution for i_L is found after the switch has been closed for a very long time (practically $t \leq 5\tau$). In this state, the inductor acts like a short-circuit such that $(i_L)_{ss} \equiv i_L(\infty) = V_B/R = 12.5$ A.

Step 7: Complete solution. The complete solution is the sum of the *transient* and *long-term steady-state* solutions.

$$i_L(t) = (i_L)_{tr} + (i_L)_{ss} = \alpha\, e^{-40t} + 12.5 \text{ A}$$

The unknown constant α is found by applying the initial condition $i_L(0^+) = 0$. The result is:

$$0 = i_L(0^+) = \alpha + 12.5 \text{ A} \qquad \text{or} \qquad \alpha = 0 - 12.5 \text{ A} = -12.5 \text{ A}$$

Finally, the complete solution is

$$i_L(t) = -12.5e^{-t/0.47} + 12.5 \text{ A}$$

The complete solution can also be expressed in terms of *natural* and *forced* responses:

$$i_L(t) = i_{LN}(t) + i_{LF}(t) = 0 + 12.5(1 - e^{-40t})$$

Page 282

The complete response and its decomposition into (a) transient plus steady-state responses, and (b) natural plus forced responses are shown in Figure 4.27.

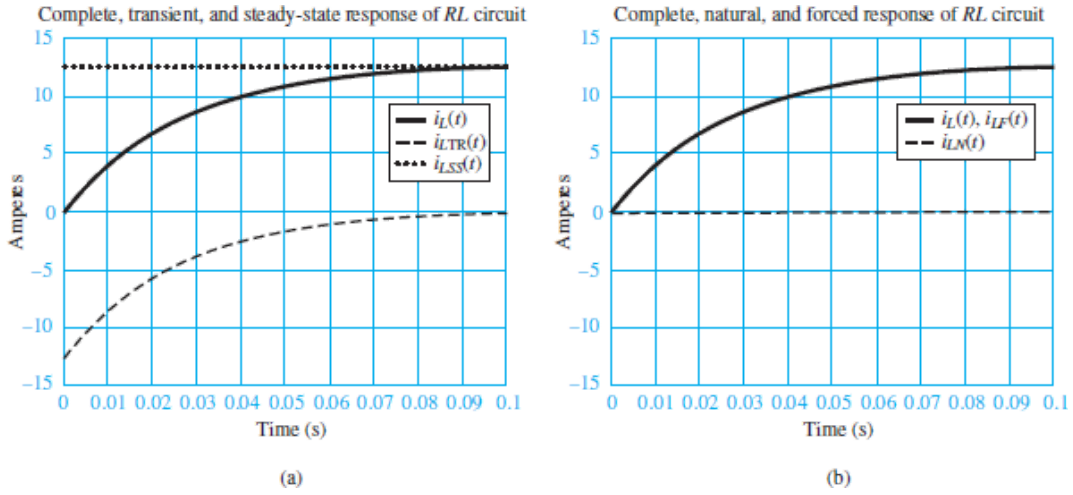


Figure 4.27 Complete response $i_L(t)$ of electric motor: (a) steady-state $i_{LSS}(t)$ plus transient $i_{LTR}(t)$ responses; (b) forced $i_{LF}(t)$ plus natural $i_{LN}(t)$ responses

Comments: As the switch opens, the inductor current is forced to change suddenly, with the result that di_L/dt, and therefore $v_L(t)$, gets very large. The large voltage transient resulting from this *inductive kick* can damage circuit components. A so-called *freewheeling diode* is used to solve this problem.

EXAMPLE 4.10 Turnoff Transient of DC Motor

Problem

Determine the motor voltage for all time in the simplified electric motor circuit model shown in Figure 4.28. The motor is represented by the series *RL* circuit in the shaded box. R_S is known as a *shunt resistor.*

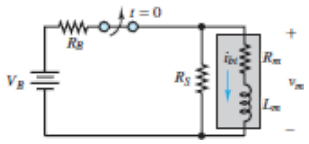


Figure 4.28

Solution

Known Quantities: V_B; R_B; R_S; R_m; L.

Find: The voltage across the motor as a function of time.

Schematics, Diagrams, Circuits, and Given Data: $R_B = 2\ \Omega$; $R_S = 20\ \Omega$; $R_m = 0.8\ \Omega$; $L = 3.0$ H; $V_B = 100.0$ V.

Assumptions: The switch has been closed for a long time prior to $t = 0$.

Analysis: With the switch closed for $t < 0$, the inductor in the circuit in Figure 4.28 behaves as a short-circuit. The current through the motor can then be calculated by current division in the modified circuit of Figure 4.29, where the inductor has been replaced Page 283with a short-circuit and the Thévenin circuit on the left has been replaced by its Norton equivalent:

$$i_m = \frac{1/R_m}{1/R_B + 1/R_s + 1/R_m}\frac{V_B}{R_B}$$
$$= \frac{1/0.8}{1/2 + 1/20 + 1/0.8}\frac{100}{2} \approx 34.7\ \text{A}$$

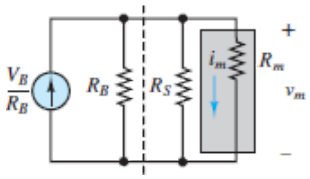


Figure 4.29

This current is the initial condition for the inductor current: $i_m(0^+) = i_m(0^-) = 34.72$ A. Since the motor inductance is effectively a short-circuit, the motor voltage for $t < 0$ is equal to

$$v_m(t) = i_m R_m = 27.8 \text{ V} \qquad t < 0$$

For $t > 0$ the switch is open and the motor sees only the *shunt* (parallel) resistance R_S, as depicted in Figure 4.30. The motor current will decay exponentially with time constant $\tau = L/(R_S + R_m) = 0.144$ s:

$$i_L(t) = i_m(t) = i_m(0^+)e^{-t/\tau} = 34.7\, e^{-t/0.144} \text{ A} \qquad t > 0$$

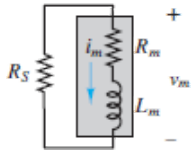


Figure 4.30

The motor voltage is then computed by adding the voltage drop across the motor resistance and inductance:

$$\begin{aligned} v_m(t) &= R_m i_m(t) + L\frac{di_m(t)}{dt} \\ &= 0.8 \times 34.7\, e^{-t/0.144} + 3\left(-\frac{34.7}{0.144}\right)e^{-t/0.144} \qquad t > 0 \\ &\approx -694\, e^{-t/0.144} \text{ V} \qquad t > 0 \end{aligned}$$

The motor voltage is plotted in Figure 4.31.

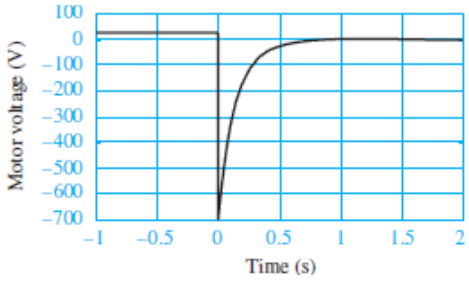


Figure 4.31 Motor voltage transient response

Comments: Notice how the motor voltage rapidly changes from the DC steady-state value of 27.8 V for $t < 0$ to a large negative value. This *inductive kick* is typical of *RL* circuits and results from the fact that although the inductor current cannot change instantaneously, the inductor voltage can and does, as it is proportional to the derivative of i_L. This example is based on a simplified representation of an electric motor but illustrates effectively the need for special starting and stopping circuits in electric motors.

Page 284

EXAMPLE 4.11 Transient Response of Ultracapacitors

Problem

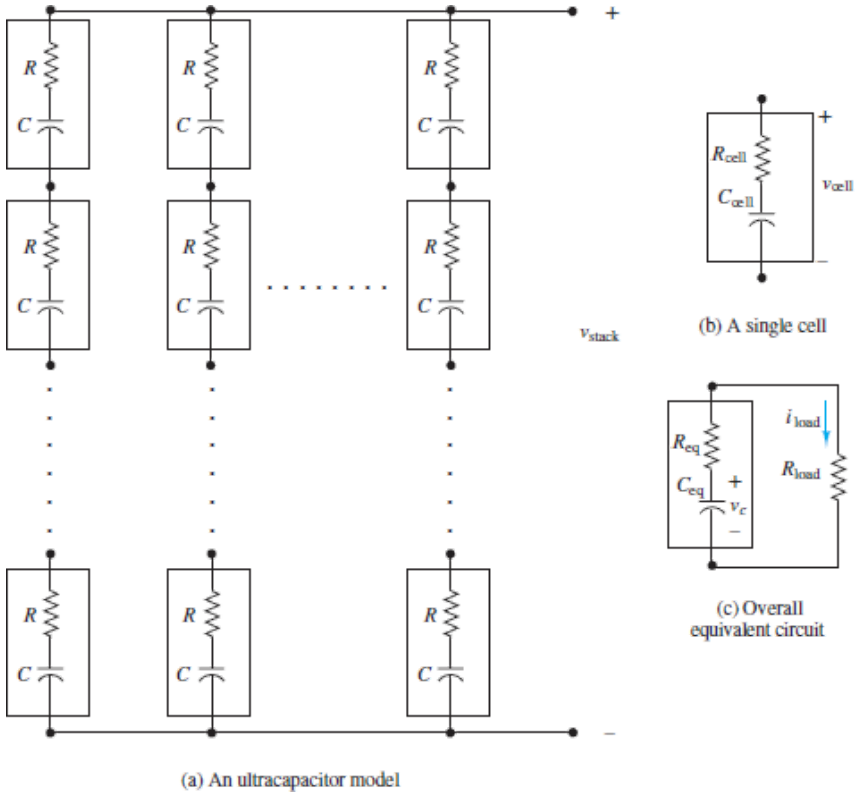
An industrial, uninterruptible power supply (UPS) is intended to provide continuous power during unexpected power outages. Ultracapacitors can store a significant amount of energy and release it during transient power outages to ensure delicate or critical electrical/electronic systems. Assume that a UPS is designed to make up for a temporary power glitch for 5 s. The system that is supported by this UPS operates at a nominal voltage of 50 V and has a maximum nominal voltage of 60 V, but it can function with a supply voltage as low as 25 V. Design a suitable UPS.

Solution

Known Quantities: Maximum, nominal, and minimum voltage; power rating and time requirements; ultracapacitor data (see Example 3.1).

Find: Number of series and parallel ultracapacitor cells needed to satisfy the specifications.

Schematics, Diagrams, Circuits, and Given Data: Figure 4.32. Capacitance of one cell: C_{cell} = 100 F; resistance of one cell: R_{cell} = 15 mΩ; nominal cell voltage V_{cell} = 2.5 V. (See Example 3.1.)



(a) An ultracapacitor model

(b) A single cell

(c) Overall equivalent circuit

Figure 4.32

Assumptions: The load can be modeled as a 0.5-Ω resistance.

Page 285

Analysis: The total capacitance of the "stack" required to satisfy the specifications is obtained by combining capacitors in series and parallel, as illustrated in Figure 4.32(a). Figure 4.32(b) depicts the electric circuit model of a single cell.

The *allowable voltage drop* in the supply is ΔV = 25 V since the load can operate with a supply as low as 25 V and nominally operates at 50 V.

The *time interval* over which the voltage will drop (but not below the allowable minimum of 25 V) is 5 s.

The ultracapacitor consists of n parallel stacks of m cells in series. Thus, the equivalent resistance of the ultracapacitor is

$$R_{eq} = mR_{cell} \| mR_{cell} \| \ldots \| mR_{cell} \qquad n \text{ times}$$
$$= \frac{m}{n} R_{cell}$$

Note that m identical capacitors C in series produce an equivalent capacitance equal to C/m and that n such capacitances in parallel produce an overall equivalent capacitance equal to nC/m. Thus:

$$C_{eq} = \frac{n}{m} C_{cell}$$

The total number of series capacitors can be calculated from the maximum required voltage:

$$m = \frac{V_{max}}{V_{cell}} = \frac{60}{2.5} = 24 \text{ series cells}$$

Define $i_C = -i_{load}$ and apply KVL to the overall equivalent circuit of Figure 4.32(c) to obtain an expression for v_C.

$$i_C(R_{eq} + R_{load}) + v_C = (R_{eq} + R_{load}) C_{eq} \frac{dv_C}{dt} + v_C = 0$$

Note that the time constant is $\tau = (R_{eq} + R_{load}) C_{eq}$. Also, assume that the ultracapacitor is fully charged prior to the power outage at $t = 0$ such that $v_C(0^+) = v_C(0^-) = 60\,\text{V}$. Since there are no independent sources in the $t > 0$ circuit the long-term DC steady-state is $v_C(\infty) = 0$. Thus, the complete solution is

$$v_C(t) = [v_C(0^+) - v_C(\infty)] e^{-t/\tau} + v_C(\infty) \qquad t \geq 0$$
$$= v_C(0^+) e^{-t/\tau} = 60 e^{-t/\tau}\ \text{V}$$

Apply voltage division to find the load voltage:

$$v_{load}(t) = \frac{R_{load}}{R_{eq} + R_{load}} v_c(t) = \frac{R_{load}}{(m/n) R_{cell} + R_{load}} v_c(t)$$
$$= \frac{0.5}{(m/n)(0.015) + 0.5} 60\, e^{-t/\tau}\ \text{V}$$

This relationship can be used to calculate the appropriate number of parallel strings n such that the load voltage is above 25 V (the minimum allowable load voltage) at $t = 5$ s. The solution could be obtained analytically, by substituting the known values $m = 24$, $R_{load} = 0.5$, $C_{cell} = 100$ F, $R_{cell} = 15$ mΩ, $t = 5$ s, and $v_{load}(t = 5)$

= 25 V and solving for n. Figure 4.33 plots the transient response of the overall equivalent circuit for n = 1 to 5. For n = 3 the requirement that v_{load} = 25 V for at least 5 s is satisfied.

Page 286

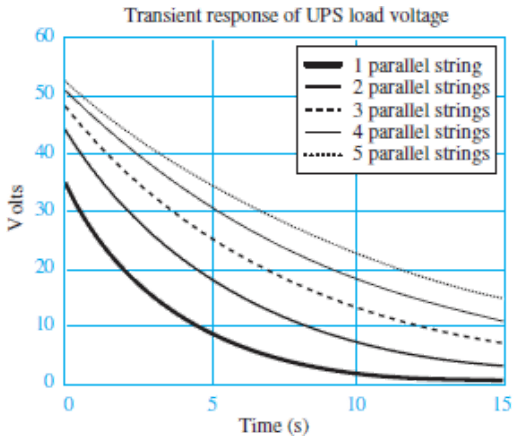


Figure 4.33 Transient response of ultracapacitor circuit

EXAMPLE 4.12 First-Order Response Due to a Pulsed Source

Problem

The circuit in Figure 4.34 includes a switch that can be used to connect and disconnect a battery. The switch has been open for a very long time. At t = 0 the switch closes, and then at t = 50 ms the switch opens again. Determine the capacitor voltage $v_C(t)$ as a function of time.

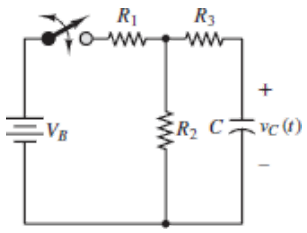


Figure 4.34

Solution

Known Quantities: V_B; R_1; R_2; R_3; C.

Find: The state variable $v_C(t)$ as a function of time for all t.

Schematics, Diagrams, Circuits, and Given Data: V_B = 15 V, $R_1 = R_2$ = 1,000 Ω, R_3 = 500 Ω, and C = 25 μF. Figure 4.34.

Assumptions: The switch has been open for a very long time for $t < 0$.

Analysis:

Part 1 ($0 \le t < 50$ ms) The switch is closed.

Step 1: DC steady-state responses. For $t < 0$ assume that the capacitor has been completely discharged through resistors R_3 and R_2 such that

$$v_C(0^-) = 0 \text{ V}$$

To determine the capacitor voltage $v_C(t)$ when the switch is closed it is necessary to calculate its long-term DC steady-state value $v_C(\infty)$ *assuming* that the switch will remain closed indefinitely. With the capacitor acting as a DC open-circuit apply voltage division to find:

$$v_C(\infty) = V_B \frac{R_2}{R_1 + R_2} = 7.5 \text{ V}$$

Step 2: Initial condition. The initial condition on $v_C(t)$ in this interval is

$$v_C(0^+) = v_C(0^-) = 0 \text{ V}$$

Step 3: Differential equation. The Thévenin equivalent resistance seen by the capacitor is $R_T = R_3 + (R_1 \| R_2) = 1$ kΩ. The Thévenin open-circuit voltage seen by the capacitor is $V_T = v_C(\infty) = 7.5$ V. Thus, the differential equation for $v_C(t)$ in this interval is

$$R_T C \frac{dv_C}{dt} + v_C = V_T \qquad 0 \le t < 50 \text{ ms}$$

Step 4: Time constant. By definition, the time constant is $\tau = R_T C = 25$msec.

Step 5: Complete solution. The complete solution is

$$\begin{aligned} v_C(t) &= v_C(\infty) + [v_C(0^+) - v_C(\infty)]\, e^{-t/\tau} \qquad 0 \le t < 50 \text{ ms} \\ &= V_T + (0 - V_T) e^{-t/R_T C} = 7.5(1 - e^{-t/0.025}) \text{ V} \qquad 0 \le t < 50 \text{ ms} \end{aligned}$$

Part 2 ($t \leq 50$ ms) The switch is open.

At $t = 50$ ms the switch opens again, and the capacitor discharges through the series combination of resistors R_3 and R_2. The independent voltage source is disconnected from the capacitor circuit so that now $R_T = R_2 + R_3$, $V_T = 0$ and the long-term DC steady-state is $v_C(\infty) = 0$. The transient for this interval begins at $t = 50$ ms and so the complete solution must be written in the form $v_C(t - t_1) = \alpha e^{-(t-t_1)/\tau}$, where $t_1 = 50$ ms.

1. The voltage v_C across the capacitor (a state variable) is continuous at $t = 50$ ms when the switch is opened.
2. The constant α is the initial condition on v_C at $t = 50$ ms.
3. The time constant for $t \leq 50$ ms is $\tau = (R_2 + R_3)C = 37.5$ ms.

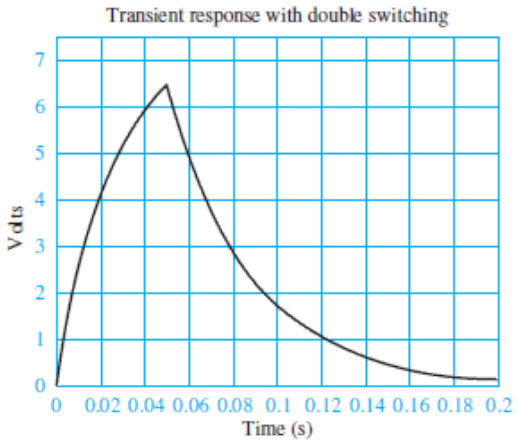
Use the solution for $0 \leq t \leq 50$ ms to calculate $v_C(t = t_1 = 50 \text{ ms})$ and determine α.

$$\alpha = 7.5(1 - e^{-0.05/0.025}) = 6.485 \text{ V}$$

Thus, the capacitor voltage for $t \leq 50$ ms is:

$$v_C(t) = 6.485\, e^{-(t-0.05)/0.0375} \text{ V}$$

The overall composite response is plotted below.



Comments: Note that the two parts of the response are based on two different time constants and that the rising portion of the response changes faster (shorter time constant) than the falling part. Also notice that the transient solution of part 2 was expressed in terms of a *time shift* $(t - 0.05)$ ms, which accounts for the fact that the switch opened at $t = 50$ ms.

EXAMPLE 4.13 First-Order Natural and Forced Responses

Problem

Determine an expression for the capacitor voltage in the circuit of Figure 4.35.

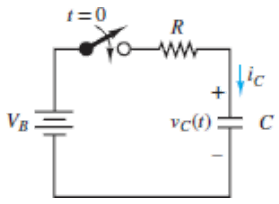


Figure 4.35

Solution

Known Quantities: $v_C(t = 0^-)$; V_B; R; C.

Find: The state variable $v_C(t)$ as a function of time for all t.

Schematics, Diagrams, Circuits, and Given Data: $v_C(t = 0^-) = 5$ V; $R = 1$ kΩ; $C = 470$ μF; $V_B = 12$ V. Figure 4.35.

Assumptions: None.

Analysis:

Step 1: **Find v_C for $t < 0$.** For $t < 0$, the capacitor is not part of a closed loop; therefore, the current through the capacitor must be zero for $t < 0$. In other words, its charge (and consequently, its energy) are constant prior to the switch closing. In this example, it is assumed that the capacitor has an initial charge $q = Cv_C(0^-) = C(5\text{ V})$. Thus:

$$v_C(t) = 5\text{ V} \qquad t < 0 \qquad \text{and} \qquad v_C(0^-) = 5\text{ V}$$

Step 2: **Find the initial condition on v_C.** Since the voltage across the capacitor is always continuous the initial condition is

$$v_C(0^+) = v_C(0^-) = 5\text{ V} \qquad \text{Continuity of capacitor voltage}$$

Step 3: Simplify the circuit for $t > 0$. For $t > 0$, the network attached to the capacitor is already in the form of a Thévenin source, so no further simplification is possible.

Step 4: Find the differential equation. Apply KVL around the loop in Figure 4.35 to yield the differential equation for $t > 0$:

$$12\text{ V} - Ri_C - v_C = 12\text{ V} - RC\frac{dv_C}{dt} - v_C = 0 \qquad t > 0$$

$$RC\frac{dv_C}{dt} + v_C = 12\text{ V} \qquad t > 0$$

The time constant τ is the coefficient of the first-derivative term:

$$\tau = RC = 0.47\text{ s}$$

Step 5: Find the transient solution. The transient solution is found by setting the right side of the differential equation to zero and solving for v_C. The solution is always of the form:

$$(v_C)_{\text{tr}} = \alpha e^{-t/\tau} = \alpha e^{-t/0.47}$$

It is important to note that the unknown constant α is found by applying the initial condition to the *complete* solution, not to the transient solution alone.

Step 6: Find the long-term steady-state solution. The long-term DC steady-state solution for v_C is found after the switch has been closed for a very long time (practically $t \leq 5\tau$). In this state, the capacitor acts like an open-circuit such that $(v_C)_{\text{ss}} \equiv v_C(\infty) = 12$ V.

Step 7: Complete solution. The complete solution is the sum of the *transient* and *long-term steady-state* solutions.

$$v_C(t) = (v_C)_{\text{tr}} + (v_C)_{\text{ss}} = \alpha\, e^{-t/0.47} + 12\text{ V}$$

Page 289

The unknown constant α is found by applying the initial condition $v_C(0^+) = 5$ V. The result is

$$5\text{ V} = v_C(0^+) = \alpha + 12\text{ V} \qquad \text{or} \qquad \alpha = 5\text{ V} - 12\text{ V} = -7\text{ V}$$

Finally, the complete solution is

$$v_C(t) = -7e^{-t/0.47} + 12$$

The complete solution can also be expressed in terms of *natural* and *forced* responses:

$$v_C(t) = v_{C\,N}(t) + v_{C\,F}(t) = 5e^{-t/0.47} + 12(1 - e^{-t/0.47})$$

The complete response and its decomposition into (a) transient plus steady-state responses, and (b) natural plus forced responses are shown in Figure 4.36.

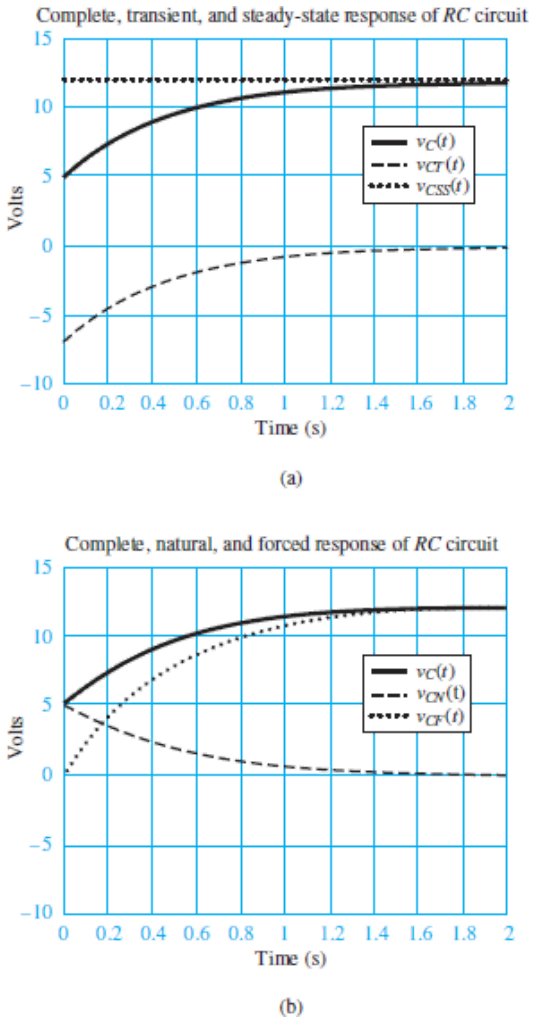


Figure 4.36 (a) Complete, transient, and steady-state responses of the circuit in Figure 4.35; (b) complete, natural, and forced responses of the same circuit

Comments: Note how in Figure 4.36(a) the long-term steady-state response $(v_C)_{ss}$ equals the battery voltage while the transient response $(v_C)_{tr}$ rises from –7 to 0 V exponentially. In Figure 4.36(b), on the other hand, the energy initially stored in the

capacitor decays to zero via its natural response $v_{C\,N}$ while the battery causes the capacitor voltage to rise exponentially to 12 V, as shown in the forced response $v_{C\,F}$.

Page 290

FOCUS ON MEASUREMENTS

Coaxial Cable Pulse Response

Problem:

A problem of great practical importance is the transmission of *pulses* along cables. Short voltage pulses are used to represent the two-level binary signals that are characteristic of digital computers; it is often necessary to transmit such voltage pulses over long distances through **coaxial cables,** which are characterized by a finite resistance per unit length and by a certain capacitance per unit length, usually expressed in picofarads per meter. A simplified model of a long coaxial cable is shown in Figure 4.37. If a 10-m cable has a capacitance of 1,000 pF/m and a series resistance of 0.2 Ω/m, what will the output of the pulse look like after traveling the length of the cable?

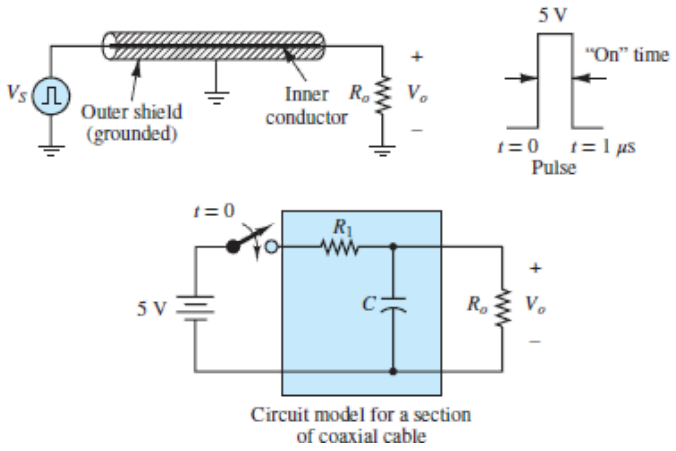


Figure 4.37 Pulse transmission in a coaxial cable

Solution:

Known Quantities—Cable length, resistance, and capacitance; voltage pulse amplitude and time duration.

Find—The cable voltage as a function of time.

Schematics, Diagrams, Circuits, and Given Data—$r_1 = 0.2\ \Omega$/m; $R_o = 150\ \Omega$; $c =$ 1,000 pF/m; $l = 10$ m; pulse duration = 1 μs.

Assumptions—The short voltage pulse is applied to the cable at $t = 0$. Assume zero initial conditions.

Analysis—The voltage pulse can be modeled by a 5-V battery connected to a switch; the switch will then close at $t = 0$ and open again at $t = 1\ \mu$s. The solution strategy will therefore proceed as follows. First, determine the initial condition; next, solve the transient problem for $t > 0$; finally, compute the value of the capacitor voltage at $t = 1\ \mu$s—that is, when the switch opens again—and solve a different transient problem. The equivalent capacitor will charge for 1 μs, and the voltage will reach a certain value. This value will be the initial condition for the capacitor voltage when the switch is opened; the capacitor voltage will then decay to zero since the voltage source has been disconnected. Note that the circuit will be characterized by two different time constants during the two transient stages of the problem. The initial condition for this problem is zero, assuming that the switch has been open for a long time.

Page 291

The differential equation for $0 < t < 1\ \mu$s is obtained by computing the Thévenin equivalent circuit relative to the capacitor when the switch is closed:

$$V_T = \frac{R_o}{R_1 + R_o} V_B \qquad R_T = R_1 \| R_o \qquad \tau = R_T C \qquad 0 < t < 1\ \mu s$$

The differential equation is given by the expression

$$R_T C \frac{dv_C}{dt} + v_C = V_T \qquad 0 < t < 1\ \mu s$$

and the solution is of the form

$$v_C(t) = -V_T e^{-t/R_T C} + V_T = V_T(1 - e^{-t/R_T C}) \qquad 0 < t < 1\ \mu s$$

Numerical values can be assigned to the solution by calculating the effective resistance and capacitance of the cable:

$$R_1 = r_1 \times l = 0.2 \times 10 = 2\ \Omega$$
$$C = c \times l = 1{,}000 \times 10 = 10{,}000\ \text{pF}$$
$$R_T = 2 \| 150 \approx 1.97\ \Omega \qquad V_T = \frac{150}{152} V_B \approx 4.93\ \text{V}$$
$$\tau_{on} = R_T C \approx 19.7 \times 10^{-9}\ \text{s}$$

so that

$$v_C(t) \approx 4.93(1 - e^{-t/19.7 \times 10^{-9}})\ \text{V} \qquad 0 < t < 1\ \mu s$$

At the time when the switch opens again, $t = 1\ \mu$s, the capacitor voltage can be found to be $v_C(t = 1\ \mu s) \approx 4.93$ V.

When the switch opens again, the capacitor will discharge through the load resistor R_o; this discharge is described by the natural response of the circuit consisting of C and R_o and is governed by the following values: $v_C(t = 1\ \mu s) = 4.93$ V, $\tau_{off} = R_o C = 1.5\ \mu$s. The natural response can be written directly as:

$$v_C(t) = v_C(t = 1 \times 10^{-6}) \times e^{-(t - 1 \times 10^{-6})/\tau_{off}}$$
$$= 4.93 \times e^{-(t - 1 \times 10^{-6})/1.5 \times 10^{-6}}\ \text{V} \qquad t \geq 1\ \mu s$$

Figure 4.38 shows a plot of the solution for $t > 0$, along with the voltage pulse.

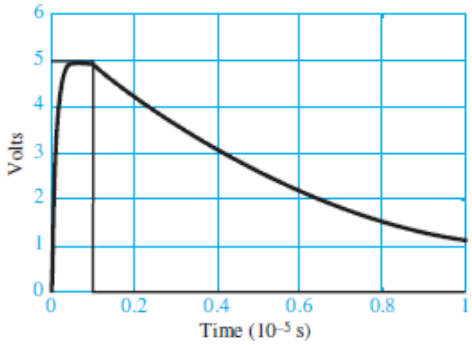


Figure 4.38 Coaxial cable pulse response

Page 292

Comments—Note that the voltage response shown in Figure 4.38 rapidly reaches the desired value, near 5 V, thanks to the very short charging time constant τ_{on}. As the length of the cable is increased, however, τ_{on} will increase, to the point that the voltage pulse may not rise sufficiently close to the desired 5-V value in the desired time. Cable length limitations exist in some applications because of intrinsic capacitance and resistance. In general, long cables such as electric transmission lines and very high-frequency circuits cannot be analyzed by way of the lumped-parameter methods presented here and require *distributed circuit analysis* techniques.

CHECK YOUR UNDERSTANDING

If another identical capacitor is placed in parallel with the capacitor in Example 4.7, how would the charging time change? How would the total stored energy change?

Answer: Both would double, as C_{eq} would be twice as large, thus doubling τ and E_{total}.

CHECK YOUR UNDERSTANDING

Derive the result obtained in Example 4.11 analytically, by solving the transient response for the unknown value n.

Answer: The solution yields 2.9, which rounds up to $n = 3$.

CHECK YOUR UNDERSTANDING

In Example 4.12, what will be the initial condition for the falling exponential decay if the switch opens at $t = 100$ ms?

Answer: 7.363 V

CHECK YOUR UNDERSTANDING

What is the complete solution for $v_C(t)$ in Example 4.13 when the initial charge on the capacitor is zero?

Answer: The complete solution is equal to the forced solution. $v_C(t) = v_F(t) = 12(1 - e^{-t/0.47})$ V

Page 293

4.4 SECOND-ORDER TRANSIENT ANALYSIS

In general, a second-order circuit has two irreducible storage elements: two capacitors, two inductors, or one capacitor and one inductor. The latter case is the most important in terms of new fundamentals; however, the important aspects of all second-order system responses are discussed in this section. Since second-order circuits have two irreducible storage elements, such circuits have two-state variables and their behavior is described by a second-order differential equation.

The simplest, yet arguably the most crucial, second-order circuits are those in which the capacitor and inductor are either in parallel or in series, as shown in

Figures 4.39 and 4.40. The circuits in these figures are drawn to suggest that the capacitor and inductor should be treated as a unified load. The rest of each circuit is either the Thévenin or Norton equivalent of the source network. The analysis of these circuits is somewhat less complicated than for other second-order circuits, which is appealing for anyone learning to analyze such circuits for the first time. The analysis of more complicated second-order circuits is treated in an example later in this section.

It is important to adopt a patient but determined attitude toward the material in this section, as it is notoriously challenging to students. Every effort has been made to walk through the material in a systematic and progressive manner. Hold on to your hat! And don't give up!

General Characteristics

Before diving into the analysis of particular second-order circuits it is worthwhile to introduce the generalized **standard form** of the differential equation for any second-order circuit.

$$\frac{1}{\omega_n^2}\frac{d^2x}{dt^2}+\frac{2\zeta}{\omega_n}\frac{dx}{dt}+x=K_S f(t) \tag{4.25}$$

The constants ω_n and ζ are the **natural frequency** and the **dimensionless damping ratio,** respectively. These parameters are characteristics of a second-order circuit and determine its response. Their values will be determined by direct comparison of equation 4.25 with the differential equation for a specific *RLC* circuit. As will be shown, second-order circuits have three distinct possible responses: *overdamped, critically damped,* and *underdamped.* The type of response for any particular second-order circuit is determined entirely by ζ.

In equation 4.25, $f(t)$ is a forcing function. K_S is the **DC gain** of a particular variable $x(t)$. Different variables in the same circuit may have different DC gains. However, all variables share the same natural frequency ω_n, the same dimensionless damping ratio ζ, and therefore also the same type of response. This fact can be an important time saver when problem solving.

Parallel *LC* Circuits

Consider the circuit shown in Figure 4.39. The two state variables are i_L and v_C. In general, at the moment ($t = 0$) of a transient event the energy of the storage elements may be nonzero; that is, the voltage $v_C(0)$ across the capacitor and the Page

294current $i_L(0)$ through the inductor may be nonzero. As always, the two-state variables are continuous such that:

$$v_C(0^+) = v_C(0^-) \qquad \text{and} \qquad i_L(0^+) = i_L(0^-)$$

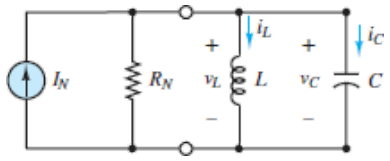


Figure 4.39 Second-order circuit with the inductor and capacitor in parallel acting as a unified load attached to a Norton source

Apply KCL to either node to find an equation in terms of both state variables.

$$I_N - \frac{v_C}{R_N} - i_L - i_C = 0 \qquad \text{KCL}$$

Here, R_N is the Norton equivalent resistance seen by the LC load. KVL around the right mesh yields the simple result $v_C = v_L$. The constitutive i-v relationships for the capacitor and inductor enable the nonstate variables i_C and v_L to be replaced by derivatives of the state variables v_C and i_L, respectively.

$$v_C = v_L = L\frac{di_L}{dt} \qquad \text{and} \qquad i_C = C\frac{dv_C}{dt} = LC\frac{d^2i_L}{dt^2}$$

Substitute for v_C and i_C in the KCL equation to find:

$$I_N - \frac{L}{R_N}\frac{di_L}{dt} - i_L - LC\frac{d^2i_L}{dt^2} = 0$$

Rearrange the order of terms to yield the following second-order differential equation in standard form:

$$\boxed{LC\frac{d^2i_L}{dt^2} + \frac{L}{R_N}\frac{di_L}{dt} + i_L = I_N} \tag{4.26}$$

Alternatively, one can differentiate both sides of the KCL equation and substitute:

$$\frac{di_L}{dt} = \frac{v_L}{L} = \frac{v_C}{L} \qquad \text{and} \qquad \frac{di_C}{dt} = C\frac{d^2v_C}{dt^2}$$

The result is

$$\frac{dI_N}{dt} - \frac{1}{R_N}\frac{dv_C}{dt} - \frac{v_C}{L} - C\frac{d^2v_C}{dt^2} = 0$$

Multiply both sides of the equation by L. If the source I_N is a constant, the resulting second-order differential equation in standard form is

$$LC\frac{d^2v_C}{dt^2} + \frac{L}{R_N}\frac{dv_C}{dt} + v_C = 0 \tag{4.27}$$

Page 295

MAKE THE CONNECTION

Automotive Suspension

The mechanical model shown below can be analyzed using Newton's second law, $ma = \sum F$, to obtain the equation

$$m\frac{d^2x(t)}{dt^2} = F(t) - b\frac{dx(t)}{dt} - kx(t)$$

This equation can be written in the *standard* form:

$$\frac{m}{k}\frac{d^2x(t)}{dt^2} + \frac{b}{k}\frac{dx(t)}{dt} + x(t) = \frac{1}{k}f(t)$$

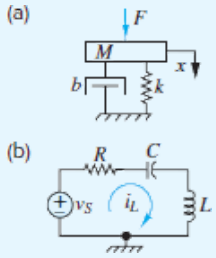
The analogous LC series circuit shown below can be analyzed using KVL:

$$v_s - Ri_L - v_C - L\frac{di_L}{dt} = 0$$

$$i_L = i_C = C\frac{dv_C}{dt}$$

$$LC\frac{d^2v_C}{dt^2} + RC\frac{dv_C}{dt} + v_C = v_S$$

Notice the similar structure of these two second-order differential equations.



Analogy between electrical and mechanical systems

Compare both second-order differential equations to the standard form of equation 4.25 to make the following observations:

$$\frac{1}{\omega_n^2}\frac{d^2x(t)}{dt^2} + \frac{2\zeta}{\omega_n}\frac{dx(t)}{dt} + x(t) = K_s f(t)$$

$$\omega_n = \sqrt{\frac{k}{m}}$$

$$\zeta = \frac{b}{k}\frac{\omega_n}{2} = \frac{b}{2}\sqrt{\frac{1}{km}}$$

Mechanical

$$\omega_n = \sqrt{\frac{1}{LC}}$$

$$\zeta = RC\frac{\omega_n}{2} = \frac{R}{2}\sqrt{\frac{C}{L}}$$

Electrical

The following analogies are apparent after comparing the expressions for the natural frequency and damping ratio in the two differential equations:

Mechanical system	Electrical system
Damping coefficient b	Resistance R
Mass m	Inductance L
Compliance $1/k$	Capacitance C

In general, the coefficient of the first-order derivative in equations such as 4.26 and 4.27 is the sum $(R_TC + L/R_N)$. Here, R_T is the Thévenin equivalent resistance

seen by the capacitor when the inductor is treated as a short-circuit and R_N is the Norton equivalent resistance seen by the inductor when the capacitor is treated as an open-circuit. In the case of a parallel LC network, $R_T = 0$.

To solve equations 4.26 and 4.27 it is first necessary to identify the *dimensionless damping ratio* ζ and the *natural frequency* ω_n. Notice that the left sides of both equations are identical, as they are for any variable in the circuit. Thus, ω_n and ζ can be found from either differential equation by comparing it to equation 4.25. The result is

$$\frac{1}{\omega_n^2} = LC \qquad \text{and} \qquad \frac{2\zeta}{\omega_n} = R_T C$$

These two equations can be solved to yield:

$$\omega_n = \frac{1}{\sqrt{LC}} \qquad \text{and} \qquad \zeta = \frac{R_T}{2}\sqrt{\frac{C}{L}} \tag{4.28}$$

The type of transient response for i_L and v_C depends upon ζ only. When ζ is greater than, equal to, or less than 1, the transient responses $(i_L)_{tr}$ and $(v_C)_{tr}$ are *overdamped, critically damped,* or *underdamped,* respectively. These three types of responses are described in detail later in this section. The complete solutions are

$$v_C(t) = (v_C)_{tr} + (v_C)_{ss} = (v_C)_{tr} + V_T$$

and

$$i_L(t) = (i_L)_{tr} + (i_L)_{ss} = (i_L)_{tr} + C\frac{dV_T}{dt}$$

Note that when I_N is a constant, $(v_C)_{ss} = 0$ and $v_C(t) = (v_C)_{tr}$.

Series *LC* Circuits

The development of the general solution for series LC circuits follows the same basic steps used above for parallel LC circuits. Consider the circuit in Figure 4.40 and note the duality between what follows and what was done above for the parallel LC circuit. In fact, the equations that follow can be found simply by starting with the equations developed above and swapping L with C, i_L with v_C, R_N with $1/R_T$, and I_N with V_T. That result is known as *duality*.

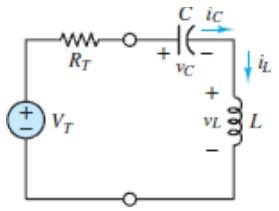


Figure 4.40 Secondorder circuit with the inductor and capacitor in series acting as a unified load attached to a Thévenin source

Page 296

Again, the two state variables are i_L and v_C. At the moment ($t = 0$) of a transient event the energy of the storage elements may be nonzero; that is, the voltage $v_C(0)$ across the capacitor and the current $i_L(0)$ through the inductor may be nonzero. As always, the two-state variables are continuous such that:

$$v_C(0^+) = v_C(0^-) \qquad \text{and} \qquad i_L(0^+) = i_L(0^-)$$

Apply KVL around the series loop to find an equation in terms of both state variables.

$$V_T - i_L R_T - v_C - v_L = 0 \qquad \text{KVL}$$

Here, R_T is the Thévenin equivalent resistance seen by the LC load. KCL at the upper right node yields the simple result $i_C = i_L$. The constitutive i-v relationships for the capacitor and inductor enable the nonstate variables i_C and v_L to be replaced by derivatives of the state variables v_C and i_L, respectively.

$$i_L = i_C = C\frac{dv_C}{dt} \qquad \text{and} \qquad v_L = L\frac{di_L}{dt} = LC\frac{d^2v_C}{dt^2}$$

Substitute for v_L and i_L in the KVL equation to find:

$$V_T - R_T C\frac{dv_C}{dt} - v_C - LC\frac{d^2v_C}{dt^2} = 0$$

Rearrange the order of terms to yield the following second-order differential equation in standard form:

$$LC\frac{d^2v_C}{dt^2} + R_T C\frac{dv_C}{dt} + v_C = V_T \tag{4.29}$$

Alternatively, one can differentiate both sides of the KVL equation and substitute:

$$\frac{dv_C}{dt} = \frac{i_C}{C} = \frac{i_L}{C} \quad \text{and} \quad \frac{dv_L}{dt} = L\frac{d^2 i_L}{dt^2}$$

The result is

$$\frac{dV_T}{dt} - R_T\frac{di_L}{dt} - \frac{i_L}{C} - L\frac{d^2 i_L}{dt^2} = 0$$

Multiply both sides of the equation by C, and if the source V_T is a constant such that its time derivative is zero, the resulting second-order differential equation in standard form is

$$LC\frac{d^2 i_L}{dt^2} + R_T C\frac{di_L}{dt} + i_L = 0 \tag{4.30}$$

In general, the coefficient of the first-order derivative in equations such as 4.29 and 4.30 is the sum (R_TC + L/R_N). Here, R_N is the Norton equivalent resistance seen by the inductor when the capacitor is treated as an open-circuit and R_T is the Thévenin equivalent resistance seen by the capacitor when the inductor is treated as a short-circuit. In the case of a series LC network, $R_N \to \infty$ such that $L/R_N \to 0$.

To solve equations 4.29 and 4.30 it is first necessary to identify the *dimensionless damping ratio* ζ and the *natural frequency* ω_n. Notice that the left sides Page 297of both equations are identical, as they are for any variable in the circuit. Thus, ω_n and ζ can be found from either differential equation by comparing it to equation 4.25. The result is

$$\frac{1}{\omega_n^2} = LC \quad \text{and} \quad \frac{2\zeta}{\omega_n} = R_T C$$

These two equations can be solved to yield:

$$\omega_n = \frac{1}{\sqrt{LC}} \quad \text{and} \quad \zeta = \frac{R_T}{2}\sqrt{\frac{C}{L}} \tag{4.31}$$

The type of transient response for i_L and v_C depends upon ζ only. As always, when ζ is greater than, equal to, or less than 1, the transient responses $(i_L)_{\text{tr}}$ and $(v_C)_{\text{tr}}$

are *overdamped, critically damped,* or *underdamped,* respectively. These three types of responses are described in detail later in this section. The complete solutions are

$$v_C(t) = (v_C)_{tr} + (v_C)_{ss} = (v_C)_{tr} + V_T$$

and

$$i_L(t) = (i_L)_{tr} + (i_L)_{ss} = (i_L)_{tr} + C\frac{dV_T}{dt}$$

Note that when V_T is a constant, $(i_L)_{ss} = 0$ and $i_L(t) = (i_L)_{tr}$.

Transient Response

The transient response $x_{tr}(t)$ is found by setting the right side of the governing differential equation equal to zero. That is

$$\frac{1}{\omega_n^2}\frac{d^2x_{tr}}{dt^2} + \frac{2\zeta}{\omega_n}\frac{dx_{tr}}{dt} + x_{tr} = 0 \tag{4.32}$$

Just as in first-order systems, the solution of this equation has an exponential form:

$$x_{tr}(t) = \alpha e^{st} \tag{4.33}$$

Substitution into the differential equation yields the *characteristic equation:*

$$\frac{s^2}{\omega_n^2} + \frac{2\zeta}{\omega_n}s + 1 = 0 \tag{4.34}$$

which, in turn, yields two **characteristic roots** s_1 and s_2. Specific values of s_1 and s_2 are found from the quadratic formula applied to the characteristic equation.

$$s_{1,2} = -\zeta\omega_n \pm \frac{1}{2}\sqrt{(2\zeta\omega_n)^2 - 4\omega_n^2} = -\omega_n\left(\zeta \pm \sqrt{\zeta^2 - 1}\right) \tag{4.35}$$

The roots s_1 and s_2 are associated with the three distinct possible responses: overdamped ($\zeta > 1$), critically damped ($\zeta = 1$), and underdamped ($\zeta < 1$).

1. Overdamped Response ($\zeta > 1$)

Two distinct, negative, and real roots: (s_1, s_2). The transient response is overdamped when $\zeta > 1$ and the roots are $s_{1,2} = \omega_n\left(-\zeta \pm \sqrt{\zeta^2 - 1}\right)$. The general form of the solution is

$$x_{tr}(t) = \alpha_1 e^{s_1 t} + \alpha_2 e^{s_2 t} = e^{-\zeta\omega_n t}\left[\alpha_1 e^{\left(\omega_n\sqrt{\zeta^2-1}\right)t} + \alpha_2 e^{\left(-\omega_n\sqrt{\zeta^2-1}\right)t}\right] \tag{4.36}$$

Page 298

Thus, an overdamped response is the sum of two decaying exponentials, as shown in Figure 4.41.

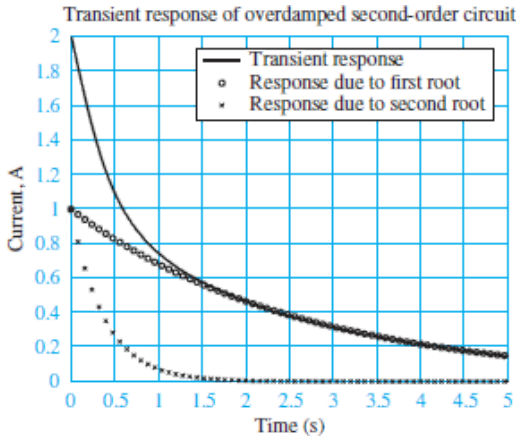


Figure 4.41 Transient response of underdamped second-order system $\alpha_1 = \alpha_2 = 1$; $\zeta = 1.5$; $\omega_n = 1$

2. Critically Damped Response ($\zeta = 1$)

Two identical, negative, and real roots: (s_1, s_2). The transient response is critically damped when $\zeta = 1$. The argument of the square root in equation 4.35 is zero, such that $s_{1,2} = -\zeta\omega_n = -\omega_n$. The general form of the solution is

$$x_{tr}(t) = \alpha_1 e^{s_1 t} + \alpha_2 t e^{s_2 t} = e^{-\omega_n t}(\alpha_1 + \alpha_2 t) \tag{4.37}$$

Note that a critically damped response is the sum of two decaying exponentials, where one is multiplied by t. These two components and the complete response are shown in Figure 4.42.

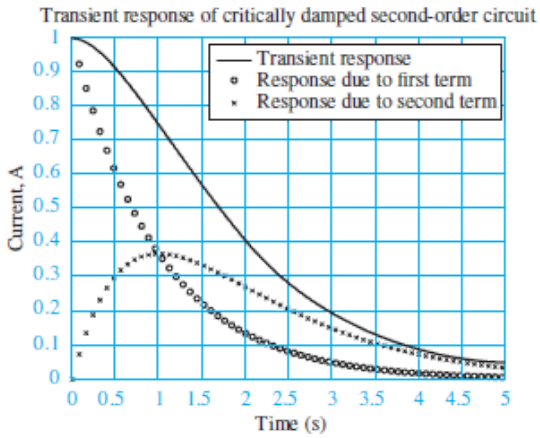


Figure 4.42 Transient response of a critically damped second-order system for $\alpha_1 = \alpha_2 = 1$; $\zeta = 1$; $\omega_n = 1$

3. Underdamped Response ($\zeta < 1$)

Two complex conjugate roots: (s_1, s_2). The transient response is underdamped when $\zeta < 1$. The argument of the square root in equation 4.35 is negative, such Page 299that $s_{1,2} = \omega_n\left(-\zeta \pm j\sqrt{1-\zeta^2}\right)$. The following complex exponentials appear in the general form of the response:

$$e^{\omega_n\left(-\zeta+j\sqrt{1-\zeta^2}\right)t} \qquad e^{\omega_n\left(-\zeta-j\sqrt{1-\zeta^2}\right)t} \tag{4.38}$$

Euler's formula can be used to express the complex exponentials in terms of sinusoids. The result is

$$x_{tr}(t) = e^{-\zeta\omega_n t}\left[\alpha_1 \sin(\omega_d t) + \alpha_2 \cos(\omega_d t)\right] \tag{4.39}$$

where $\omega_d = \omega_n\sqrt{1-\zeta^2}$ is the **damped natural frequency.** Note that ω_d is the frequency of oscillation and is related to the period T of oscillation by $\omega_d T = 2\pi$. Also note that ω_d approaches the natural frequency ω_n as ζ approaches zero. The oscillation is *damped* by the decaying exponential $e^{-\zeta\omega nt}$, which has a time constant $\tau = 1/\zeta\omega_n$, as shown in Figure 4.43. As ζ increases toward 1 (more damping), τ decreases and the oscillations decay more quickly. In the limit $\zeta \to 0$, the response is a pure sinusoid.

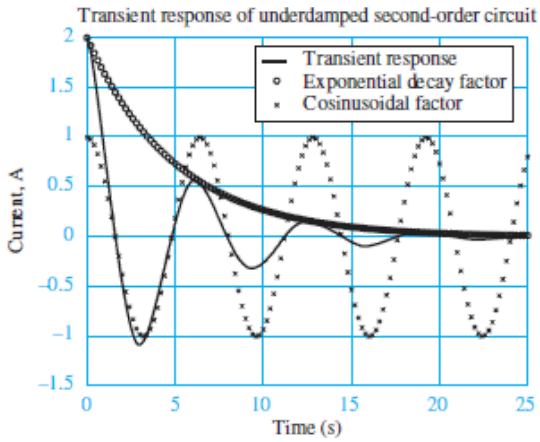


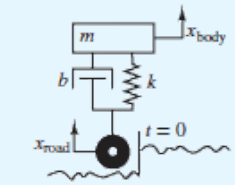
Figure 4.43 Transient response of an underdamped second-order system for $\alpha_1 = \alpha_2 = 1$; $\zeta = 0.2$; $\omega_n = 1$

MAKE THE CONNECTION

Automotive Suspension

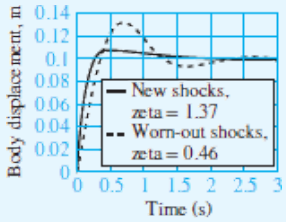
The mechanical model described in the previous sidebar can represent an automotive suspension system. The mass m, spring k, and damper b model the vehicle mass, the suspension struts (or coils), and the shock absorbers, respectively. The differential equation is

$$\frac{m}{k}\frac{d^2x_{body}(t)}{dt^2}+\frac{b}{k}\frac{dx_{body}(t)}{dt}+x_{body}(t)=\frac{1}{k}x_{road}(t)+\frac{b}{k}\frac{dx_{road}(t)}{dt}$$



$m = 1{,}500$ kg
$k = 20{,}000$ N/m
$b_{new} = 15{,}000$ N-s/m
$b_{old} = 5{,}000$ N-s/m

The input to the suspension system is the road surface profile, which generates both displacement and velocity inputs x_{road} and $\dot{x}_{road}$. One objective of the suspension is to isolate the body of the car (i.e., the passengers) from any vibration caused by unevenness in the road surface. Automotive suspension systems are also very important in guaranteeing vehicle stability and in providing acceptable handling. In this illustration consider the response of the vehicle to a sharp step of amplitude 10 cm for two cases, corresponding to new and worn-out shock absorbers, respectively. Which ride is more comfortable?



"Step" response of automotive suspension

Long-Term Steady-State Response

For switched DC sources, the forcing function F in equation 4.40 is a constant. The result is a constant long-term ($t \to \infty$) steady-state response x_{SS}.

$$\frac{1}{\omega_n^2}\frac{d^2x_{SS}}{dt^2}+\frac{2\zeta}{\omega_n}\frac{dx_{SS}}{dt}+x_{SS}=K_SF \tag{4.40}$$

Since x_{ss} must also be a constant the solution is

$$x_{ss} = x(\infty) = K_S F \qquad t \to \infty \tag{4.41}$$

Complete Response

As with first-order systems, the complete response is the sum of the transient and long-term steady-state responses. The complete mathematical solutions for the overdamped, critically damped, and underdamped cases are shown in the highlighted Focus on Problem Solving section. In each of these cases, the initial conditions on the state variables must be used to solve for the unknown constants α_1 and α_2. The required procedure uses the two initial conditions to evaluate $x(t)$ and Page 300dx/dt at $t = 0^+$. The details of the procedure vary slightly in each of the three cases and are illustrated in the example problems.

One particularly useful complete solution is the *unit-step response* brought about by letting $K_S f(t)$ (see equation 4.25) be a *unit step,* which equals 0 for $t < 0$ and 1 for $t > 0$. To illustrate, assume a dimensionless damping coefficient $\zeta = 0.1$ and an underdamped period of oscillation $T = 2\pi$, such that the damped natural frequency is $\omega_d = 1$. The corresponding *unit-step response,* shown in Figure 4.44, asymptotically approaches the *long-term DC steady-state* value of 1 dictated by the unit-step input.

Also, as seen in the underdamped transient response, the magnitude of the oscillation *decays exponentially* over time. The time constant for the surrounding envelope (see dashed lines in Figure 4.44) is $\omega_d \tau = \sqrt{1 - \zeta^2}/\zeta \approx 10$, such that by $t = 5\tau$ the oscillations are within 1 percent of the long-term DC steady-state value.

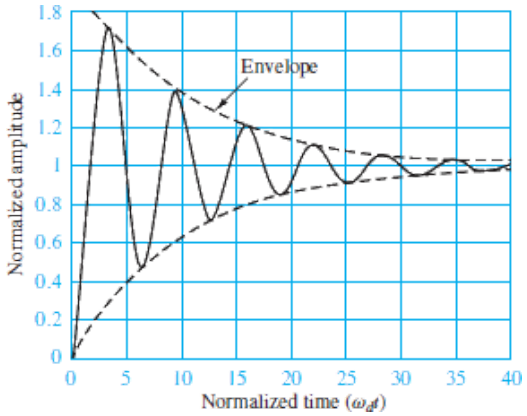


Figure 4.44 Second-order transient response with $K_S = 1$, $\omega_d = 1$, and $\zeta = 0.1$

Note that the rate of decay of the oscillations is governed by ζ. Figure 4.45 shows that as ζ increases the *overshoot* of the long-term DC steady-state response decreases until, when $\zeta = 1$ (critically damped), the response no longer oscillates and the

overshoot is zero. The response for $\zeta > 1$ (overdamped) has no oscillations and zero overshoot.

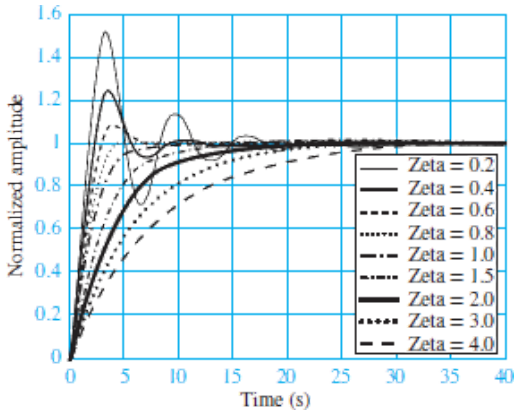


Figure 4.45 Second-order unit-step response with $K_S = 1$, $\omega_d = 1$, and ζ ranging from 0.2 to 4.0

Page 301

LO

FOCUS ON PROBLEM SOLVING

SECOND-ORDER TRANSIENT RESPONSE

The following steps determine the response of a generic second-order *RLC* circu a transient event, such as the throwing of a switch, at $t = 0$. Here, $x(t)$ is a variable, either the capacitor voltage $v_C(t)$ or the inductor current $i_L(t)$.

1. **DC steady-state responses:** Assume all independent sources are DC and a steady-state prior to the transient event.
   - Apply DC analysis to solve for $v_C(0^-)$ and $i_L(0^-)$ just prior to the event.
   - Apply DC analysis to solve for the long-term steady-state values $v_C(\infty)$ $i_L(\infty)$ after the event.
2. **Differential equation for $t > 0$:**
   - Identify the simplest one-port load network that contains both sto elements. Simplify the remaining one-port source network into eith Thévenin or Norton source.

- Apply KVL, KCL and the *i-v* relations for capacitors and inductors to two first-order ordinary differential equations involving the state varia only.
- Use the first-order ODEs to find a second-order ODE in standard form in state variable.
- Check that the coefficient of the first derivative term is $R_T C+L/R_N$. Here, R_T the Thévenin equivalent resistance seen by capacitor with the inductor ac as a short-circuit and R_N is the Norton equivalent resistance seen by inductor with the capacitor acting as an open-circuit.
- Check that the right side of the second-order ODE is $v_C(\infty)$ or $i_L(\infty)$.

3. **Solve for ω_n** and ζ for $t > 0$:
   - Compare the second-order ODE in standard form to its generalized f (equation 4.25).
   - Set the coefficient of the second derivative term equal to $1/\omega_n^2$. Solve for ω_n
   - Set the coefficient of the first derivative term equal to $2\zeta/\omega_n$. Solve for ζ.
4. **The transient response $x_{tr}(t)$:**

   Overdamped case ($\zeta > 1$):

$$x_{tr}(t) = e^{-\zeta\omega_n t}\left(\alpha_1 e^{\omega_n t\sqrt{\zeta^2-1}} + \alpha_2 e^{-\omega_n t\sqrt{\zeta^2-1}}\right) \qquad t \geq 0$$

   Critically damped case ($\zeta = 1$):

$$x_{tr}(t) = e^{-\omega_n t}(\alpha_1 + \alpha_2 t) \quad t \geq 0$$

   Underdamped case ($\zeta < 1$):

$$x_{tr}(t) = e^{-\zeta\omega_n t}[\alpha_1 \sin(\omega_d t) + \alpha_2 \cos(\omega_d t)] \qquad t \geq 0$$
$$\omega_d = \omega_n\sqrt{1-\zeta^2}$$

5. **The complete solution $x(t)$:**

$$x(t) = x_{tr}(t) + x(\infty)$$

6. **Solve for unknown constants α_1 and α_2:**

- Apply continuity of the state variables to set the initial conditions $v_C(0$ $v_C(0^-)$ and $i_L(0^+) = i_L(0^-)$.
- Express $x(0^+)$ in terms of α_1 and α_2. Plug in the initial condition for $x(0^+)$.
- Differentiate $x(t)$ found in step 5. Then, use the first-order ODEs from st to find another expression for dx/dt. Set these expressions equal to each c at $t = 0^+$.
- Solve the system of two equations in the two unknowns α_1 and α_2.
- Whew! Voila!! Done!!!

Comments: It is important to keep in mind that ωn and ζ are the same fo variables in the circuit. Thus, the *form* of the transient response is also the same all such variables. However, the values of α_1, α_2, and the long-term DC steady-s are specific to each variable. A good problem-solving method is to use the in conditions to determine α_1 and α_2 for each state variable. Then, the response of c variables in the circuit can be found by applying simple circuit analysis method relate those variables to the state variables. This approach avoids errors commonly occur when attempting to find the response of nonstate variables direc

LO

FOCUS ON PROBLEM SOLVING

ROOTS OF SECOND-ORDER SYSTEMS

The general form of the roots s_1 and s_2 is $s_{1,2} = -\zeta\omega_n \pm \omega_n\sqrt{\zeta}$.

The nature of these roots depends upon the argument of the square root.

Case 1: **Distinct, negative, real roots.** This case occurs when $\zeta > 1$ since the under the square root sign is positive. The result is $s_{1,2} = -\omega_n\left[\zeta \pm \sqrt{\zeta^2 - 1}\right]$ a second-order **overdamped response.**

Case 2: **Identical, negative, real roots.** This case occurs when $\zeta = 1$ since the under the square root is zero. The result is a repeated root $s = -\zeta\omega_n = -\omega_n$ a second-order **critically damped response.**

Case 3: **Complex conjugate roots.** This case holds when $\zeta < 1$ since the term u the square root is negative. The result is a pair of complex conjugate r $s_{1,2} = -\omega_n\left[\zeta \pm j\sqrt{1 - \zeta^2}\right]$ and a second-order **underdamped response.**

EXAMPLE 4.14 Complete Response of an Underdamped Parallel *LC* Circuit

Problem

Find the natural frequency, ω_n, the dimensionless damping coefficient, ζ, and the form of the transient response of $i_L(t)$ in the circuit shown in Figure 4.46.

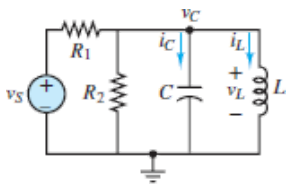


Figure 4.46

Solution

Known Quantities: v_S; R_1; R_2; C; L.

Find: The transient response of $i_L(t)$ for the circuit in Figure 4.46.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 8$ k Ω; $R_2 = 8$ k Ω; $C = 10$ μF; $L = 1$ H.

Assumptions: None

Analysis: Refer to the Focus on Problem Solving box "Second-Order Transient Response" and the subsection on parallel *LC* circuits. The load is L in parallel with C.

Step 1: **DC steady-state responses:** There is no information about a transient event, such as the throwing of a switch, so it is not possible to describe a DC steady-state prior to such an event. However, it is still possible to describe how the circuit in Figure 4.46 would respond to such an event. For example, the inductor and capacitor will act as short- and open-circuits as $t \rightarrow \infty$ such that the long-term DC steady-states for each state variable are $i_L(\infty) = v_S/R_1$ and $v_C(\infty) = 0$.

Step 2: **Differential equation:** Treat everything to the left of the inductor as a one-port source network for the parallel LC load. Basic DC analysis yields the Norton equivalent network $I_N = v_S/R_1$ and $R_N = R_1 \;/\!/\; R_2 = 4$ kΩ. Replace the source network with its Norton equivalent so that the circuit now has the same form as Figure 4.39.

Apply KCL at the upper node and use the i-v relationship for a capacitor to write:

$$I_N - \frac{v_C}{R_N} - i_C - i_L = I_N - \frac{v_C}{R_N} - C\frac{dv_C}{dt} - i_L = 0$$

Apply KVL around the right mesh and use the i-v relationship for an inductor to write:

$$v_C = v_L = L\frac{di_L}{dt}$$

Use the KVL equation to plug in for v_C in the KCL equation to find:

$$I_N - \frac{L}{R_N}\frac{di_L}{dt} - LC\frac{d^2i_L}{dt^2} - i_L = 0$$

or

$$LC\frac{d^2i_L}{dt^2} + \frac{L}{R_N}\frac{di_L}{dt} + i_L = I_N$$

Step 3: **Solve for ω_n and ζ:**

$$\frac{1}{\omega_n^2} = LC = 10 \times 10^{-6}\ \text{sec}^2 \qquad \text{and} \qquad \frac{2\zeta}{\omega_n} = \frac{L}{R_N} = 250 \times 10^{-6}\ \text{sec}$$

Solve these two equations to find $\omega_n \approx 316$ rad/sec and $\zeta \approx 0.04 < 1$. The transient response is *underdamped*. The damped natural frequency is $\omega_d = \omega_n\,[1 - \zeta^2]^{1/2} \approx 316$ rad/sec. Note that when $\zeta \ll 1$, $\omega_d \approx \omega_n$.

Step 4: **The transient response $x_{tr}(t)$:** Plug in for ω_n and ζ to write:

$$\begin{aligned} i_{L_{tr}}(t) &= e^{-\zeta\omega_n t}[\alpha_1 \sin(\omega_d t) + \alpha_2 \cos(\omega_d t)] \qquad t \geq 0 \\ &= e^{-12.5t}[\alpha_1 \sin(316t) + \alpha_2 \cos(316t)] \end{aligned}$$

Page 303

Step 5: **The complete solution $x(t)$:** From step 1 $i_L(\infty) = I_N = v_S/R1$such that the complete solution for $i_L(t)$ is

$$i_L(t) = e^{-12.5t}[\alpha_1 \sin(316t) + \alpha_2 \cos(316t)] + \frac{v_S}{R_1}$$

Step 6: **Solve for the unknown constants α_1 and α_2:** The constants α_1 and α_2 can be determined once the initial conditions on v_C and i_L are known.

Note that the roots of the characteristic equation associated with the second-order ODE are

$$s_{1,2} = \omega_n\left[-\zeta \pm j\sqrt{1-\zeta^2}\right]$$

Substitute numerical values to find $s_{1,2} = -12.5 \pm j316.0$ rad/sec. Both parts of these roots are observed in the transient response.

EXAMPLE 4.15 Complete Response of an Overdamped Series *LC* Circuit

Problem

Determine the complete response for the inductor current i_L shown in Figure 4.47.

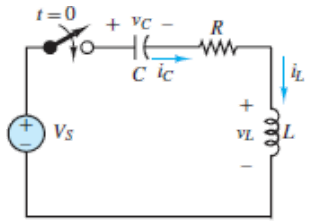


Figure 4.47

Solution

Known Quantities: V_S; R; C; L.

Find: The complete response for the inductor current i_L in the circuit of Figure 4.47.

Schematics, Diagrams, Circuits, and Given Data: $V_S = 25$ V; $R = 5$ kΩ; $C = 1$ μ F; $L = 1$ H.

Assumptions: The capacitor has been charged prior to the switch closing, such that $v_C(0) = 5$ V.

Analysis: Refer to the Focus on Problem Solving box "Second-Order Transient Response" and the subsection on series *LC* circuits. The load is *L* in series with *C*.

Step 1: **DC steady-state responses:** For $t < 0$ the switch is open such that $i_C = i_L = 0$ A. For $t \to \infty$ the switch is closed and the capacitor acts as a DC open-circuit while the inductor acts as a DC short-circuit. Thus, $i_C = i_L = 0$ A, $v_L = 0$ V and the voltage across the resistor must also be zero due to Ohm's law. By KVL, $v_C(\infty) = V_S = 25$ V.

Step 2: **Differential equation for $t > 0$:** With the switch closed the circuit is already in the series *LC* form shown in Figure 4.40. No further simplification is possible. Thus, the Thévenin equivalent network seen by the *LC* load is $V_T = V_S = 25$ V and $R_T = R = 5$ kΩ. Refer to Figure 4.40, apply KVL around the loop, and use the *i*-*v* relationship for an inductor to find:

$$V_T - v_C - i_L R_T - v_L = V_T - v_C - i_L R_T - L\frac{di_L}{dt} = 0$$

Also apply KCL at a closed boundary surrounding the resistor and use the *i*-*v* relationship for the capacitor to find:

$$i_L = i_C = C\frac{dv_C}{dt}$$

Use the KCL equation to plug in for i_L in the KVL equation to find:

$$V_T - v_C - R_T C\frac{dv_C}{dt} - LC\frac{d^2v_C}{dt^2} = 0$$

or

$$LC\frac{d^2v_C}{dt^2} + R_T C\frac{dv_C}{dt} + v_C = V_T$$

Notice that the Norton equivalent resistance seen by the inductor with the capacitor acting as an open-circuit is $R_N \to \infty$ such that $L/R_N \to 0$. Thus, the coefficient of the first derivative term is $R_T C + L/R_N = R_T C$. Also note that the right side of the differential equation is $v_C(\infty) = V_S = V_T$.

Step 3: **Solve for ω_n and ζ for $t > 0$:** Compare the second-order ODE in standard form to its generalized form (see equation 4.25) to find:

$$\frac{1}{\omega_n^2} = LC = 10^{-6}\ \text{sec}^2 \qquad \text{and} \qquad \frac{2\zeta}{\omega_n} = R_T C = 5 \times 10^{-3}\ \text{sec}$$

Solve these two equations to find $\omega_n = 10^3$ rad/sec and $\zeta = 2.5 > 1$. The transient response is *overdamped.*

Step 4: **The transient response $x_{tr}(t)$:** The differential equation in step 2 was found for v_C because it was the easiest variable for which to find a second-order ODE. However, since the transient response for v_C is overdamped, the same is true for every other variable in the circuit. Thus, for the overdamped case:

$$\begin{aligned} i_{L_{tr}}(t) &= e^{-\zeta\omega_n t}\left(\alpha_1 e^{\omega_n t\sqrt{\zeta^2-1}} + \alpha_2 e^{-\omega_n t\sqrt{\zeta^2-1}}\right) \\ &= e^{-2500t}\left(\alpha_1 e^{2291t} + \alpha_2 e^{-2291t}\right)\ \text{A} \qquad t \geq 0 \end{aligned}$$

Step 5: **The complete solution *x(t)*:** From step 1 $i_L(\infty) = 0$ A such that the complete solution for $i_L(t)$ is

$$i_L(t) = e^{-2500t}(\alpha_1 e^{2291t} + \alpha_2 e^{-2291t}) + 0\ \text{A} \qquad t \geq 0$$

Step 6: **Solve for the unknown constants α_1 and α_2:** The initial conditions are $v_C(0^+) = v_C(0^-) = 5$ V and $i_L(0^+) = i_L(0^-) = 0$ A. From step 5:

$$i_L(0^+) = \alpha_1 + \alpha_2 = 0\ \text{A}$$

Differentiate $i_L(t)$ and set $t = 0^+$ to find:

$$\frac{di_L(0^+)}{dt} = -2500 i_L(0^+) + 2291(\alpha_1 - \alpha_2) = 2291(\alpha_1 - \alpha_2)\ \text{A/sec}$$

Also, from the first-order KVL equation found in step 2:

$$\frac{di_L(0^+)}{dt} = \frac{1}{L}[V_T - v_C(0^+) - i_L(0^+)R_T] = 20\ \text{A/sec}$$

Thus:

$$\alpha_1 + \alpha_2 = 0\ \text{A} \qquad \text{and} \qquad 2291(\alpha_1 - \alpha_2) = 20\ \text{A/sec}$$

Solve this system of two equations in two unknowns to find:

$$\alpha_1 = 4.36 \times 10^{-3}\ \text{A} \qquad \text{and} \qquad \alpha_2 = -4.36 \times 10^{-3}\ \text{A}$$

Finally, plug in to find:

$$i_L(t) = e^{-2500t}[4.36 \times 10^{-3}\, e^{2291t} - 4.36 \times 10^{-3}\, e^{-2291t}] + 0 \text{ A}$$
$$= 4.36 \times 10^{-3}\, e^{-2500t}[e^{2291t} - e^{-2291t}] \text{ A} \qquad t \geq 0$$

Page 305

A plot of the complete solution and of its components is given in Figure 4.48.

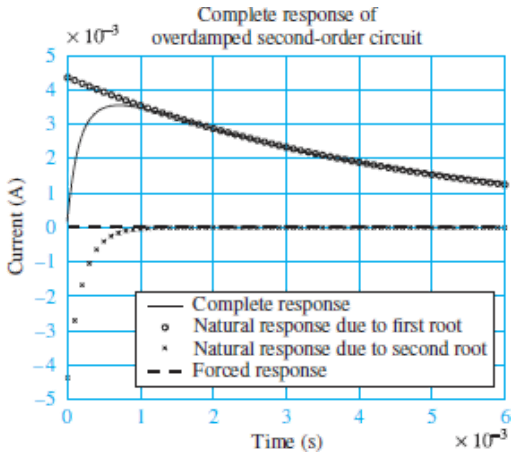


Figure 4.48 Complete response of overdamped second-order circuit

EXAMPLE 4.16 Complete Response of Critically Damped Parallel *LC* Circuit

Problem

Determine the complete response for the voltage v_C shown in Figure 4.49.

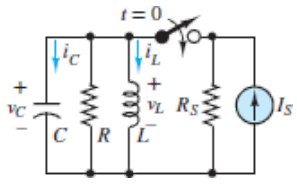


Figure 4.49

Solution

Known Quantities: I_S; R; R_S; C; L.

Find: The complete response of the differential equation in v_C describing the circuit in Figure 4.49.

Schematics, Diagrams, Circuits, and Given Data: $I_S = 5$ A; $R = R_S = 500\ \Omega$; $C = 2$ μF; $L = 500$ mH.

Assumptions: The network is in a DC steady-state prior to $t = 0$.

Analysis: Refer to the Focus on Problem Solving box "Second-Order Transient Response" and the subsection on parallel LC circuits. The load is L in parallel with C.

Step 1: **DC steady-state responses.** For $t < 0$, the switch is open such that the current source is separated from the parallel RLC network, which is in a DC steady-state such that the inductor acts as a short-circuit and the capacitor acts as an open-circuit. The result is that $v_C(0^-) = 0$ V and $i_L(0^-) = 0$ A. For $t \to \infty$ the switch is closed and the capacitor acts as a DC open-circuit while the inductor acts as a DC short-circuit. Thus, all of the current I_S is through the inductor such that $i_L(\infty) = I_S$. Likewise, since no current is through either resistor the voltage across the two nodes is zero such that $v_C(\infty) = 0$.

Step 2: **Differential equation for $t > 0$:** With the switch closed the Norton equivalent network seen by the parallel L_C load is $R_N = R \,/\!/\, R_S = 250\Omega$ and $I_N = I_S = 5$ A such that Page 306the network is simplified to the form shown in Figure 4.39. Apply KCL at the upper node and use the i-v relationship for a capacitor to write:

$$I_N - \frac{v_C}{R_N} - i_C - i_L = I_N - \frac{v_C}{R_N} - C\frac{dv_C}{dt} - i_L = 0$$

Apply KVL around the right mesh in Figure 4.39 and use the i-v relationship for an inductor to write:

$$v_C = v_L = L\frac{di_L}{dt}$$

Use the KVL equation to plug in for v_C in the KCL equation to find:

$$I_N - \frac{L}{R_N}\frac{di_L}{dt} - LC\frac{d^2i_L}{dt^2} - i_L = 0$$

or

$$LC\frac{d^2 i_L}{dt^2} + \frac{L}{R_N}\frac{di_L}{dt} + i_L = I_N$$

Notice that the Thévenin equivalent resistance seen by the capacitor with the inductor acting as a short-circuit is $R_T = 0$ such that $R_T C = 0$. Thus, the coefficient of the first derivative term is $R_T C + L/R_N = L/R_N$. Also note that the right side of the differential equation is $i_L(\infty) = I_S = I_N$.

Since the left side of the differential equation is the same for every variable it is a simple matter to substitute v_C for i_L and replace the right side with $v_C(\infty)$ to find:

$$LC\frac{d^2 v_C}{dt^2} + \frac{L}{R_N}\frac{dv_C}{dt} + v_C = v_C(\infty) = 0$$

The differential equation for i_L was found first because it was the easier to do!

Step 3: **Solve for ω_n and ζ for $t > 0$:** Compare the second-order ODE in standard form to its generalized form (see equation 4.25) to find:

$$\frac{1}{\omega_n^2} = LC = 10^{-6}\ \text{sec}^2 \qquad \text{and} \qquad \frac{2\zeta}{\omega_n} = \frac{L}{R_N} = 2 \times 10^{-3}\ \text{sec}$$

Solve these two equations to find $\omega_n = 10^3$ rad/sec and $\zeta = 1$. The transient response is *critically damped.*

Step 4: **The transient response $\mathbf{x_{tr}(t)}$:** For the critically damped case:

$$v_{C_{tr}}(t) = e^{-\omega_n t}(\alpha_1 + \alpha_2 t) = e^{-1000t}(\alpha_1 + \alpha_2 t) \qquad t \geq 0$$

Step 5: **The complete solution $x(t)$:** From step 1, $v_C(\infty) = 0$ V such that the complete solution for $v_C(t)$ is

$$v_C(t) = e^{-1000t}(\alpha_1 + \alpha_2 t) + 0\ \text{V} \qquad t \geq 0$$

Step 6: **Solve for the unknown constants α_1 and α_2:** The initial conditions are $v_C(0^+) = v_C(0^-) = 0$ V and $i_L(0^+) = i_L(0^-) = 5$ A. From step 5:

$$v_C(0^+) = \alpha_1 = 0\ \text{V}$$

Differentiate $v_C(t)$ and set $t = 0^+$ to find:

$$\frac{dv_C(0^+)}{dt} = -1000\, v_C(0^+) + \alpha_2 = \alpha_2\ \text{V/sec}$$

Also, from the first-order KCL equation found in step 2:

$$\frac{dv_C(0^+)}{dt} = \frac{1}{C}\left[I_N - \frac{v_C(0^+)}{R_N} - i_L(0^+)\right] = \frac{I_N}{C} = 2.5 \times 10^6 \text{ V/sec}$$

Thus:

$$\alpha_2 = 2.5 \times 10^6 \text{ V/sec}$$

Page 307

Finally, plug in to find:

$$v_C(t) = 2.5 \times 10^6 te^{-1000t} \text{ V} \qquad t \geq 0$$

A plot of the complete solution and of its components is given in Figure 4.50. Note that the maximum value occurs when $\omega_n t = 1$.

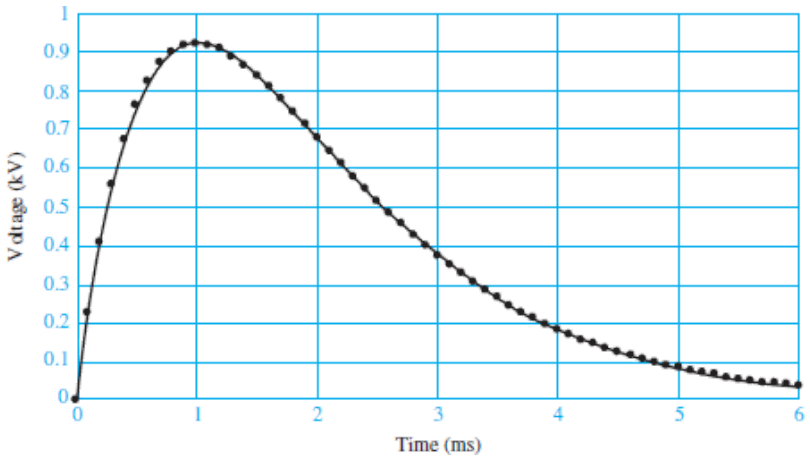


Figure 4.50 Complete response of critically damped second-order circuit

EXAMPLE 4.17 Complete Response of Underdamped Series *LC* Circuit

Problem

Determine the complete response for the current i_L shown in Figure 4.51.

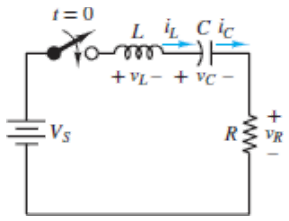


Figure 4.51

Solution

Known Quantities: V_S; R; C; L.

Find: The complete response for the current i_L shown in Figure 4.51.

Schematics, Diagrams, Circuits, and Given Data: $V_S = 12$ V; $R = 200\ \Omega$; $C = 10\ \mu$F; $L = 0.5$ H.

Assumptions: The capacitor has an initial charge such that $v_C(0^-) = v_C(0^+) = 2$ V.

Analysis: Refer to the Focus on Problem Solving box "Second-Order Transient Response" and the subsection on series LC circuits. The load is L in series with C.

Step 1: **DC steady-state responses.** For $t < 0$ the switch is open such that $i_C = i_L = 0$ A. The capacitor has an initial charge such that its voltage is $v_C = 2$ V for $t < 0$. For $t \to \infty$ the switch is closed and the capacitor acts as a DC open-circuit while the inductor acts as a DC short-circuit. Thus, $i_C = i_L = 0$, $v_L = 0$ and the voltage across the resistor must also be zero due to Ohm's law. By KVL, $v_C(\infty) = V_S = 12$ V.

Step 2: **Differential equation for $t > 0$:** With the switch closed the circuit is already in the series LC form shown in Figure 4.40. No further simplification is possible. Thus, the Thévenin equivalent network seen by the LC load is $V_T = V_S = 12$ V and $R_T = R = 200\Omega$. Refer to Figure 4.40, apply KVL around the loop and use the i-v relationship for an inductor to find:

$$V_T - v_C - i_L R_T - v_L = V_T - v_C - i_L R_T - L\frac{di_L}{dt} = 0$$

Page 308

Also apply KCL at a closed boundary surrounding the resistor and use the i-v relationship for the capacitor to find:

$$i_L = i_C = C\frac{dv_C}{dt}$$

Use the KCL equation to plug in for iL in the KVL equation to find:

$$V_T - v_C - R_T C\frac{dv_C}{dt} - LC\frac{d^2v_C}{dt^2} = 0$$

or

$$LC\frac{d^2v_C}{dt^2} + R_T C\frac{dv_C}{dt} + v_C = V_T$$

Notice that the Norton equivalent resistance seen by the inductor with the capacitor acting as an open-circuit is $R_N \to \infty$ such that $L/R_N \to 0$. Thus, the coefficient of the first derivative term is $R_TC + L/R_N = R_TC$. Also note that the right side of the differential equation is $v_C(\infty) = V_S = V_T$.

Since the left side of the differential equation is the same for every variable it is a simple matter to substitute i_L for v_C and replace the right side with $iL(\infty)$ to find:

$$LC\frac{d^2i_L}{dt^2} + R_T C\frac{di_L}{dt} + i_L = i_L(\infty) = 0$$

The differential equation for v_C was found first because it was the easier to do!

Step 3: **Solve for ω_n and ζ for $t > 0$:** Compare the second-order ODE in standard form to its generalized form (see equation 4.25) to find:

$$\frac{1}{\omega_n^2} = LC = 5 \times 10^{-6}\ \text{sec}^2 \qquad \text{and} \qquad \frac{2\zeta}{\omega_n} = R_T C = 2 \times 10^{-3}\ \text{sec}$$

Solve these two equations to find $\omega_n = \sqrt{20} \times 10^2$ and $\zeta = \sqrt{5}/5 < 1$. The transient response is *underdamped.*

Step 4: **The transient response $\mathbf{x_{tr}}$ (t):** The transient response is underdamped, thus:

$$i_{L_{tr}}(t) = e^{-\zeta\omega_n t}[\alpha_1 \sin(\omega_d t) + \alpha_2 \cos(\omega_d t)] \qquad t \geq 0$$

$$\omega_d = \omega_n\sqrt{1 - \zeta^2} = 400\ \text{rad/sec}$$

or

$$i_{L_{tr}}(t) = e^{-200t}[\alpha_1 \sin(400\ t) + \alpha_2 \cos(400\ t)] \qquad t \geq 0$$

Step 5: **The complete solution $x(t)$:** From step 1 $i_L(\infty) = 0$ such that the complete solution for $i_L(t)$ is the same as the transient solution.

$$i_L(t) = e^{-200t}[\alpha_1 \sin(400\ t) + \alpha_2 \cos(400\ t)] \qquad t \geq 0$$

Step 6: **Solve for the unknown constants α_1 and α_2:** The initial conditions are $v_C(0^+) = v_C(0^-) = 2$ V and $i_L(0^+) = i_L(0^-) = 0$ A. From step 5:

$$i_L(0^+) = \alpha_2 = 0 \text{ A}$$

Differentiate $i_L(t)$ and set $t = 0^+$ to find:

$$\frac{di_L(0^+)}{dt} = -200\, i_L(0^+) + 400\,[\alpha_1 \cos(0^+) - \alpha_2 \sin(0^+)] = 400\,\alpha_1 \text{ A/sec}$$

Also, from the first-order KVL equation found in step 2:

$$\frac{di_L(0^+)}{dt} = \frac{1}{L}[V_T - v_C(0^+) - i_L(0^+)R_T] = 20 \text{ A/sec}$$

Page 309

Thus:

$$\alpha_1 = 0.05 \text{ A/sec}$$

Finally, plug in to find:

$$i_L(t) = 0.05\, e^{-200t} \sin(400\, t) \text{ A} \qquad t \geq 0$$

A plot of the complete solution and of its components is given in Figure 4.52.

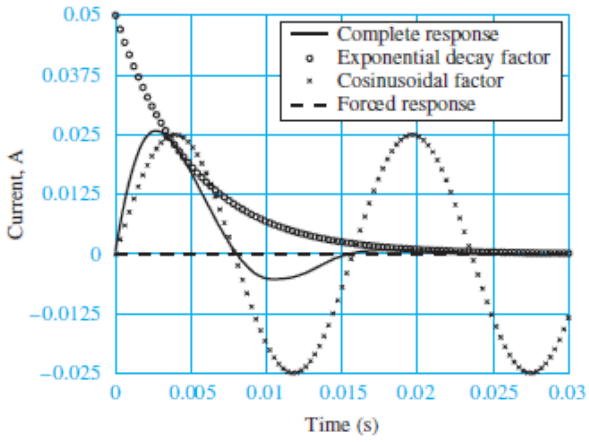


Figure 4.52 Complete response of underdamped second-order circuit

EXAMPLE 4.18 Analysis of Nonseries, Nonparallel *LC* Circuit

Problem

Assume the circuit shown in Figure 4.53 is in DC steady-state for $t < 0$. The switch closes at $t = 0$. Find the differential equations for the voltage v_C across the capacitor and the current i_L through the inductor for $t > 0$.

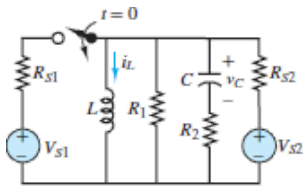


Figure 4.53

Solution

Known Quantities: V_{S1}; R_{S1}; V_{S2}; R_{S2}; R_1; R_2; L; C.

Find: For $t > 0$, find the differential equations for the voltage v_C across the capacitor and the current i_L through the inductor shown in Figure 4.53.

Assumptions: DC steady-state for $t < 0$.

Analysis: The critical difference between the circuit in this example and those in the previous examples is that the capacitor and inductor are neither in series nor in parallel. It is necessary to find two first-order differential equations in the state variables v_C and i_L to find the second-order differential equation in either state variable. Refer to the Focus on Problem Solving box "Second-Order Transient Response."

Step 1: **DC steady-state responses:** For $t < 0$ the switch is open, and V_{S1} and R_{S1} are disconnected from the rest of the circuit. Assuming a DC steady-state the inductor acts as a short-circuit and the capacitor acts as an open-circuit such that the voltage across R_1 is Page 310zero and the current through R_2 is zero. Thus, KVL requires that the voltage across R_{S2} is V_{S2} and all of the current through R_{S2} is also through the inductor to complete a circuit. Ohm's law requires that the voltage across R_2 is zero

such that KVL around a loop containing the inductor and capacitor results in $v_C = 0$. Therefore, the values of the state variables just prior to the switch event are

$$i_L(0^-) = \frac{V_{S2}}{R_{S2}} \quad \text{and} \quad v_C(0^-) = 0$$

For $t > 0$ the switch is closed and the circuit can be redrawn in the source-load perspective as shown in Figure 4.54. The two Thévenin sources can be transformed into Norton sources as shown in Figure 4.55. The two parallel current sources can be summed to produce $I_N = I_{S1} + I_2$, and the three parallel resistors can be replaced with an equivalent resistance $R_N = R_{S1} /\!/ R_{S2} /\!/ R_1$ as shown in Figure 4.56. As $t \to \infty$ the inductor and capacitor again act as short- and open-circuits, respectively, such that:

$$i_L(\infty) = I_N \quad \text{and} \quad v_C(\infty) = 0 \text{ V}$$

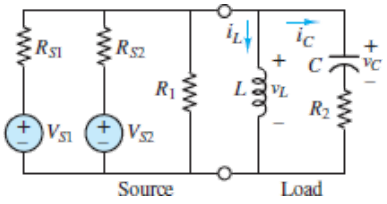


Figure 4.54

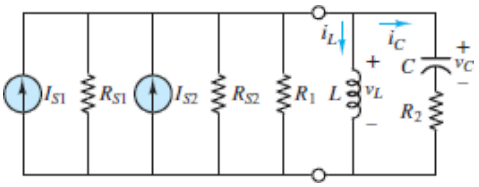


Figure 4.55

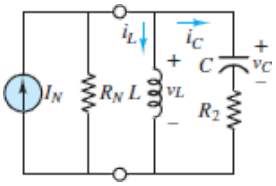


Figure 4.56

Step 2: **Differential equation for $t > 0$:** Refer to Figure 4.56. The source network is already in the form of a Norton source so no further simplification is possible. Apply KCL at the top node and the *i-v* relationships for the inductor and capacitor to find:

$$I_N - \frac{v_L}{R_N} - i_L - i_C = I_N - \frac{L}{R_N}\frac{di_L}{dt} - i_L - C\frac{dv_C}{dt} = 0 \qquad \text{KCL}$$

Also apply KVL around the right-most mesh and the *i*-*v* relationships for the inductor and capacitor to find:

$$v_L - v_C - i_C R_2 = L\frac{di_L}{dt} - v_C - R_2 C\frac{dv_C}{dt} = 0 \qquad \text{KVL}$$

These two first-order differential equations in the two state variables i_L and v_C can be combined to find a second-order differential equation in one state variable. One way to accomplish this task is to multiply the KCL equation by R_2 and subtract the result from the KVL equation to yield:

$$v_C = L\frac{R_N + R_2}{R_N}\frac{di_L}{dt} + i_L R_2 - I_N R_2$$

Differentiate both sides of this equation to find:

$$\frac{dv_C}{dt} = L\frac{R_N + R_2}{R_N}\frac{d^2 i_L}{dt^2} + R_2\frac{di_L}{dt}$$

Substitute this result into the KCL equation to yield a second-order differential equation in standard form.

$$LC\frac{R_N + R_2}{R_N}\frac{d^2 i_L}{dt^2} + \left(R_2 C + \frac{L}{R_N}\right)\frac{di_L}{dt} + i_L = I_N$$

Notice that the coefficient of *iL* is unity (standard form) and the coefficient of the first-order derivative is the sum of the time constants $R_T C + L/R_N$ associated with the capacitor and inductor. Here, R_T is the Thévenin equivalent resistance seen by capacitor Page 311with the inductor acting as a short-circuit and R_N is the Norton equivalent resistance seen by the inductor with the capacitor acting as an open-circuit. (Confirm that R_T seen by the capacitor is R_2.) Also, the right-hand side is $i_L(\infty) = I_N$.

Likewise, the second-order differential equation in v_C must be

$$LC\frac{R_N + R_2}{R_N}\frac{d^2 v_C}{dt^2} + \left(R_2 C + \frac{L}{R_N}\right)\frac{dv_C}{dt} + v_C = v_C(\infty) = 0$$

Step 3: **Solve for ω_n and ζ for $t > 0$:** Compare the second-order ODE in standard form to its generalized form (see equation 4.25) to find:

$$\frac{1}{\omega_n^2} = LC\frac{R_N + R_2}{R_N} \qquad \text{and} \qquad \frac{2\zeta}{\omega_n} = R_2C + \frac{L}{R_N}$$

These two equations can be reexpressed as:

$$\omega_n = \sqrt{\frac{1}{LC}}\sqrt{\frac{R_N}{R_N + R_2}} \qquad \text{and} \qquad \zeta = \frac{\omega_n}{2}\left(R_2C + \frac{L}{R_N}\right)$$

Step 4: **The transient response x_{tr} (t):** The form (overdamped, critically damped, underdamped) of the transient solution depends upon the value of ζ, which itself depends upon the values of the various circuit elements.

Step 5: **The complete solution $x(t)$:** The complete solution is the sum of the transient solution and the long-term DC steady-state value.

$$i_L(t) = i_{L_{tr}}(t) + i_L(\infty) \qquad \text{and} \qquad v_C(t) = v_{C_{tr}}(t) + v_C(\infty)$$

Regardless of the form of the transient solution, the complete solution will contain two unknown constants α_1 and α_2.

Step 6: **Solve for the unknown constants α_1 and α_2:** The initial conditions are

$$v_C(0^+) = v_C(0^-) = 0 \text{ V} \qquad \text{and} \qquad i_L(0^+) = i_L(0^-) = \frac{V_{S2}}{R_{S2}} \text{ A}$$

In general, to solve for the unknown constants it is necessary to find two linearly independent algebraic equations in them. The first such equation is found by simply setting $t = 0^+$ in the complete solution and setting the result equal to the initial condition for the variable. To find a second equation take the derivative of the complete solution and evaluate it at $t = 0^+$. Then, use the first-order KCL and KVL differential equations found in step 2 to find another expression of v_C or i_L. Evaluate it at $t = 0^+$ to find an expression in terms of $i_L(0^+)$ and $v_C(0^+)$.

In this particular example, one of the results found in step 2 is

$$v_C = L\frac{R_N + R_2}{R_N}\frac{di_L}{dt} + i_L R_2 - I_N R_2$$

This equation can be rearranged and evaluated at $t = 0^+$ to find:

$$\left.\frac{di_L(t)}{dt}\right|_{t=0^+} = \frac{[I_N R_2 - i_L(0^+)R_2 + v_C(0^+)]}{L}\frac{R_N}{R_N + R_2}$$

Comments: Recall that the values of the unknown constants and the long-term steady state are, in general, different for different variables. However, all variables in a circuit share the same natural frequency ω_n and dimensionless damping coefficient ζ.

That is, the left side of the second-order differential equation is the same for all variables. Also, keep in mind that the initial conditions for any variable and its derivative must be related to the initial conditions on the state variables since only the state variables are guaranteed to be continuous across the transient event.

EXAMPLE 4.19 Transient Response of Automotive Ignition Circuit

Problem

The circuit shown in Figure 4.57 is a simplified but realistic representation of an automotive ignition system. The circuit includes an **automotive battery,** a transformer **(ignition coil),** a capacitor (known as a *condenser* in old-fashioned automotive parlance), and a switch. The switch is usually an electronic switch (e.g., a transistor—see Chapter 9) and can be treated as an ideal switch. The circuit on the left represents the ignition circuit immediately after the electronic switch has closed, following a spark discharge. Thus, one can assume that no energy is stored in the inductor prior to the switch closing, say at $t = 0$. Furthermore, no energy is stored in the capacitor, as the short-circuit (closed switch) across it would have dissipated any charge in the capacitor. The primary winding of the ignition coil (left-hand side inductor) is then given a suitable length of time to build up stored energy, and then the switch opens, say at $t = \Delta t$, leading to a rapid voltage buildup across the secondary winding of the coil (right-hand side inductor). The voltage rises to a very high value because of two effects: an *inductive voltage kick* due to the fact that a large change in the *rate* of current through the coil requires a large voltage (see the constitutive i-v relation for an inductor) and the voltage multiplying effect of the transformer. The result is a very short high-voltage transient (reaching thousands of volts), which causes a spark to be generated across the spark plug.

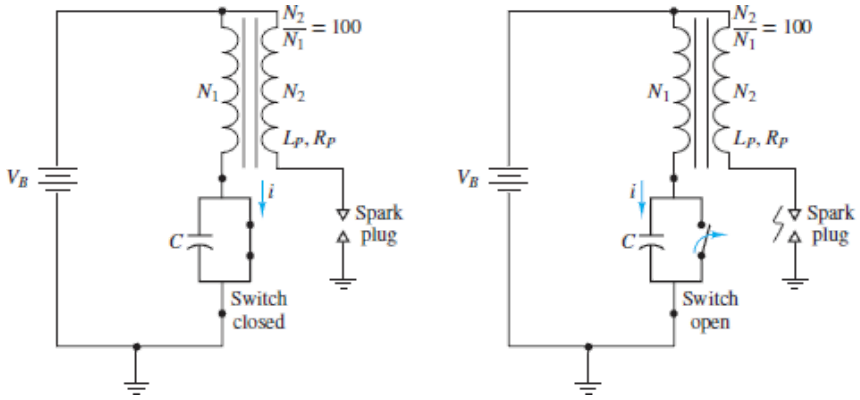


Figure 4.57

Solution

Known Quantities: V_B; N_2/N_1; L_p; R_p; C.

Find: The ignition coil current $i(t)$ and the open-circuit voltage across the spark plug $v_{OC}(t)$.

Schematics, Diagrams, Circuits, and Given Data: $V_B = 12$ V; $N_2/N_1 = 100$; $L_p = 5$ mH; $R_p = 2\ \Omega$; $C = 10\ \mu$F.

Assumptions: The switch has been open for a long time before it closes at $t = 0$. The switch opens again at $t = \Delta t$.

Analysis: Initially, the switch is open and no energy is stored in either the inductor or the capacitor. Then, the switch is closed, as shown in Figure 4.58. When the switch is closed, a first-order RL circuit is formed by the primary coil inductance L_P and resistance R_P. Page 313The solution of this circuit gives the initial condition that will be in effect when the switch is opened again.

$$i_L(t) = i_L(\infty) + [i_L(0) - i_L(\infty)]e^{-t/\tau} \qquad 0 \le t < \Delta t$$
$$i_L(t) = 6(1 - e^{-t/2.5\times10^{-3}})$$

where

$$i_L(\infty) = \frac{V_B}{R_p} = 6\ \text{A} \qquad \text{Long term first-order steady-state}$$
$$\tau = \frac{L_p}{R_p} = 2.5 \times 10^{-3}\ \text{s} \qquad \text{First-order time constant}$$

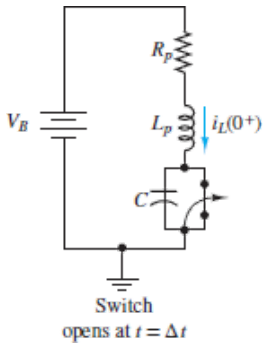


Figure 4.58 When the switch is opened at $t = \Delta t$, the capacitor is no longer bypassed, resulting in a second-order transient.

The switch remains closed until $t = \Delta t = 12.5$ ms $= 5\tau$. At time Δt, the value of the inductor current will be

$$i_L(t = \Delta t) = 6(1 - e^{-5}) = 5.96 \text{ A}$$

that is, the current reaches roughly 99 percent of its long term steady-state value in five time constants.

Now, when the switch opens at $t = \Delta t$, the result is a series LC circuit.

Step 1: Steady-state response for $t > \Delta t$. After the switch has been open for a long time, the capacitor acts as an open-circuit and the inductor acts as a short-circuit. In this case, all the source voltage will appear across the capacitor, and, of course, the inductor current is zero: $i_L(\infty) = 0$ A, $v_C(\infty) = V_S = 12$ V.

Step 2: Differential equation. The differential equation for the series circuit can be obtained by KCL:

$$L_p C\frac{d^2 i_L}{dt^2} + R_p C\frac{di_L}{dt} + i_L = C\frac{dV_B}{dt} = 0 \qquad t > \Delta t$$

Step 3: Solve for ω_n and ζ.

$$\omega_n = \sqrt{\frac{1}{L_p C}} = 4{,}472 \text{ rad/s}$$

$$\zeta = R_p C\frac{\omega_n}{2} = \frac{R_p}{2}\sqrt{\frac{C}{L_p}} = 0.0447$$

Thus, the ignition circuit is underdamped.

Step 4: The complete solution. The long term steady-state inductor current is zero so the transient solution is also the complete solution for $t > \Delta t$.

$$i_L(t) = \alpha_1 e^{\left(-\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)} + \alpha_2 e^{\left(-\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)} \qquad t > \Delta t$$

Step 5: Solve for the constants α_1 and α_2. Finally, solve for the initial conditions to evaluate the constants α_1 and α_2. At $t = \Delta t = 12.5$ ms, $i_L(\Delta t^-) = 5.96$ A and $v_C(\Delta t^-) = 0$ V. Since the differential equation is in the variable i_L, the two needed initial conditions are $i_L(\Delta t^+)$ and $di_L(\Delta t^+)/dt$. The first initial condition is found directly from the solution $i_L(\Delta t^+) = i_L(\Delta t^-) = 5.96$ A. The second initial condition is found by applying KVL at $t = \Delta t^+$:

$$V_B - v_C(\Delta t^+) - Ri_L(\Delta t^+) - L\frac{di_L(\Delta t^+)}{dt} = 0$$

$$\frac{di_L(\Delta t^+)}{dt} = \frac{V_B}{L} - \frac{v_C(\Delta t^+)}{L} - \frac{R}{L}i_L(\Delta t^+) = \frac{12}{5 \times 10^{-3}} - 0 - \frac{2 \times 5.96}{5 \times 10^{-3}}$$

$$= 16.0 \text{ A/s}$$

The first initial condition at $t = \Delta t^+$ yields

$$i_L(\Delta t^+) = \alpha_1 e^0 + \alpha_2 e^0 = 5.96 \text{ A}$$

$$\alpha_1 = 5.96 - \alpha_2$$

Page 314

The second initial condition at $t = \Delta t^+$ is evaluated as follows:

$$\frac{di_L(\Delta t^+)}{dt} = \left(-\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2}\right)\alpha_1 e^0 + \left(-\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}\right)\alpha_2 e^0$$

Substituting $\alpha_1 = 5.96 - \alpha_2$, obtain:

$$\frac{di_L(\Delta t^+)}{dt} = \left(-\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2}\right)\alpha_1$$

$$+ \left(-\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}\right)(5.96 - \alpha_1) = 16.0 \text{ A/s}$$

$$2\left(j\omega_n\sqrt{1-\zeta^2}\right)\alpha_1 + 5.96\left(-\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}\right) = 16.0 \text{ V}$$

$$\alpha_1 = \frac{16.0 - 5.96\left(-\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}\right)}{2j\omega_n\sqrt{1-\zeta^2}} = 2.98 - j0.135 \text{ A}$$

$$\alpha_2 = 5.96 - \alpha_1 = 2.98 + j0.135 \text{ A}$$

The final complete solution is:

$$
\begin{aligned}
i_L(t) &= (2.98 - j0.135)\,e^{\left(-\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)} \\
&\quad + (2.98 + j0.135)\,e^{\left(-\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)} \qquad t > \Delta t \\
i_L(t) &= 2.98\,e^{(-\zeta\omega_n)(t-\Delta t)}\left(e^{\left(+j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)} + e^{\left(-j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)}\right) \\
&\quad - j0.135e^{(-\zeta\omega_n)(t-\Delta t)}\left(e^{\left(+j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)} - e^{\left(-j\omega_n\sqrt{1-\zeta^2}\right)(t-\Delta t)}\right) \\
&= 2.98e^{(-200)(t-\Delta t)}\left(e^{(+j4{,}468)(t-\Delta t)} + e^{(-j4{,}468)(t-\Delta t)}\right) \\
&\quad - j0.135e^{(-200)(t-\Delta t)}\left(e^{(+j4{,}468)(t-\Delta t)} - e^{(-j4{,}468)(t-\Delta t)}\right) \\
&= 5.96e^{(-200)(t-\Delta t)}\cos[4{,}468(t-\Delta t)] + 0.27e^{(-200)(t-\Delta t)}\sin[4{,}468(t-\Delta t)] \qquad t > \Delta t
\end{aligned}
$$

A plot of the inductor current for $-10 \le t \le 50$ ms is shown in Figure 4.59. Notice the initial first-order transient at $t = 0$ followed by a second-order transient at $t = 12.5$ ms.

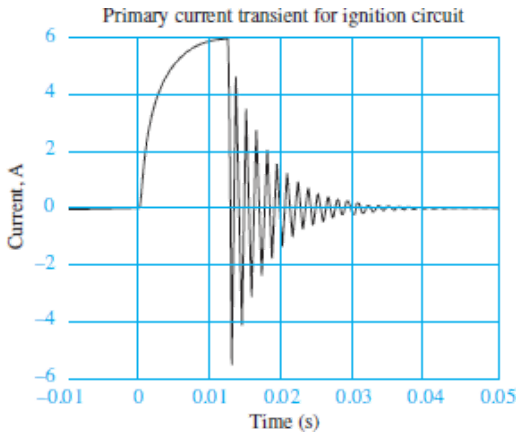


Figure 4.59 Transient current response of ignition current

To compute the primary voltage, differentiate the inductor current and multiply by L; to determine the secondary voltage, which is that applied to the spark plug, the 1:100 transformer Page 315increases the secondary voltage by a factor of 100 relative to the primary voltage.[1] Thus, the expression for the secondary voltage is:

$$
\begin{aligned}
v_{\text{spark plug}} &= 100 \times L\frac{di_L}{dt} = 0.5 \times \frac{d}{dt}\{5.96\,e^{(-200)(t-\Delta t)}\cos[4{,}468(t-\Delta t)] \\
&\quad + 0.27\,e^{(-200)(t-\Delta t)}\sin[4{,}468(t-\Delta t)]\} \\
&= 0.5 \times \{5.96(-200)\,e^{(-200)(t-\Delta t)}\cos[4{,}468(t-\Delta t)] \\
&\quad + 5.96\,e^{(-200)(t-\Delta t)}(-4{,}468)\sin[4{,}468(t-\Delta t)]\} \\
&\quad + 0.5 \times \{0.27(-200)\,e^{(-200)(t-\Delta t)}\sin[4{,}468(t-\Delta t)] \\
&\quad - 0.27\,e^{(-200)(t-\Delta t)}4{,}468\cos[4{,}468(t-\Delta t)]\}
\end{aligned}
$$

The voltage near $t = \Delta t$ will generate the spark. Evaluating at $t = \Delta t$:

$$v_{\text{spark plug}}(t = 0) = 0.5 \times [5.96(-200)] - 0.5 \times (0.27 \times 4{,}468) = -1{,}199.18 \text{ V}$$

It is important to note that $v_{\text{spark plug}}$ oscillates rapidly and that its first peak voltage occurs near 0.32 ms at an approximate value of –12,550 V. A plot of the inductor voltage starting at the time when the switch is opened is shown in Figure 4.60. The result of the switching is a succession of large (negative) voltage spikes, capable of generating a series of sparks across the plug gap. However, once a single spark is generated, the entire dynamics of the spark plug changes since the spark itself acts as a low-resistance ionized path to ground.

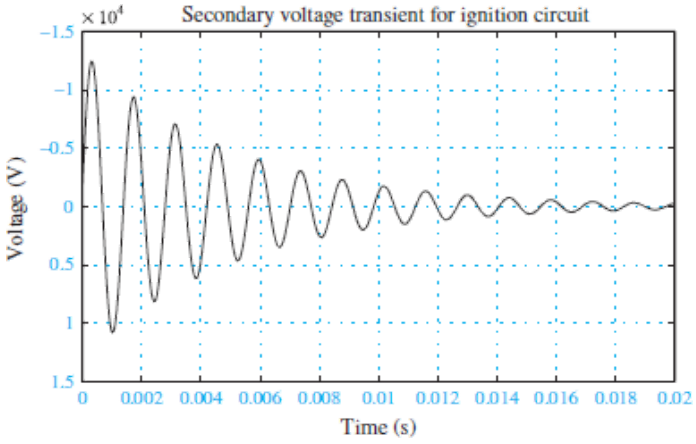


Figure 4.60 Secondary ignition voltage response

CHECK YOUR UNDERSTANDING

For what value of R_N in Example 4.14 will the circuit response become critically damped?

Answer: $R = 158.1\ \Omega$

Page 316

CHECK YOUR UNDERSTANDING

If the inductance in Example 4.17 is reduced to one-half of its original value (from 0.5 to 0.25 H), for what range of values of R will the circuit be underdamped?

Answer: $R \leq 316\ \Omega$

Conclusion

Chapter 4 has focused on the solution of first- and second-order differential equations for the case of DC switched transients, and it has presented a number of analogies between electric circuits and other physical systems, such as thermal, hydraulic and mechanical.

While many other forms of excitation exist, turning a DC supply on and off is a very common occurrence in electrical, electronic, and electromechanical systems. Further, the methods discussed in this chapter can be readily extended to the solution of more general problems.

A thorough study of this chapter should result in the acquisition of the following learning objectives:

1. *Write differential equations for circuits containing inductors and capacitors.* This process involves the application of KVL and/or KCL to produce first-order differential equations and the use of constitutive i-v relationships for inductors and capacitors to produce differential equations in the state variables.
2. *Determine the DC steady-state solution of circuits containing inductors and capacitors.* The DC steady-state solution of any differential equation can be easily obtained by setting the derivative terms equal to zero. Alternatively, the DC steady-state response for any circuit variable can be acquired directly from the circuit since an inductor acts as a short-circuit and a capacitor acts as an open-circuit under DC conditions.
3. *Write the differential equation of first-order circuits in standard form, and determine the complete solution of first-order circuits excited by switched DC sources.* First-order systems are most commonly described by way of two constants: the DC gain and the time constant. You have learned how to recognize these constants, how to compute the initial and final conditions, and how to write the complete solution of all first-order circuits almost by inspection.
4. *Write the differential equation of second-order circuits in standard form, and determine the complete solution of second-order circuits excited by switched DC sources.* Second-order circuits are described by three constants: the DC gain, the natural frequency, and the dimensionless damping coefficient. While the method

for obtaining the complete solution for a second-order circuit is logically the same as that used for a first-order circuit, the details are more involved.

HOMEWORK PROBLEMS

Section 4.2: Elements of Transient Problem Solving

4.1 Write the differential equations for $t > 0$ for i_L and v_3 in Figure P4.21. How are they related?

4.2 Write the differential equation for $t > 0$ for v_C in Figure P4.23.

4.3 Write the differential equation for $t > 0$ for i_C in Figure P4.27.

4.4 Write the differential equation for $t > 0$ for i_L in Figure P4.29.

4.5 Write the differential equation for $t > 0$ for v_C in Figure P4.32.

Page 317

4.6 Write the differential equations for $t > 0$ for i_C and v_3 in Figure P4.34. How are they related?

4.7 Write the differential equation for $t > 0$ for v_C in Figure P4.41. Assume S_1 is open, S_2 is closed and that $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, and $C_1 = C_2 = 4$ F.

4.8 Write the differential equation for $t > 0$ for i_C in Figure P4.47. Assume $V_S = 9$ V, $C = 1\ \mu$F, $R_S = 5\ \text{k}\Omega$, $R_1 = 10\ \text{k}\Omega$, and $R_2 = R_3 = 20\ \text{k}\Omega$.

4.9 Write the differential equation for $t > 0$ for i_L in Figure P4.49.

4.10 Write the differential equations for $t > 0$ for i_L and v_1 in Figure P4.52. How are they related? Assume $L_1 = 1$ H and $L_2 = 5$ H.

4.11 Determine the initial and final conditions on i_L and v_3 in Figure P4.21.

4.12 Determine the initial and final conditions on v_C in Figure P4.23.

4.13 Determine the initial and final conditions on i_C in Figure P4.27.

4.14 Determine the initial and final conditions on i_L in Figure P4.29.

4.15 Determine the initial and final conditions on v_C in Figure P4.32.

4.16 Determine the initial and final conditions on i_C and v_3 in Figure P4.34.

4.17 Determine the initial and final conditions on v_C in Figure P4.41. Assume S_1 is always open and S_2 is closed at $t = 0$.

4.18 Determine the initial and final conditions on i_C in Figure P4.47. Assume $V_S = 9$ V, $C = 1$ μF, $R_S = 5$ kΩ, $R_1 = 10$ kΩ, and $R_2 = R_3 = 20$ kΩ.

4.19 Determine the initial and final conditions on i_L in Figure P4.49.

4.20 Determine the initial and final conditions on i_L and v_1 in Figure P4.52. Assume $L_1 = 1$ H and $L_2 = 5$ H.

4.21 At $t = 0^-$, just before the switch is opened, the current through the inductor in Figure P4.21 is $i_L = 140$ mA. Is this value the same as that for DC steady-state? Was the circuit in steady state just before the switch was opened? Assume $VS =$ 10 V, $R_1 = 1$ kΩ, $R_2 = 5$ kΩ, $R_3 = 2$ kΩ, and $L = 1$ mH.

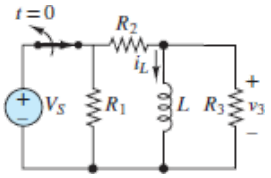


Figure P4.21

4.22 For $t < 0$, the circuit shown in Figure P4.22 is at DC steady-state. The switch is thrown at $t = 0$.

$VS1 = 35$ V$VS2 = 130$ V

$V_{S1} = 35$ V	$V_{S2} = 130$ V
$C = 11$ μF	$R_1 = 17$ kΩ
$R_2 = 7$ kΩ	$R_3 = 23$ kΩ

Determine the current through R_3 just after the switch is thrown at $t = 0^+$.

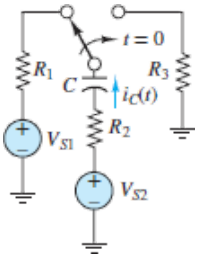


Figure P4.22

4.23 Determine the current i_C through the capacitor just before and just after the switch is closed in Figure P4.23. Assume steady-state conditions for $t < 0$. $V_1 = 15$ V, $R_1 = 0.5$ kΩ, $R_2 = 2$ kΩ, and $C = 0.4$ μF.

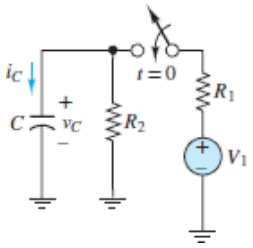


Figure P4.23

4.24 Assume the switch in Figure P4.23 has been closed for a very long time and then is opened. Determine the current i_C through the capacitor immediately after the switch is opened. $V_1 = 10$ V, $R_1 = 200$ mΩ, $R_2 = 5$ kΩ, and $C = 300$ μF.

4.25 Just before the switch is opened at $t = 0$ in Figure P4.21, assume the current through the inductor Page 318is $i_L = 1.5$ mA. Determine the voltage v_3 across R_3 immediately after the switch is opened. Assume $V_S = 12$ V, $R_1 = 6$ kΩ, $R_2 = 6$ kΩ, $R_3 = 3$ kΩ, and $L = 0.9$ mH.

4.26 Assume that steady-state conditions exist in the circuit shown in Figure P4.26 for $t < 0$. Determine the current through the inductor immediately after the switch is thrown. Assume $L = 0.5$ H, $R_1 = 100$ kΩ, $R_S = 5$ Ω, and $V_S = 24$ V.

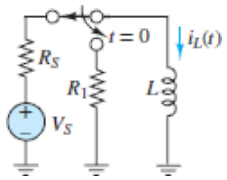


Figure P4.26

4.27 Assume that steady-state conditions exist in the circuit shown in Figure P4.27 for $t < 0$ and that $V_1 = 15$ V, $R_1 = 100$ Ω, $R_2 = 1.2$ kΩ, $R_3 = 400$ Ω, $C = 4.0$ μF. Determine the current i_C through the capacitor at $t = 0^+$, just after the switch is closed.

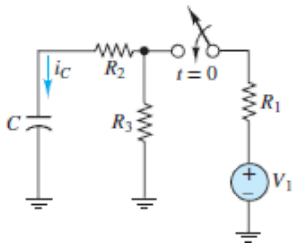


Figure P4.27

4.28 For $t > 0$, find the Norton equivalent network seen by the inductor in Figure P4.28. Use that result to determine the associated time constant. Assume:

$V_1 = 12$ V $\quad V_2 = 5$ V
$L = 3$ H $\quad R_1 = R_2 = 2\ \Omega$
$R_3 = 4\ \Omega$

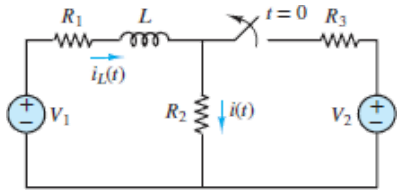


Figure P4.28

4.29 For $t > 0$, find the Norton equivalent network seen by the inductor in Figure P4.29. Use that result to determine the associated time constant. Assume:

$V_{S1} = 9$ V $\quad V_{S2} = 12$ V
$L = 120$ mH $\quad R_1 = 2.2\ \Omega$
$R_2 = 4.7\ \Omega$ $\quad R_3 = 18$ kΩ

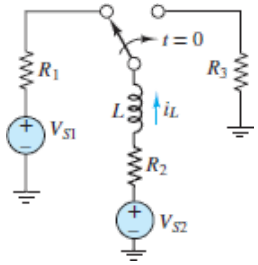


Figure P4.29

4.30 For $t > 0$, find the Thévenin equivalent network seen by the capacitor in Figure P4.30. Use that result to determine the associated time constant. Assume: $R_1 = 3\ \Omega, R_2 = 1\ \Omega, R_3 = 4\ \Omega, C = 0.2\text{F}, I_S = 3\text{A}$.

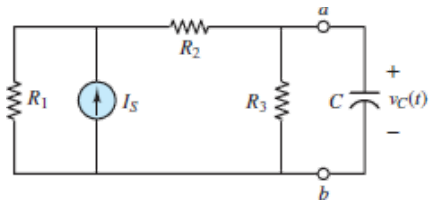


Figure P4.30

4.31 For $t > 0$, find the Thévenin equivalent network seen by the capacitor in Figure P4.31. Use that result to determine the associated time constant. Assume: $R_S = 8$ kΩ, $V_S = 40$ V, $C = 350$ μF, and $R = 24$ kΩ.

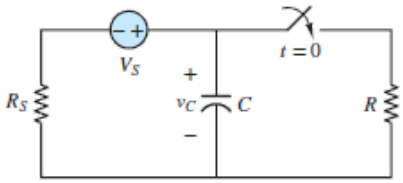


Figure P4.31

Section 4.3: First-Order Transient Analysis

4.32 Determine the voltage v_C across the capacitor shown in Figure P4.32 for $t > 0$. Assume a DC steady-state for $t < 0$ and:

$I_o = 17$ mA $\quad C = 0.55$ μF
$R_1 = 7$ kΩ $\quad R_2 = 3.3$ kΩ

Page 319

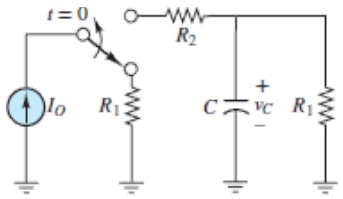


Figure P4.32

4.33 For $t < 0$, the circuit shown in Figure P4.29 is at steady state. The switch is thrown at $t = 0$. Determine the current i_L through the inductor for $t > 0$. Assume:

$V_{S1} = 9$ V $\quad V_{S2} = 12$ V
$L = 120$ mH $\quad R_1 = 2.2$ Ω
$R_2 = 4.7$ Ω $\quad R_3 = 18$ kΩ

4.34 For $t < 0$, the circuit shown in Figure P4.34 is at steady state. The switch is thrown at $t = 0$. Assume:

$$V_{S1} = 17\ \text{V} \qquad V_{S2} = 11\ \text{V}$$
$$R_1 = 14\ \text{k}\Omega \qquad R_2 = 13\ \text{k}\Omega$$
$$R_3 = 14\ \text{k}\Omega \qquad C = 70\ \text{nF}$$

Determine the

a. Current i_C through the capacitor for $t > 0$.

b. Voltage v_3 across R_3 for $t > 0$.

c. Time required for i_C and v_3 to change by 98 percent of their initial values at $t = 0^+$.

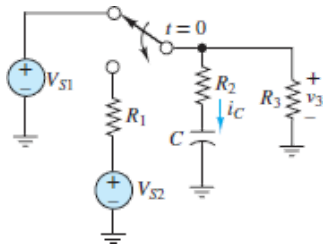


Figure P4.34

4.35 The circuit in Figure P4.35 is a simple model of an automotive ignition system. The switch models the "points" that switch electric power to the cylinder when the fuel-air mixture is compressed. R is the resistance across the gap between the electrodes of the spark plug.

$$V_G = 12\ \text{V} \qquad R_G = 0.37\ \Omega$$
$$R = 1.7\ \text{k}\Omega$$

Determine the value of L and R_1 so that the voltage across the spark plug gap just after the switch is changed is 23 kV and so that this voltage will change exponentially with a time constant $\tau = 13$ ms.

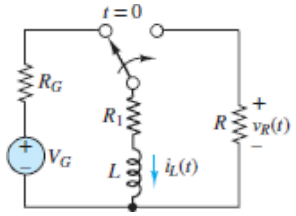


Figure P4.35

4.36 The inductor L in the circuit shown in Figure P4.36 is the coil of a relay. When the current i_L through the coil is equal to or greater than 2 mA, the relay is activated. Assume DC steady-state conditions at $t < 0$ and the following values:

$$V_S = 12 \text{ V}$$
$$L = 10.9 \text{ mH}$$
$$R_1 = 3.1 \text{ k}\Omega$$

Determine R_2 so that the relay activates 2.3 seconds after the switch is thrown.

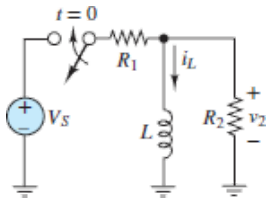


Figure P4.36

4.37 Determine the current i_C through the capacitor in Figure P4.37 for all time. Assume DC steady-state conditions for $t < 0$. Also assume: $V_1 = 10$ V, $C = 200$ μF, $R_1 = 300$ mΩ, and $R_2 = R_3 = 1.2$ kΩ.

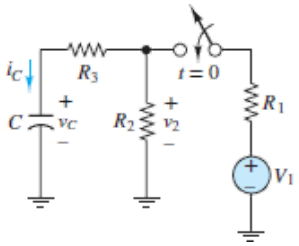


Figure P4.37

4.38 Determine the voltage v_L across the inductor in Figure P4.38 for all time. Assume DC steady-state conditions for $t < 0$. Also assume: $V_S = 15$ V, $L = 200$ mH, $R_S = 1$ Ω, and $R_1 = 20$ kΩ.

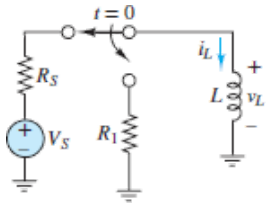


Figure P4.38

4.39 For $t < 0$, the circuit shown in Figure P4.39 is at DC steady-state. The switch is closed at $t = 0$. Determine the voltage v_C for all time. Assume: $R_1 = R_3 = 3\ \Omega$, $R_2 = 6\ \Omega$, $V_1 = 15$ V, and $C = 0.5$ F.

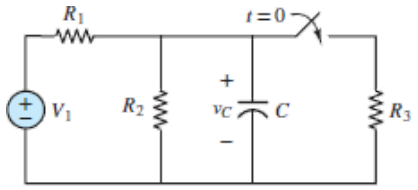


Figure P4.39

4.40 For $t < 0$, the circuit shown in Figure P4.21 is at DC steady-state. The switch is opened at $t = 0$. Determine the current i_L through the inductor for all time. Assume:

$V_S = 12$ V $\quad L = 100$ mH
$R_1 = 400\ \Omega \quad R_2 = 400\ \Omega$
$R_3 = 600\ \Omega$

4.41 For the circuit shown in Figure P4.41, assume that switch S_1 is always held open and that switch S_2 is open until being closed at $t = 0$. Assume DC steady-state conditions for $t < 0$. Also assume $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, and $C_1 = C_2 = 4$ F.

a. Find the capacitor voltage v_C at $t = 0^+$.

b. Find the time constant τ for $t > 0$.

c. Find v_C for all time and sketch the function.

d. Evaluate the ratio v_C to $v_C(\infty)$ at each of the following times: $t = 0$, τ, 2τ, 5τ, 10τ.

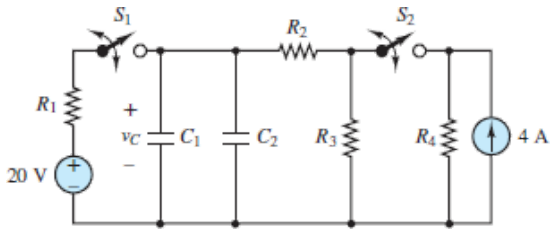


Figure P4.41

4.42 For the circuit shown in Figure P4.41, assume that switches S_1 and S_2 have been held open and closed, respectively, for a long time prior to $t = 0$. Then, simultaneously at $t = 0$, S_1 closes and S_2 opens. Also assume $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, and $C_1 = C_2 = 4$ F.

a. Find the capacitor voltage v_C at $t = 0^+$.

b. Find the time constant τ for $t > 0$.

c. Find v_C for all time and sketch the function.

d. Evaluate the ratio v_C to $v_C(\infty)$ at each of the following times: $t = 0$, τ, 2τ, 5τ, 10τ.

4.43 For the circuit shown in Figure P4.41, assume that switch S_2 is always held open and that switch S_1 is closed until being opened at $t = 0$. Subsequently, S_1 closes at $t = 3\ \tau$ and remains closed. Also assume DC steady-state conditions for $t < 0$ and $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, $C_1 = C_2 = 4$ F.

a. Find the capacitor voltage v_C at $t = 0$.

b. Find v_C for $0 < t < 3\tau$.

c. Use part b to find the capacitor voltage v_C at $t = 3\tau$, and use it to find v_C for $t > 3\tau$.

d. Compare the two time constants for $0 < t < 3\tau$ and $t > 3\tau$.

e. Sketch v_C for all time.

4.44 For the circuit shown in Figure P4.41, assume that switches S_1 and S_2 have been held open for a long time prior to $t = 0$ but then close at $t = 0$. Also assume $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, and $C_1 = C_2 = 4$ F.

a. Find the capacitor voltage v_C at $t = 0$.

b. Find the time constant τ for $t > 0$.

c. Find v_C and sketch the function.

d. Evaluate the ratio v_C to $v_C(\infty)$ at each of the following times: $t = 0$, τ, 2τ, 5τ, 10τ.

4.45 For the circuit shown in Figure P4.41, assume that switches S_1 and S_2 have been held closed for a long time prior to $t = 0$. S_1 then opens at $t = 0$; however, S_2 does not open until $t = 48$ s. Also assume $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, and $C_1 = C_2 = 4$ F.

a. Find the capacitor voltage v_C at $t = 0$.

b. Find the time constant τ for $0 < t < 48$ s.

c. Find v_C for $0 < t < 48$ s.

d. Find τ for $t > 48$ s.

e. Find v_C for $t > 48$ s.

f. Sketch v_C for all time.

Page 321

4.46 For the circuit shown in Figure P4.41, assume that switches S_1 and S_2 have been held closed for a long time prior to $t = 0$. S_2 then opens at $t = 0$; however, S_1 does not open until $t = 96$ s. Also assume $R_1 = 5\ \Omega$, $R_2 = 4\ \Omega$, $R_3 = 3\ \Omega$, $R_4 = 6\ \Omega$, and $C_1 = C_2 = 4$ F.

a. Find the capacitor voltage v_C at $t = 0$.

b. Find the time constant for $0 < t < 96$ s.

c. Find v_C for $0 < t < 96$ s.

d. Find the time constant for $t > 96$ s.

e. Use part c to find the capacitor voltage v_C at $t = 96$ s, and use it to find v_C for $t > 96$ s

f. Sketch v_C for all time.

4.47 For the circuit in Figure P4.47, determine the value of resistors R_1 and R_2, knowing that the time constant before the switch opens is 1.5 ms, and it is 10 ms after the switch opens. Assume: $R_S = 15$ kΩ, $R_3 = 30$ kΩ, and $C = 1\ \mu$F.

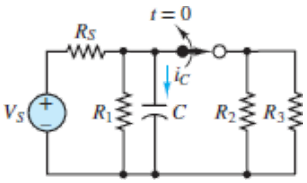


Figure P4.47

4.48 For the circuit in Figure P4.47, assume $V_S = 100$ V, $R_S = 4$ kΩ, $R_1 = 2$ kΩ, $R_2 = R_3 = 6$ kΩ, $C = 1\ \mu$F, and the circuit is in a steady-state condition before the switch opens. Find the value of v_C at $t = 8/3$ ms after the switch opens.

4.49 In the circuit in Figure P4.49, how long after the switch is thrown at $t = 0$ will $i_L = 5$ A? Assume a DC steady-state for $t < 0$. Plot $i_L(t)$.

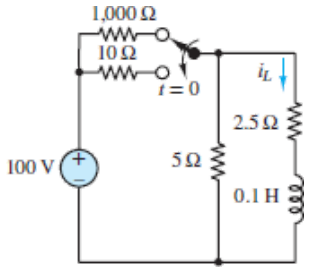


Figure P4.49

4.50 Refer to Figure P4.49 and assume that the switch takes 5 ms to move from one contact to the other. Also assume that during this time neither switch position has electrical contact. Find:

a. $i_L(t)$ for $0 < t < 5$ ms.

b. The maximum voltage between the contacts during the 5-ms duration of the switching.

4.51 The circuit in Figure P4.51 includes a voltage-controlled switch. The switch closes or opens when the voltage across the capacitor reaches the value v_M^c or v_M^o, respectively. If $v_M^o = 1$ and the period of the capacitor voltage waveform is 200 ms, find v_M^c.

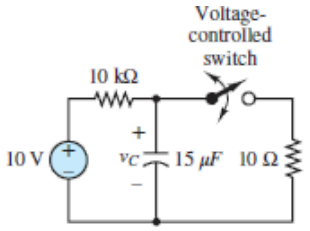


Figure P4.51

4.52 At $t = 0$ the switch in the circuit in Figure P4.52 closes. Assume that $L_1 = 1$ H, $L_2 = 5$ H, and that the circuit is in DC steady-state for $t < 0$. Find $i_L(t)$ for all time.

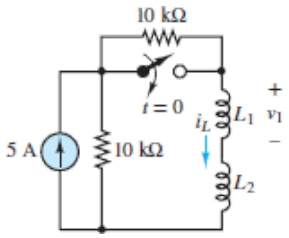


Figure P4.52

4.53 Repeat Problem P4.52 to find $v_1(t)$ for all time.

4.54 The analogy between electrical and thermal systems can be used to analyze the behavior of a pot heating on an electric stove. The heating element is modeled as shown in Figure P4.54. Find the "heat capacity" of the burner, C_S, if the burner reaches 90 percent of the desired temperature in 10 s. Assume R_S = 1.5 Ω.

Page 322

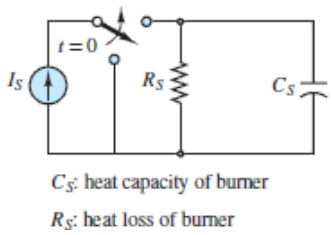


Figure P4.54

4.55 The burner and pot of Problem 4.54 can be modeled as shown in Figure P4.55. R_0 models the thermal loss between the burner and the pot. The pot is modeled by a thermal capacitance C_P in parallel with a thermal resistance R_P.

a. Find the final temperature of the water in the pot— that is, find v_o as $t \to \infty$ if I_S = 75 A, C_P = 80 F, R_0 = 0.8 Ω, R_P = 2.5 Ω, and the burner is the same as in Problem 4.54.

b. How long will it take for the water to reach 80 percent of its final temperature?

Hint: Assume $C_S \ll C_P$ such that C_S effectively acts an open-circuit.

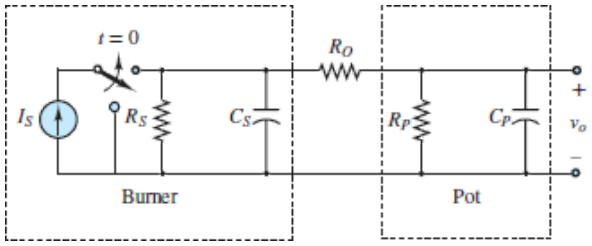


Figure P4.55

4.56 The circuit in Figure P4.56 is used as a variable delay in a burglar alarm. The alarm is a siren with an internal resistance of 1 kΩ. The alarm will not sound until the current i_0 exceeds 100 μA. Use a graphical solution or a computer simulation to find the range of the variable resistor R for which the delay is between 1 and 2 s. Assume the capacitor is initially uncharged.

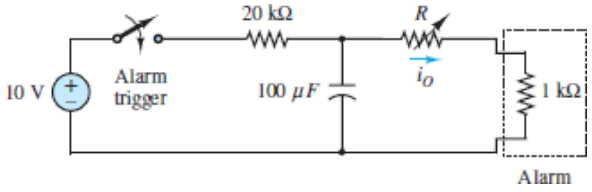


Figure P4.56

4.57 For $t > 0$, find the voltage v_1 across C_1 shown in Figure P4.57. Let $C_1 = 5$ μF and $C_2 = 10$ μF. Assume the capacitors are initially uncharged.

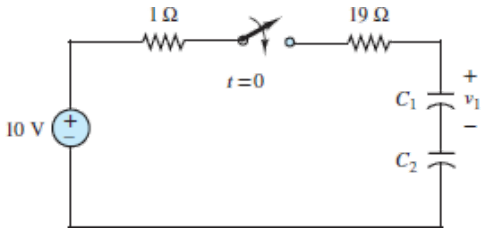


Figure P4.57

4.58 For the circuit shown in Figure P4.58 determine the time constants when the switch is open and when it is closed.

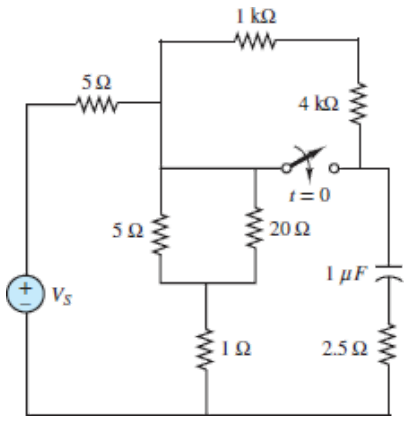


Figure P4.58

4.59 The circuit in Figure P4.59 models the charging circuit of an electronic camera flash. The flash should be charged to $v_C \leq 7.425$ V for each use. Assume $C = 1.5$ mF, $R_1 = 1$ kΩ, and $R_2 = 1$ Ω.

a. How long does it take the flash to recharge after taking a picture?

b. The shutter button stays closed for 1/30 s. How much energy is delivered to the flash bulb R_2 in that interval? Assume the capacitor is fully charged.

c. If the shutter button is pressed 3 s after a flash, how much energy is delivered to the bulb R_2?

Page 323

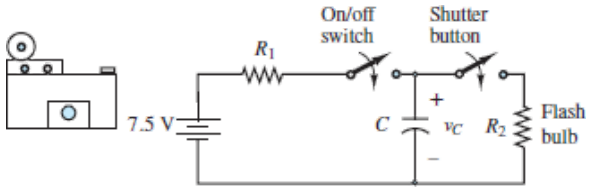


Figure P4.59

4.60 The ideal current source $i_s(t)$ in Figure P4.60 switches levels as shown. Determine and sketch the voltage $v_o(t)$ across the inductor for $0 < t < 2$ s. Assume the inductor current is zero before $t = 0$, $R_S = 500$ Ω, and $L = 50$ H.

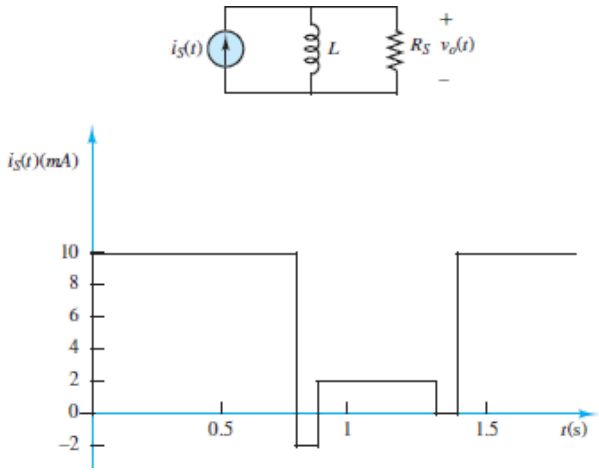


Figure P4.60

Section 4.4: Second-Order Transient Analysis

4.61 In the circuit shown in Figure P4.61:

$V_{S1} = 15$ V	$V_{S2} = 9$ V
$R_{S1} = 130\ \Omega$	$R_{S2} = 290\ \Omega$
$R_1 = 1.1$ kΩ	$R_2 = 700\ \Omega$
$L = 17$ mH	$C = 0.35\ \mu$F

Determine the voltage v_C across the capacitor and the current i_L through the inductor as $t \to \infty$.

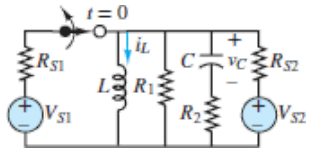


Figure P4.61

4.62 For $t > 0$, determine the current i_L through the inductor and the voltage v_C across the capacitor in Figure P4.62. Assume $v_S = -1$ V for $t < 0$ but is reversed to $v_S = 1$ V for $t > 0$. Also assume $R = 10\ \Omega$, $L = 5$ mH, $C = 100\ \mu$F, and that the circuit was in DC steady-state prior to when the source was reversed.

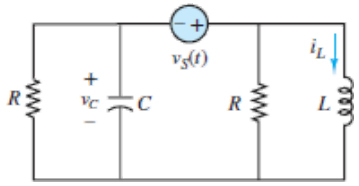


Figure P4.62

4.63 The switch shown in Figure P4.63 closes at $t = 0$. Assume a DC steady-state for $t < 0$ and:

$V_S = 170$ V $\quad R_S = 7$ kΩ
$R_1 = 2.3$ kΩ $\quad R_2 = 7$ kΩ
$L = 30$ mH $\quad C = 130$ μF

Determine the current i_L through the inductor and the voltage v_C across the capacitor for $t > 0$.

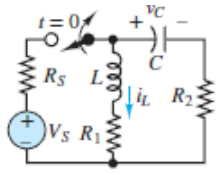


Figure P4.63

4.64 The switch in the circuit shown in Figure P4.64 closes at $t = 0$. Assume a DC steady-state for $t < 0$ and:

$V_S = 12$ V $\quad C = 130$ μF
$R_1 = 2.3$ kΩ $\quad R_2 = 7$ kΩ
$L = 30$ mH

Determine the current i_L through the inductor and the voltage v_C across the capacitor for $t > 0$.

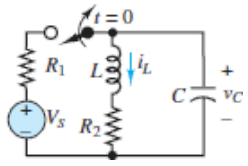


Figure P4.64

Page 324

4.65 The switch shown in Figure P4.65 is thrown at $t = 0$. Assume a DC steady-state for $t < 0$ and:

$V_S = 12$ V $\quad R_S = 100$ Ω
$R_1 = 31$ kΩ $\quad R_2 = 22$ kΩ
$L = 0.9$ mH $\quad C = 0.5$ μF

Determine the current i_1 through R_1 and the voltage v_2 across R_2 for $t > 0$.

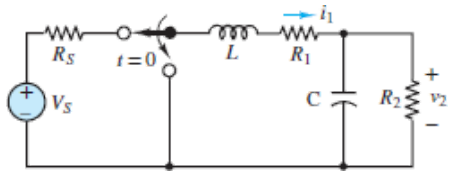


Figure P4.65

4.66 For $t < 0$, the circuit shown in Figure P4.66 is at DC steady-state and the voltage across the capacitor is +7 V. The switch is thrown at $t = 0$. Assume:

$V_S = 12$ V $\quad C = 3,300\ \mu$F
$R_1 = 9.1$ kΩ $\quad R_2 = 4.3$ kΩ
$R_3 = 4.3$ kΩ $\quad L = 16$ mH

Determine the current i_L through the inductor, the voltage v_C across the capacitor, and the current i_2 through R_2 for $t > 0$.

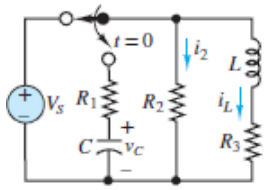


Figure P4.66

4.67 For $t < 0$, the circuit shown in Figure P4.67 is in DC steady-state. Determine the current i_L through the inductor and the voltage v_C across the capacitor for $t > 0$.

$V_{S1} = 15$ V $\quad V_{S2} = 9$ V
$R_{S1} = 130\ \Omega$ $\quad R_{S2} = 290\ \Omega$
$R_1 = 1.1$ kΩ $\quad R_2 = 700\ \Omega$
$L = 17$ mH $\quad C = 0.35\ \mu$F

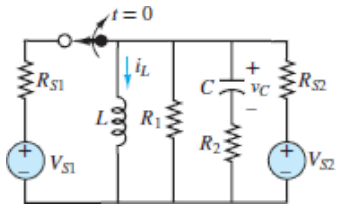


Figure P4.67

4.68 For $t < 0$, the circuit shown in Figure P4.68 is in DC steady-state. The switch is closed at $t = 0$. Determine the current i_L through the inductor and the voltage v_C across the capacitor for $t > 0$. Assume $R = 3$ kΩ, $R_S = 600$ Ω, $V_S = 2$ V, $C = 2$ mF, and $L = 1$ mH.

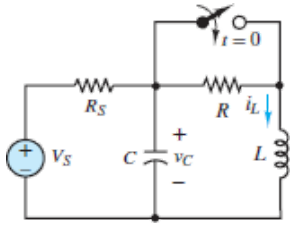


Figure P4.68

4.69 Assume the switch in the circuit in Figure P4.69 has been closed for a very long time. It is suddenly opened at $t = 0$ and then reclosed at $t = 5$ s. Determine the current i_L through the inductor, the voltage v_C across the capacitor, and the voltage v across the 2-Ω resistor for $t > 0$.

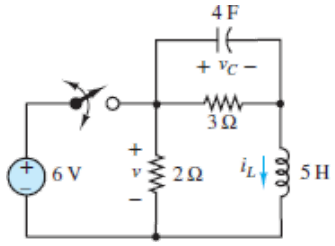


Figure P4.69

4.70 Determine whether the circuit in Figure P4.70 is overdamped or underdamped for $t > 0$. Assume $V_S = 15$ V, $R = 200$ Ω, $L = 20$ mH, and $C = 0.1$ μF. Determine the capacitance that results in critical damping.

Page 325

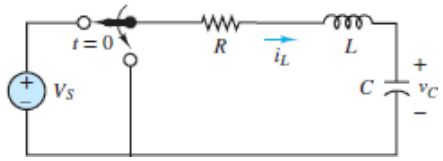


Figure P4.70

4.71 For $t < 0$, assume the circuit in Figure P4.70 is in DC steady-state. Assume V_S = 15 V, R = 200 Ω, L = 20 mH and C = 0.1 μF. If the switch is thrown at t = 0, find the:

a. Initial capacitor voltage v_C at $t = 0^+$.

b. Capacitor voltage v_C at t = 20 μs.

c. Capacitor voltage v_C as $t \rightarrow \infty$.

d. Maximum capacitor voltage.

4.72 Assume the switch in the circuit in Figure P4.69 has been open for a very long time. It is suddenly closed at t = 0 and then reopened at t = 5 s. Determine the current i_L through the inductor, the voltage v_C across the capacitor, and the voltage v across the 2-Ω resistor for $t > 0$.

4.73 Assume that the circuit shown in Figure P4.70 is underdamped, and for $t < 0$, the circuit is in DC steady-state with $v_C = V_S$. After the switch is thrown at t = 0, the first two zero crossings of the capacitor voltage v_C occur at $t = 5\pi/3$ μs and $t = 5\pi$ μs. At $t = 20\pi/3$ μs, the capacitor voltage v_C peaks at 0.6 V_S. If C = 1.6 μF, what are the values of R and L?

4.74 Given the information provided in Problem 4.73, what are the values of R and L so that the peak at $20\pi/3$ μs is $v_C = 0.7\ V_S$? Assume C = 1.6 μF.

4.75 Determine i_L for $t > 0$ in Figure P4.75, assuming $i_L(0)$ = 2.5 A and $v_C(0)$ = 10 V.

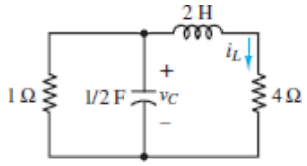


Figure P4.75

4.76 Find the maximum value of v_C for $t > 0$ in Figure P4.76, assuming DC steady-state for $t < 0$.

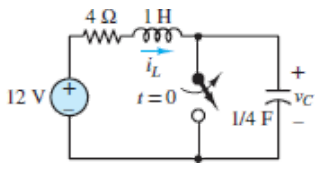


Figure P4.76

4.77 For $t > 0$, determine the time t at which $i = 2.5$ A in Figure P4.77, assuming DC steady-state for $t < 0$.

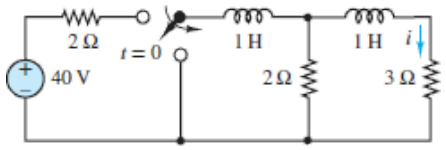


Figure P4.77

4.78 For $t > 0$, determine the time t at which $i = 6$ A in Figure P4.78, assuming DC steady-state for $t < 0$.

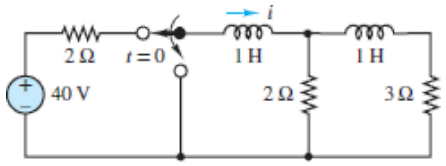


Figure P4.78

4.79 For $t > 0$, determine the time t at which $v = 7.5$ V in Figure P4.79, assuming DC steady-state for $t < 0$.

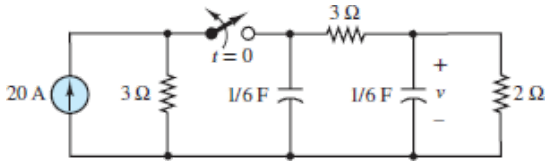


Figure P4.79

Page 326

4.80 Assume the circuit in Figure P4.80 is in DC steady-state for $t < 0$ and $L = 3$ H. Find the maximum value of v_C for $t > 0$.

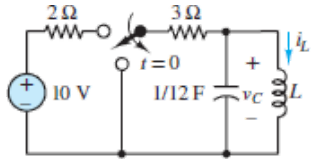


Figure P4.80

4.81 Assume the circuit in Figure P4.80 is in DC steady-state for $t < 0$. Find the value of the inductance L that makes the circuit critically damped for $t > 0$. Find the maximum value of v_C for $t > 0$.

4.82 For $t > 0$, determine v in Figure P4.82, assuming DC steady-state for $t < 0$.

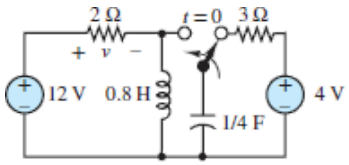


Figure P4.82

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]The secondary current will decrease by a factor of 100 so that power is conserved.

C H A P T E R

5

FREQUENCY RESPONSE AND SYSTEM CONCEPTS

Frequency-dependent phenomena are commonly encountered in engineering problems. For example, structures vibrate at a characteristic frequency when excited by wind forces (some high-rise buildings experience perceptible oscillation!). The propeller on a ship excites the shaft at a vibration frequency related to the engine's speed of rotation and to the number of blades on the propeller. An internal combustion engine is excited periodically by the combustion events in the individual cylinder, at a frequency determined by the firing of the cylinders. Wind blowing across a pipe excites a resonant vibration that is perceived as sound (wind instruments operate on this principle). Filters of all types depend upon frequency. In this respect, electric circuits are no different from other dynamic systems. A large body of knowledge has been developed related to the frequency response of electric circuits, most of it based on the ideas of phasors and impedance. The ideas developed in this chapter are applied, by analogy, to the analysis of other physical systems to illustrate the generality of the concepts.

In this chapter, quantities often involve angles. Unless indicated otherwise, angles are given in units of radians.

Learning Objectives

Students will learn to...

1. Understand the physical significance of frequency domain analysis, and compute the frequency response of circuits using AC circuit analysis tools. *Section 5.1*.
2. Compute the Fourier spectrum of periodic signals by using the Fourier series representation, and use this representation in connection with frequency response ideas to compute the response of circuits to periodic inputs. *Section 5.2*
3. Analyze simple first- and second-order electrical filters, and determine their frequency response and filtering properties. *Sections 5.3*–*5.4*.
4. Compute the frequency response of a circuit and its graphical representation in the form of a Bode plot. *Section 5.5*.

5.1 SINUSOIDAL FREQUENCY RESPONSE

The **sinusoidal frequency response** (or, simply, **frequency response**) of a circuit provides a measure of how the circuit responds to sinusoidal inputs of arbitrary frequency. In other words, for a given input signal with a particular amplitude, phase, and frequency, the frequency response of a circuit permits the computation of a particular output signal. For example, suppose you wanted to determine how the load voltage $\mathbf{V}_o$ or current $\mathbf{I}_o$ varied in response to different frequencies in the circuit of Figure 5.1. An analogy could be made, for example, with how an earbud (the load) responds to the audio signal generated by a smartphone (the source) when an amplifier (the circuit) is placed between the two.[1] In the circuit of Figure 5.1, the signal source circuitry is represented by a Thévenin source. The impedances are, in general, functions of frequency. The amplifier circuit is represented by the idealized connection of two impedances $\mathbf{Z}_1$ and $\mathbf{Z}_2$, and the load is represented by an additional impedance $\mathbf{Z}_o$. The following statement provides a general definition of the frequency response of such a system:

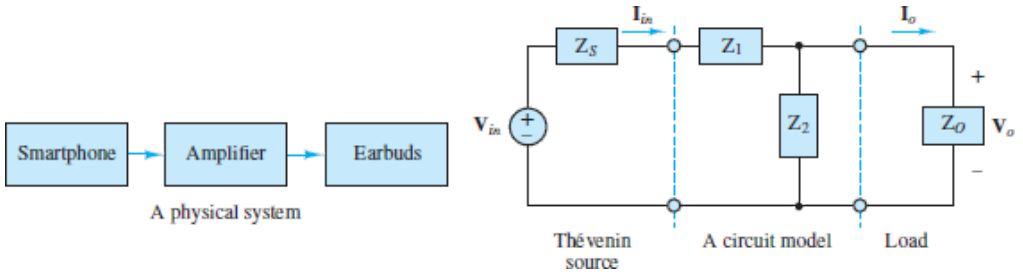


Figure 5.1 A circuit model

The frequency response of a circuit is a measure of the variation of a load-related voltage or current as a function of the frequency of the excitation signal.

Frequency Response Functions

A frequency response function is the ratio of a chosen *output* to a chosen *input.* In circuit analysis, the chosen input is often an independent voltage or current source. The chosen output can be any voltage or current elsewhere in the circuit. Page 329By convention, frequency response functions are represented by either **G** or **H**, where **G** is a dimensionless *gain* and **H** can represent a gain, an impedance, or a conductance. Four distinct versions of frequency response function follow:

$$\mathbf{G}_V(j\omega) \equiv \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_{\text{in}}(j\omega)} \qquad \mathbf{G}_I(j\omega) \equiv \frac{\mathbf{I}_o(j\omega)}{\mathbf{I}_{\text{in}}(j\omega)}$$
$$\mathbf{H}_Z(j\omega) \equiv \frac{\mathbf{V}_o(j\omega)}{\mathbf{I}_{\text{in}}(j\omega)} \qquad \mathbf{H}_Y(j\omega) \equiv \frac{\mathbf{I}_o(j\omega)}{\mathbf{V}_{\text{in}}(j\omega)} \tag{5.1}$$

In many cases the inputs $\mathbf{V}_{\text{in}}$ and $\mathbf{I}_{\text{in}}$ are chosen to be independent voltage and current sources, respectively. The outputs $\mathbf{V}_o$ and $\mathbf{I}_o$ are freely chosen and, as such, represent the load in a circuit.

The above frequency response functions are related by the impedance $\mathbf{Z}_o$ of the load. For example, if $\mathbf{G}_V(j\omega)$ and $\mathbf{G}_I(j\omega)$ are known, the other two can be derived directly:

$$\mathbf{H}_Z(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{I}_{\text{in}}(j\omega)} = \frac{\mathbf{I}_o(j\omega)}{\mathbf{I}_{\text{in}}(j\omega)}\mathbf{Z}_o(j\omega) = \mathbf{G}_I(j\omega)\mathbf{Z} \tag{5.2}$$

$$\mathbf{H}_Y(j\omega) = \frac{\mathbf{I}_o(j\omega)}{\mathbf{V}_{\text{in}}(j\omega)} = \frac{\mathbf{V}_o(j\omega)}{\mathbf{Z}_o(j\omega)}\frac{1}{\mathbf{V}_{\text{in}}(j\omega)} = \frac{\mathbf{G}_V(j\omega)}{\mathbf{Z}_o(j\omega)} \tag{5.3}$$

Circuit Simplification

In general, the first step in determining the details of a chosen frequency response function is to divide the circuit into a load (in accord with the chosen output) and a source. Consider again the circuit shown in Figure 5.1. The network attached to the load can be replaced by its Thévenin equivalent as shown in Figure 5.2. Once Page 330the load is reattached as in Figure 5.3, voltage division can be applied to express $\mathbf{V}_o$ in terms of $\mathbf{V}_T$, and then eventually in terms of $\mathbf{V}_{\text{in}}$.

$$\begin{aligned}\mathbf{V}_o &= \frac{\mathbf{Z}_o}{\mathbf{Z}_o + \mathbf{Z}_T}\mathbf{V}_T \\ &= \frac{\mathbf{Z}_o}{\mathbf{Z}_o + (\mathbf{Z}_{\text{in}} + \mathbf{Z}_1)\mathbf{Z}_2/(\mathbf{Z}_{\text{in}} + \mathbf{Z}_1 + \mathbf{Z}_2)} \cdot \frac{\mathbf{Z}_2}{\mathbf{Z}_{\text{in}} + \mathbf{Z}_1 + \mathbf{Z}_2}\mathbf{V}_{\text{in}} \\ &= \frac{\mathbf{Z}_o\mathbf{Z}_2}{\mathbf{Z}_o(\mathbf{Z}_{\text{in}} + \mathbf{Z}_1 + \mathbf{Z}_2) + (\mathbf{Z}_{\text{in}} + \mathbf{Z}_1)\mathbf{Z}_2}\mathbf{V}_{\text{in}}\end{aligned} \tag{5.4}$$

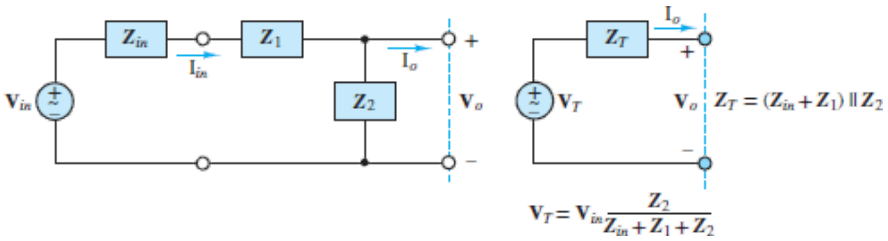


Figure 5.2 Thévenin equivalent source network

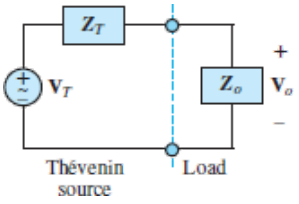


Figure 5.3 Equivalent circuit from the perspective of the load

The gain, $\mathbf{G}_V(j\omega)$, is a dimensionless complex quantity, given by:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_{\text{in}}}(j\omega) = \frac{\mathbf{Z}_o\mathbf{Z}_2}{\mathbf{Z}_o(\mathbf{Z}_{\text{in}} + \mathbf{Z}_1 + \mathbf{Z}_2) + (\mathbf{Z}_{\text{in}} + \mathbf{Z}_1)\mathbf{Z}_2} \tag{5.5}$$

Thus, the gain is known if the circuit element impedances are known.

$\mathbf{V}_o(j\omega)$ is a phase-shifted and amplitude-scaled version of $\mathbf{V}_{\text{in}}(j\omega)$.

If the phasor source voltage and the frequency response of the circuit are known, the phasor load voltage can be computed as follows:

$$\mathbf{V}_o(j\omega) = \mathbf{G}_V(j\omega) \cdot \mathbf{V}_{\text{in}}(j\omega) \tag{5.6}$$

$$V_o e^{j\phi_o} = |\mathbf{G}_V| e^{j\angle \mathbf{G}_v} \cdot V_{\text{in}} e^{j\phi_{\text{in}}} \tag{5.7}$$

such that

$$V_o = |\mathbf{G}_V| \cdot V_{\text{in}} \tag{5.8}$$

and

$$\phi_o = \angle \mathbf{G}_v + \phi_{\text{in}} \tag{5.9}$$

At any given angular frequency ω, the load voltage is a sinusoid with the same frequency as the source voltage.

First- and Second-Order Archetypes

Whenever possible, the first step toward deriving a frequency response function is to use Thévenin's or Norton's theorem to simplify the circuit. If the circuit is first order, or second order with the storage elements in series or parallel, it can be simplified to one of the four archetypes shown in Figures 5.4 to 5.7.

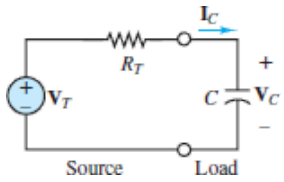


Figure 5.4 Simplified first-order circuit with one capacitor

In the first-order circuit of Figure 5.4, the loop current $\mathbf{I}_C$ is related to the Thévenin source voltage $\mathbf{V}_T$ by the generalized Ohm's law:

$$\mathbf{I}_C = \frac{\mathbf{V}_T}{R_T + \mathbf{Z}_C} \tag{5.10}$$

Multiply the numerator and denominator by$(j\omega)C$ and divide both sides by $\mathbf{V}_T$ to find the frequency response function:

$$\mathbf{H}_Y(j\omega) = \frac{\mathbf{I}_C}{\mathbf{V}_T} = \frac{(j\omega)C}{1 + (j\omega)\tau_C} \tag{5.11}$$

where $\tau_C = R_T C$.

Page 331

It is now a simple matter to find the frequency response function relating $\mathbf{V}_C$ to $\mathbf{V}_T$:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_C}{\mathbf{V}_T} = \frac{\mathbf{I}_C}{\mathbf{V}_T}\mathbf{Z}_C = \frac{(j\omega)C}{1 + (j\omega)\tau_C}\frac{1}{(j\omega)C} = \frac{1}{1 + (j\omega)\tau_C} \tag{5.12}$$

Notice that the denominator is the same as in $\mathbf{H}_Y$. This is a common result because the denominator expresses the characteristic dynamics of the circuit. The numerator expresses differences in the circuit variables. It is a useful exercise to derive $\mathbf{G}_V$ directly from voltage division. Try it!

A similar approach can be taken to find the frequency response relating the voltage $\mathbf{V}_L$ to the Norton source current $\mathbf{I}_N$ in the first-order circuit of Figure 5.5. Apply the generalized Ohm's law to write:

$$\mathbf{V}_L = \mathbf{I}_N(R_N \| \mathbf{Z}_L) = \mathbf{I}_N\frac{R_N(j\omega)L}{R_N + (j\omega)L} \tag{5.13}$$

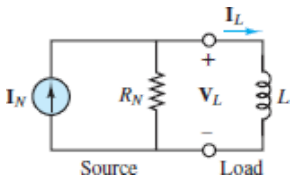


Figure 5.5 Simplified first-order circuit with one inductor

The frequency response function is found by dividing both sides by $\mathbf{I}_N$ and then dividing the numerator and denominator by R_N.

$$\mathbf{H}_Z(j\omega) = \frac{\mathbf{V}_L}{\mathbf{I}_N} = \frac{(j\omega)L}{1 + (j\omega)\tau_L} \tag{5.14}$$

where, in this case, $\tau_L = L/R_N$.

Again, it is a simple matter to find the frequency response function relating $\mathbf{I}_L$ to $\mathbf{I}_N$:

$$\mathbf{G}_I(j\omega) = \frac{\mathbf{I}_L}{\mathbf{I}_N} = \frac{\mathbf{V}_L}{\mathbf{I}_N}\frac{1}{\mathbf{Z}_L} = \frac{1}{1 + (j\omega)\tau_L} \tag{5.15}$$

Notice that the denominator is the same as in $\mathbf{H}_Z$. It is a useful exercise to derive $\mathbf{G}_I$ directly from current division. Try it!

Second-order circuits are handled in much the same way. Consider the series *LC* circuit of Figure 5.6. The common loop current $\mathbf{I}_L$ is related to the Thévenin source voltage $\mathbf{V}_T$ by the generalized Ohm's law:

$$\mathbf{V}_T = \mathbf{I}_L(R_T + \mathbf{Z}_C + \mathbf{Z}_L) \tag{5.16}$$

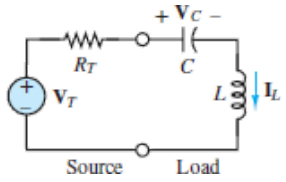


Figure 5.6 Simplified second-order circuit with one capacitor and one inductor in series

Divide both sides by $\mathbf{I}_L$, invert both sides, and multiply the resulting numerator and denominator by $j\omega C$ to find:

$$\mathbf{H}_Y(j\omega) = \frac{\mathbf{I}_L}{\mathbf{V}_T} = \frac{(j\omega)C}{1 + (j\omega)R_T C + (j\omega)^2 LC} = \frac{(j\omega)C}{1 + (j\omega)\tau_C + (j\omega/\omega_n)^2} \tag{5.17}$$

where $\omega_n^2 = 1/LC$ is the natural frequency always found in second-order series and parallel *LC* circuits.

The voltage gain $\mathbf{G}_V$ for the second-order series *LC* circuit can be found using the result for $\mathbf{H}_Y$.

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_C}{\mathbf{V}_T} = \frac{\mathbf{I}_C}{\mathbf{V}_T}\mathbf{Z}_C \tag{5.18}$$

Of course, $\mathbf{I}_C = \mathbf{I}_L$ for the series loop, so:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{I}_L}{\mathbf{V}_T}\mathbf{Z}_C = \mathbf{H}_Y\mathbf{Z}_C = \frac{1}{1 + (j\omega)\tau_C + (j\omega/\omega_n)^2} \tag{5.19}$$

Finally, Figure 5.7 shows a second-order parallel *LC* circuit. The common voltage $\mathbf{V}_C$ is related to the Norton source current $\mathbf{I}_N$ by the generalized Ohm's law:

$$\mathbf{V}_C = \mathbf{I}_N(R_N \| \mathbf{Z}_L \| \mathbf{Z}_C) \tag{5.20}$$

Divide both sides by $\mathbf{I}_N$ to obtain:

$$\mathbf{H}_Z(j\omega) = \frac{\mathbf{V}_C}{\mathbf{I}_N} = \frac{1}{1/R_N + 1/j\omega L + j\omega C} \tag{5.21}$$

Multiply the numerator and denominator by $j\omega L$ to obtain:

$$\mathbf{H}_Z(j\omega) = \frac{(j\omega)L}{1 + (j\omega)L/R_N + (j\omega)^2 LC} = \frac{(j\omega)L}{1 + (j\omega)\tau_L + (j\omega/\omega_n)^2} \tag{5.22}$$

where $\omega_n^2 = 1/LC$ is the natural frequency always found in second-order series and parallel LC circuits.

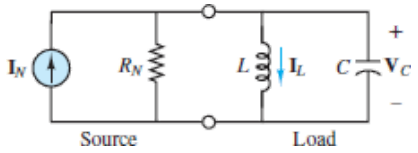


Figure 5.7 Simplified second-order circuit with one capacitor and one inductor in parallel

The current gain $\mathbf{G}_I$ for the second-order parallel LC circuit can be found using the result for $\mathbf{H}_Z$.

$$\mathbf{G}_I(j\omega) = \frac{\mathbf{I}_L}{\mathbf{I}_N} = \frac{\mathbf{V}_L}{\mathbf{I}_N}\frac{1}{\mathbf{Z}_L} \tag{5.23}$$

Of course, $\mathbf{V}_L = \mathbf{V}_C$, so:

$$\mathbf{G}_I(j\omega) = \frac{\mathbf{V}_C}{\mathbf{I}_N}\frac{1}{\mathbf{Z}_L} = \frac{\mathbf{H}_Z}{\mathbf{Z}_L} = \frac{1}{1 + (j\omega)\tau_L + (j\omega/\omega_n)^2} \tag{5.24}$$

Poles and Zeros

By definition, a frequency response function is the ratio of an output to an input. Consequently, the development of any specific frequency response function will, in general, also result in a ratio. The numerator and denominator can always be expressed as the product of four distinct types of **standard terms**. One of these terms is simply a constant. The other three terms are known as *zeros* or *poles,* depending upon whether they appear in the numerator or denominator, respectively.

1. K A constant

2. ($j\omega$) Pole or zero at the origin
3. ($1 + j\omega\tau$) Simple pole or zero
4. $[1 + j\omega\tau + (j\omega/\omega_n)^2]$ Quadratic (complex) pole or zero

A simple pole or zero may also take the form $(1 + j\omega/\omega_0)$, where $\omega_0 = 1/\tau$. In this context, τ and ω_n are a generic time constant and a generic natural frequency, respectively, each of which may differ from one zero/pole to another. A frequency response function may have multiples of each type of term listed above.

Page 333

The first- and second-order frequency response functions developed in the previous section are good examples of the standard form in which the numerator and denominator are expressed as products of these four terms. These same terms will appear repeatedly when filters and Bode plots are discussed.

EXAMPLE 5.1 Computing the Frequency Response Using Thévenin's Theorem

Problem

Compute the frequency response $\mathbf{G}_V(j\omega) = \mathbf{V}_o/\mathbf{V}_S$ for Figure 5.8.

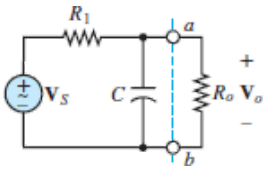


Figure 5.8

Solution

Known Quantities: $R_1 = 1$ kΩ; $C = 10$ μF; $R_o = 10$ kΩ.

Find: The frequency response $\mathbf{G}_V(j\omega) = \mathbf{V}_o/\mathbf{V}_S$.

Assumptions: None.

Analysis: With R_o as the load resistance, the approach is to use Thévenin's theorem to determine the equivalent network of the source network; that is, the equivalent network of everything to the left of terminals *a* and *b*. The Thévenin equivalent impedance $\mathbf{Z}_T$ of the source network is

$$\mathbf{Z}_T = (R_1 \| \mathbf{Z}_C)$$

The Thévenin (open-circuit) voltage $\mathbf{V}_T$ across terminals *a* and *b* is found from voltage division:

$$\mathbf{V}_T = \mathbf{V}_S \frac{\mathbf{Z}_C}{R_1 + \mathbf{Z}_C}$$

After reattaching the load to the terminals shown in Figure 5.9, the voltage $\mathbf{V}_o$ across the load can be found by applying voltage division once more:

$$\mathbf{V}_o = \frac{R_o}{\mathbf{Z}_T + R_o}\mathbf{V}_T$$

$$= \frac{R_o}{R_1\mathbf{Z}_C/(R_1 + \mathbf{Z}_C) + R_o}\frac{\mathbf{Z}_C}{R_1 + \mathbf{Z}_C}\mathbf{V}_S$$

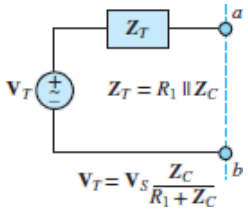


Figure 5.9

Thus:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_S}(j\omega) = \frac{R_o\mathbf{Z}_C}{R_o(R_1 + \mathbf{Z}_C) + R_1\mathbf{Z}_C} = \frac{R_o}{R_o + R_1}\frac{1}{1 + j\omega R_T C}$$

where $R_o/(R_o + R_1)$ is the DC gain when $\omega \rightarrow 0$ and the capacitor acts as an open-circuit. $R_T = R_1 \| R_o$ is the Thévenin equivalent resistance seen by the capacitor with the load attached. The impedances of the circuit elements are $R_1 = 10^3\ \Omega$, $\mathbf{Z}_C = 1/(j\omega \times 10^{-5})\ \Omega$, and $R_o = 10^4\ \Omega$. The resulting frequency response is

$$\mathbf{G}_V(j\omega) = \frac{\dfrac{10^4}{j\omega \times 10^{-5}}}{10^4\left(10^3 + \dfrac{1}{j\omega \times 10^{-5}}\right) + \dfrac{10^3}{j\omega \times 10^{-5}}} = \frac{100}{110 + j\omega}$$

$$= \frac{100}{\sqrt{110^2 + \omega^2}\, e^{j\tan^{-1}(\omega/110)}} = \frac{100}{\sqrt{110^2 + \omega^2}} \angle -\tan^{-1}\left(\frac{\omega}{110}\right)$$

Page 334

Comments: The use of equivalent circuit ideas is often helpful in deriving frequency response functions. However, it is certainly not the only method of solution. For example, node analysis would have yielded the same results just as easily, by recognizing that the top node voltage is equal to the load voltage and by solving directly for $\mathbf{V}_o$ as a function of $\mathbf{V}_S$, without going through the intermediate step of computing the Thévenin equivalent source circuit.

EXAMPLE 5.2 Computing the Frequency Response

Problem

Compute the frequency response $\mathbf{H}_Z(j\omega) = \mathbf{V}_o/\mathbf{I}_S$ for Figure 5.10.

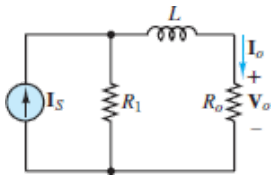


Figure 5.10

Solution

Known Quantities: $R_1 = 1$ kΩ; $L = 2$ mH; $R_o = 4$ kΩ.

Find: The frequency response $\mathbf{H}_Z(j\omega) = \mathbf{V}_o/\mathbf{I}_S$.

Assumptions: None.

Analysis: While it is possible to find the Thévenin or Norton equivalent network of everything attached to R_o and proceed as in the previous example to find the frequency response function, it is also possible to apply current division to find $\mathbf{I}_o$ and then apply Ohm's law to find $\mathbf{V}_o$ and thus the frequency response function.

Apply current division to write:

$$\mathbf{I}_o = \frac{R_1}{R_1 + R_o + j\omega L}\mathbf{I}_S$$

Factor out $R_1 + R_o$ in the denominator to find:

$$\mathbf{I}_o = \frac{R_1}{R_1 + R_o}\frac{1}{1 + j\omega L/(R_1 + R_o)}\mathbf{I}_S$$

Then:

$$\frac{\mathbf{V}_o}{\mathbf{I}_S}(j\omega) = \mathbf{H}_Z(j\omega) = \frac{\mathbf{I}_o R_o}{\mathbf{I}_S}$$

$$= \frac{R_1 R_o}{R_1 + R_o}\frac{1}{1 + j\omega L/R_N}$$

where $R_1R_o/(R_1 + R_o)$ is the DC gain when $\omega \rightarrow 0$ and the inductor acts as a short-circuit. $R_N = R_1 + R_o$ is the Norton equivalent resistance seen by the inductor. Substitute numerical values to obtain:

$$\mathbf{H}_Z(j\omega) = \frac{(10^3)(4 \times 10^3)}{10^3 + 4 \times 10^3}\frac{1}{1 + (j\omega)(2 \times 10^{-3})/(5 \times 10^3)}$$

$$= (0.8 \times 10^3)\frac{1}{1 + j0.4 \times 10^{-6}\omega}$$

Comments: The units of $\mathbf{H}_Z(j\omega)$ should be ohms. Verify that they are!

Page 335

CHECK YOUR UNDERSTANDING

Refer to Example 5.1 and compute the magnitude and phase of $\mathbf{G}_V$ at the frequencies ω = 10, 100, and 1,000 rad/s.

Answer: Magnitude = 0.9054, 0.6727, and 0.0994; phase (degrees) = −5.1944, −42.2737, and −83.7227

CHECK YOUR UNDERSTANDING

Refer to Example 5.2 and compute the magnitude and phase of $\mathbf{H}_Z$ at the frequencies ω = 1, 10, and 100 Mrad/s.

Answer: Magnitude = 742.78 Ω, 194.03 Ω, and 19.99 Ω; phase (degrees) = −21.8°, −75.96°, and −88.57°

5.2 FOURIER ANALYSIS

It is possible to represent periodic signals by means of the superposition of various sinusoidal signals of different amplitude, phase, and frequency. Let the signal $x(t)$ be periodic with period T, that is,

$$x(t) = x(t+T) = x(t+nT) \qquad n = 1,2,3,\ldots;\ T = \text{period} \tag{5.25}$$

An example of a periodic signal is shown in Figure 5.11.

Figure 5.11 A periodic signal

The signal $x(t)$ can be expressed as an infinite summation of sinusoidal components, known as a **Fourier series**, using either of the following two representations.

Fourier Series

1. Sine-cosine (quadrature) representation

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(n\frac{2\pi}{T}t\right) + \sum_{n=1}^{\infty} b_n \sin\left(n\frac{2\pi}{T}t\right) \tag{5.26}$$

2. Magnitude and phase form

$$x(t) = c_0 + \sum_{n=1}^{\infty} c_n \sin\left(n\frac{2\pi}{T}t + \theta_n\right) \tag{5.27}$$

$$x(t) = c_0 + \sum_{n=1}^{\infty} c_n \cos\left(n\frac{2\pi}{T}t - \psi_n\right) \tag{5.28}$$

In each of these expressions, the period T is related to the **fundamental frequency** of the signal ω_0 by

$$\omega_0 = 2\pi f_0 = \frac{2\pi}{T} \quad \text{rad/s} \tag{5.29}$$

The frequencies $2\omega_0$, $3\omega_0$, $4\omega_0$, etc., are called **harmonics**.

Page 336

It is straightforward to show that equations 5.26 and 5.27 are equivalent when the parameters a_n, b_n, c_n, and θ_n are related by:

$$\sqrt{a_n^2 + b_n^2} = c_n \quad \text{and} \quad \frac{b_n}{a_n} = \cot(\theta_n) \tag{5.30}$$

Similarly, one can show that equations 5.26 and 5.28 are equivalent when the parameters a_n, b_n, c_n, and ψ_n are related by:

$$\sqrt{a_n^2 + b_n^2} = c_n \quad \text{and} \quad \frac{b_n}{a_n} = \tan(\psi_n) \tag{5.31}$$

Figure 5.12 is a graphical representation of the equivalence of the $\{a_n, b_n\}$ and $\{c_n, \theta_n\}$ forms of the Fourier series.

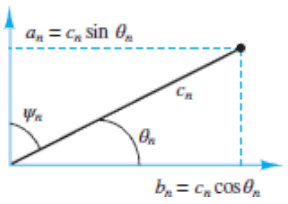


Figure 5.12 Relationship between $\{a_n, b_n\}$ and $\{c_n, \theta_n\}$ forms

Each form of the Fourier series has its distinct advantages. The sine-cosine representation uses odd and even functions of the independent variable. Odd functions are antisymmetric about the origin and satisfy the condition:

$$f(-t) = -f(t) \qquad \text{odd function} \tag{5.32}$$

Sine functions are odd. Even functions are symmetric about the origin and satisfy the condition:

$$f(-t) = f(t) \qquad \text{even function} \tag{5.33}$$

Cosine functions are even, as is any constant, such as a_0. Figure 5.13 shows even and odd function examples.

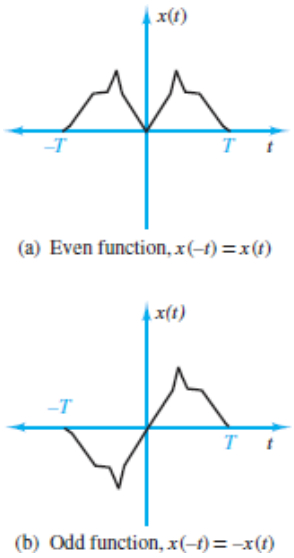


Figure 5.13 Definition of even and odd functions

The advantage of the representation in equation 5.26 is that if *x (t)* is known to be odd (even), it can be represented as the sum of only odd (even) functions [i.e., using only the sine (cosine) terms], thus resulting in easier evaluation of the Fourier series coefficients.

The magnitude and phase forms, equations 5.27 and 5.28, separate out the magnitude information c_n from the phase information θ_n or ψ_n. In this form, Fourier series may be combined readily with magnitude and phase responses of linear systems to periodic inputs. The magnitude and phase components are often represented as a discrete **frequency spectrum**, as shown in Figure 5.14.

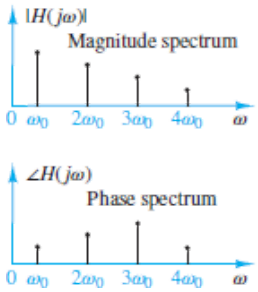


Figure 5.14 Discrete frequency spectrum

Computation of Fourier Series Coefficients

The computation of the $\{a_n, b_n\}$ or $\{c_n, \theta_n\}$ coefficients for the periodic function $x(t)$ is based on the following formulas:

$$a_0 = \frac{1}{T}\int_0^T x(t)\,dt = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\,dt = \text{average value of } x(t) \tag{5.34}$$

$$a_n = \frac{2}{T}\int_0^T x(t)\cos\left(n\frac{2\pi}{T}t\right)dt = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos\left(n\frac{2\pi}{T}t\right)dt \tag{5.35}$$

$$b_n = \frac{2}{T}\int_0^T x(t)\sin\left(n\frac{2\pi}{T}t\right)dt = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\sin\left(n\frac{2\pi}{T}t\right)dt \tag{5.36}$$

The integral limits are written in two different forms to illustrate that it does not matter where the integration starts, provided that it is carried out over one entire period. The c_n and θ_n (or ψ_n) values can be derived from the a_n and b_n coefficients by using equations 5.30 and 5.31.

To illustrate the significance of the Fourier series decomposition, consider the square wave of Figure 5.15(a), which is an even function. Only even function (cosine) terms are nonzero. The first six nonzero Fourier series terms are shown in Figure 5.15(b). Note that the first term corresponds to a_0, which is the average value $A/2$ of the function. The other five terms correspond to n odd (1, 3, 5, 7, and 9). Note also that the coefficients for $n = 1$, 5, and 9 are positive, and those for $n = 3$ and 7 are negative. This result is apparent when comparing the peaks of the cosine waveforms. This alternation of positive and negative cosines is required so that each term can add or subtract from the previous terms as needed to "flatten" the waveform into a square wave. Figure 5.15(c) compares the original square wave with this six-term Fourier series approximation. It is evident that this approximation does not reproduce the sharp edges of the square wave. As more terms are added, the resulting approximation improves.

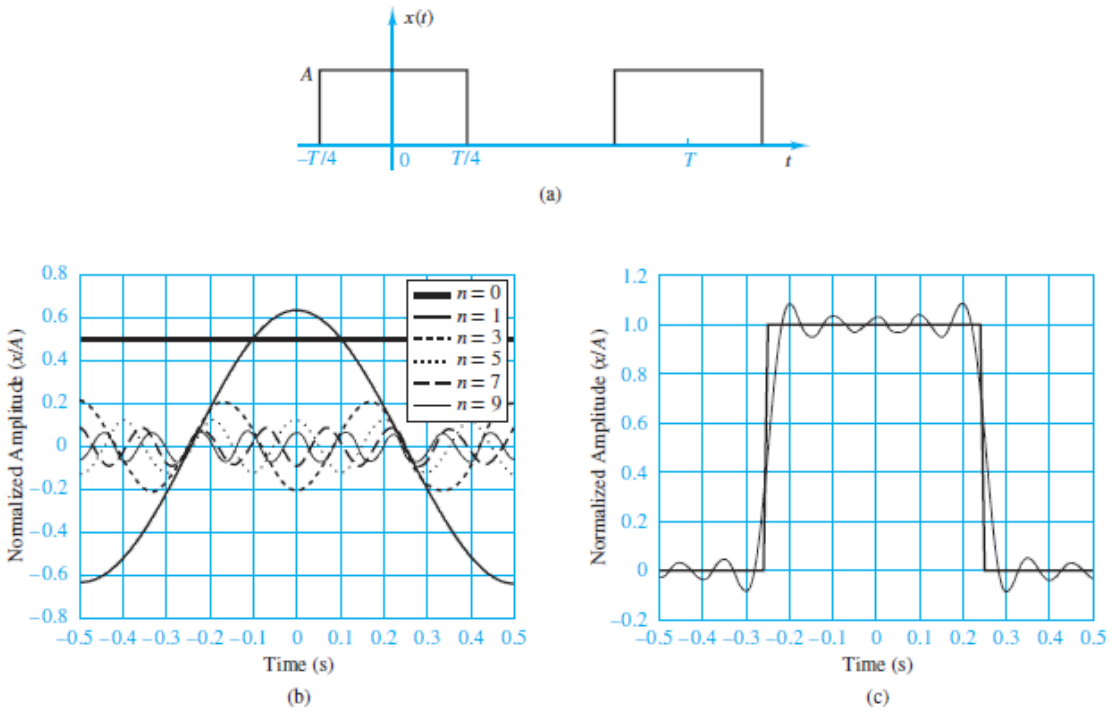


Figure 5.15 Square wave and its representation by a Fourier series. (a) Square wave (even function); (b) first six Fourier series terms of square ware; (c) sum of first six Fourier series terms superimposed upon a square wave

The complete Fourier series for the square wave shown in Figure 5.15(a) is

$$f(t) = \frac{A}{2} + \frac{2A}{\pi}\sum_{n=1}^{\infty}(-1)^{(n-1)}\frac{\cos\left[(2n-1)(2\pi)t/T\right]}{2n-1}$$

It is interesting to note that if the square wave shown in Figure 5.15(a) were shifted to the right by $T/4$ the resulting waveform would be an odd function. The complete Fourier series now contains only odd functions (sines) and is

$$f(t) = \frac{A}{2} + \frac{2A}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)(2\pi)t/T]}{2n-1}$$

The Fourier series for two other common waveforms are listed below. In each case the peak-to-trough amplitude is A, the period is T, and the average value is $A/2$.

$$f(t) = \frac{A}{2} + \frac{A}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n\pi t/T)}{n} \qquad \text{Sawtooth wave}$$

$$f(t) = \frac{A}{2} - \frac{4A}{\pi^2}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)(2\pi)t/T]}{(2n-1)^2} \qquad \text{Triangle wave}$$

These Fourier series converge everywhere that the function itself is differentiable. For example, the Fourier series for the square wave in Figure 5.15 (a) does not converge at $t = -T/4$, $T/4$, $3T/4$, . . . while the Fourier series for the sawtooth wave does not converge at $t = 0$, T, $2T$, ...

Response of Linear Systems to Periodic Inputs

The frequency response concept is particularly useful when one deals with a system excited by a periodic input, which can be modeled by a Fourier series of sinusoids of known amplitude and phase, but different frequencies. Assume that the Fourier series has a finite number of terms:

$$x(t) = c_0 + \sum_{n=1}^{N} c_n \sin\left(n\frac{2\pi}{T}t + \theta_n\right) \tag{5.37}$$

Each of the N sinusoids is characterized by amplitude c_n, phase θ_n, and frequency $\omega_n = n\omega_0$, where $\omega_0 = 2\pi/T$ and T is the period of the input signal. For example, the periodic input could be a sawtooth waveform.

Figure 5.16 illustrates the input-output representation of a system, making use of the frequency response concept. The figure shows that if the input $q_{in}(t)$ to a linear system can be represented by the phasor $Q_{in}(j\omega)$, then the output can be computed by multiplying the phasor input by the frequency response function of the linear system. This product is a complex number that can be computed by multiplying the magnitude of the input phasor by the magnitude of the frequency response function and by adding the phase angle of the input phasor to the phase angle of the frequency response.

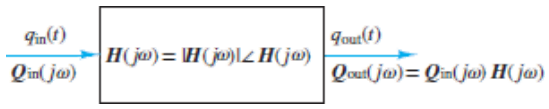


Figure 5.16 Response of a linear system to a phasor input

It is important to recognize that each of the input sinusoidal components propagates through the system according to the frequency response. Thus, the discrete magnitude spectrum of the periodic output signal in the steady state is equal Page 339to the discrete magnitude spectrum of the input signal *multiplied* by the amplitude ratio of the frequency response of the system *at each discrete frequency.* The phase spectrum of the output signal in the steady state is equal to the phase spectrum of the input signal *added* to the phase angle frequency response of the system *at each discrete frequency*. If $x\ (t)$ is the input to a linear system in the form given by equation 5.37, and if the linear system has a frequency response function $\mathbf{H}(j\omega)$, then the time-dependent output of the system $y\ (t)$ is given by

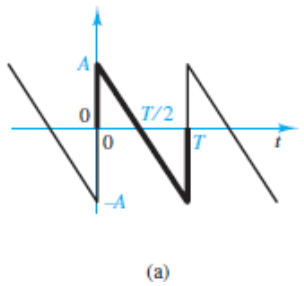
$$y(t) = \sum_{n=1}^{N} |\mathbf{H}(j\omega_n)| c_n \sin[\omega_n t + \theta_n + \angle \mathbf{H}(j\omega_n)] \tag{5.38}$$

where $|\ \mathbf{H}(j\omega_n)\ |$ and $\angle \mathbf{H}(j\omega_n)$ are the magnitude and phase, respectively, at the frequency corresponding to the nth harmonic $n\omega_0$ of the input.

EXAMPLE 5.3 Computation of Fourier Series Coefficients

Problem

Compute the complete Fourier spectrum of the *sawtooth* function shown in Figure 5.17; that is, find a general expression for the coefficients a_n and b_n as a function of n, and then compute the spectrum of $x\ (t)$, that is, the coefficients c_n and θ_n. Plot the spectrum of the signal.



(a)

Figure 5.17 (a) Periodic (sawtooth) function

Solution

Known Quantities: Amplitude and period of sawtooth waveform.

Find: Fourier series coefficients a_n and b_n.

Schematics, Diagrams, Circuits, and Given Data: The function is periodic, with period $T = 1$ s and peak amplitude $A = 1$.

Assumptions: None.

Analysis: The function in Figure 5.17(a) is odd since it is antisymmetric about the vertical axis. Thus, only the b_n coefficients are nonzero. First, express $x(t)$ over one period:

$$x(t) = A\left(1 - \frac{2t}{T}\right) \qquad 0 \le t < T$$

Then evaluate the integral in equation 5.36:

$$\begin{aligned} b_n &= \frac{2}{T}\int_0^T A\left(1 - \frac{2t}{T}\right)\sin\left(n\frac{2\pi}{T}t\right)dt \\ &= \frac{2}{T}\int_0^T A\sin\left(n\frac{2\pi}{T}t\right)dt + \frac{2A}{T}\int_0^T\left(-\frac{2t}{T}\right)\sin\left(n\frac{2\pi}{T}t\right)dt \\ &= \frac{2A}{T}\left[-\frac{T}{2n\pi}\cos\left(n\frac{2\pi}{T}t\right)\right]_0^T - \frac{4A}{T^2}\int_0^T t\cdot\sin\left(n\frac{2\pi}{T}t\right)dt \\ &= 0 - \frac{4A}{T^2}\left[\frac{1}{n^2(2\pi/T)^2}\sin\left(n\frac{2\pi}{T}t\right) - \frac{t}{n(2\pi/T)}\cos\left(n\frac{2\pi}{T}\right)\right]\Bigg|_0^T \\ &= -\frac{4A}{T^2}\left[-\frac{T^2}{2n\pi}\cos(2n\pi)\right] = \frac{2A}{n\pi} \qquad n = 1, 2, 3, \ldots \end{aligned}$$

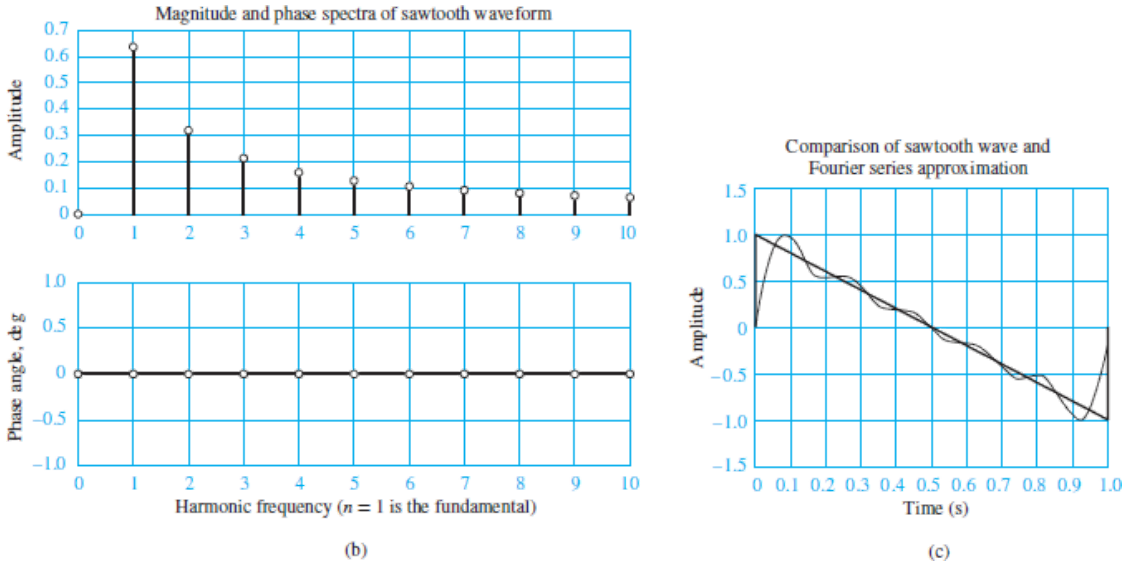


Figure 5.17 (b) Spectrum of sawtooth waveform; (c) approximation of sawtooth waveform for $N = 5$

To compute the spectrum of the signal, apply equation 5.30:

$$c_n = \sqrt{a_n^2 + b_n^2} = |b_n|$$

$$\theta_n = \cot^{-1}\frac{b_n}{a_n} = \cot^{-1}\frac{b_n}{0} = 0$$

The individual components of the spectrum of $x(t)$ are shown in Figure 5.17(b).

Comments: A computer program, such as MatLab®, can be used to visualize the result of a Fourier series approximation. Figure 5.17(c) shows one such approximation.

EXAMPLE 5.4 Computation of Fourier Series Coefficients

Problem

Compute the complete Fourier series expansion of the pulse waveform shown in Figure 5.18(a) for $\tau/T = 0.2$. Plot the spectrum of the signal.

Figure 5.18 (a) Pulse train

Solution

Known Quantities: Amplitude and period of pulse train waveform.

Find: Fourier series coefficients a_n and b_n; Fourier spectrum.

Schematics, Diagrams, Circuits, and Given Data: The function is periodic, with period $T = 1$ s, peak amplitude $A = 1$.

Assumptions: None.

Page 341

Analysis: The function in Figure 5.18(a) is neither odd nor even. Thus, all the a_n and b_n coefficients are needed. First, express $x(t)$ over one period:

$$x(t) = \begin{cases} A & 0 \le t < \tau \\ 0 & \tau \le t < T \end{cases}$$

Then evaluate the integrals of equations 5.34 through 5.36:

$$a_0 = \frac{1}{T}\int_0^{T/5} A\,dt = \frac{A}{5}$$

$$a_n = \frac{2}{T}\int_0^{T} x(t)\cos\left(n\frac{2\pi}{T}t\right)dt = \frac{2}{T}\int_0^{T/5} A\cos\left(n\frac{2\pi}{T}t\right)dt + \int_{T/5}^{T} 0\cos\left(n\frac{2\pi}{T}t\right)dt$$

$$= \frac{2}{T}\frac{AT}{2n\pi}\sin\left(\frac{2n\pi}{T}t\right)\Big|_0^{T/5}$$

$$= \frac{2}{T}\frac{AT}{2n\pi}\left[\sin\left(\frac{2n\pi}{5}\right) - 0\right] = \frac{A}{n\pi}\sin\left(\frac{2n\pi}{5}\right)$$

$$b_n = \frac{2}{T}\int_0^{T/5} A\sin\left(n\frac{2\pi}{T}t\right)dt = \frac{2}{T}\frac{AT}{2n\pi}\left[-\cos\left(\frac{2n\pi}{T}t\right)\right]\Big|_0^{T/5}$$

$$= \frac{2}{T}\frac{AT}{2n\pi}\left[-\cos\left(\frac{2n\pi}{5}\right) + \cos(0)\right] = \frac{A}{n\pi}\left[1 - \cos\left(\frac{2n\pi}{5}\right)\right]$$

To compute the spectrum of the signal, apply equation 5.30:

$$c_n = \sqrt{a_n^2 + b_n^2} = \sqrt{\left[\frac{A}{n\pi}\sin\left(\frac{2n\pi}{5}\right)\right]^2 + \left\{\frac{A}{n\pi}\left[1 - \cos\left(\frac{2n\pi}{5}\right)\right]\right\}^2}$$

$$\theta_n = \cot^{-1}\left(\frac{b_n}{a_n}\right) = \cot^{-1}\left\{\frac{(A/n\pi)[1 - \cos(2n\pi/5)]}{(A/n\pi)\sin(2n\pi/5)}\right\}$$

The frequency spectrum of $x(t)$ (magnitude and phase) is shown in Figure 5.18(b). Table 5.1 lists the first seven coefficients in both forms.

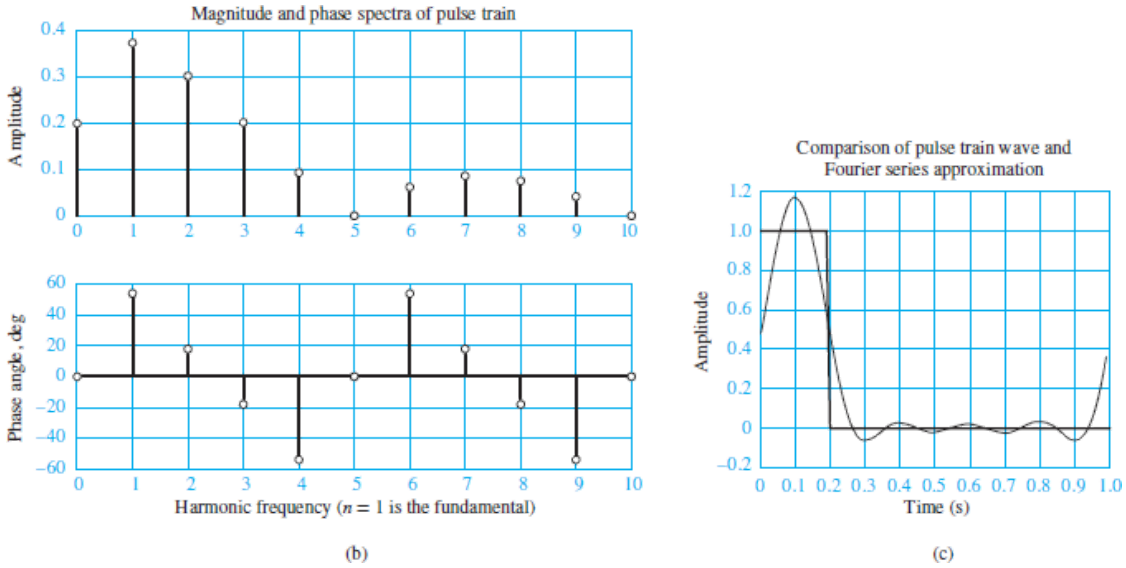


Figure 5.18 (b) Signal spectrum; (c) approximation obtained using 11 Fourier coefficients

Page 342

Table 5.1 Fourier coefficients of pulse train

n	a_n	b_n	c_n	θ_n (deg)
0	0.2	0	0.2	
1	0.3027	0.2199	0.3742	54
2	0.0935	0.2879	0.3027	18
3	−0.0624	0.1919	0.2018	−18
4	−0.0757	0.0550	0.0935	−54
5	0	0	0	0
6	0.0505	0.0367	0.0624	54
7	0.0267	0.0823	0.0865	18
8	−0.0234	0.0720	0.0757	−18
9	−0.0336	0.0244	0.0416	−54
10	0	0	0	0

Comments: A Fourier series approximation with the first 10 frequency components included can be generated and visualized with MatLab. Figure 5.18(c) shows the result.

EXAMPLE 5.5 Response of Linear System to a Periodic Input

Problem

A linear system with $\mathbf{H}(j\omega) = 2/(0.2j\omega + 1)$ is excited by the sawtooth waveform of Example 5.3 with $T = 0.25$ s and $A = 2$.

Solution

Known Quantities: $T = 0.25$ s; $A = 2$.

Find: Output $y(t)$ in response to input $x(t)$.

Assumptions: The waveform is well approximated by the first two terms of its Fourier series.

Analysis: The Fourier approximation of the sawtooth waveform of Example 5.3 is

$$x(t) = \frac{2A}{\pi}\sin\left(\frac{2\pi}{0.25}t\right) + \frac{A}{\pi}\sin\left(\frac{4\pi}{0.25}t\right) = \frac{4}{\pi}\sin(8\pi t) + \frac{2}{\pi}\sin(16\pi t)$$

Then:

$$c_1 = \sqrt{a_1^2 + b_1^2} = |b_1| = \frac{4}{\pi} \qquad \omega_1 = 1\,\omega_0 = 8\pi$$

and

$$c_2 = \sqrt{a_2^2 + b_2^2} = |b_2| = \frac{2}{\pi} \qquad \omega_2 = 2\,\omega_0 = 16\pi$$

The frequency response of the system can then be expressed in magnitude and phase form:

$$\mathbf{H}(j\omega) = \frac{2}{0.2j\omega + 1} = |\mathbf{H}(j\omega)|\angle\mathbf{H}(j\omega) = \frac{2}{\sqrt{(0.2\omega)^2 + 1}}\angle\left(-\arctan\frac{\omega}{5}\right)$$

The magnitude and phase of the frequency response are shown in Figure 5.19(a). Observe that the system is excited only at the frequencies $\omega_1 = 8\pi = 25.1$ rad/s and $\omega_2 = 16\pi = 50.2$ rad/s. Page 343The frequency response of the system can be evaluated at these frequencies either graphically, as in Figure 5.19(a), or analytically, as shown here.

$$|\mathbf{H}(j\omega_1)| = \frac{2}{\sqrt{(0.2\omega_1)^2 + 1}} = 0.3902 \angle \mathbf{H}(j\omega_1) = -1.37 \text{ rad} = -78.75°$$

$$|\mathbf{H}(j\omega_2)| = \frac{2}{\sqrt{(0.2\omega_2)^2 + 1}} = 0.1980 \angle \mathbf{H}(j\omega_2) = -1.47 \text{ rad} = -84.32°$$

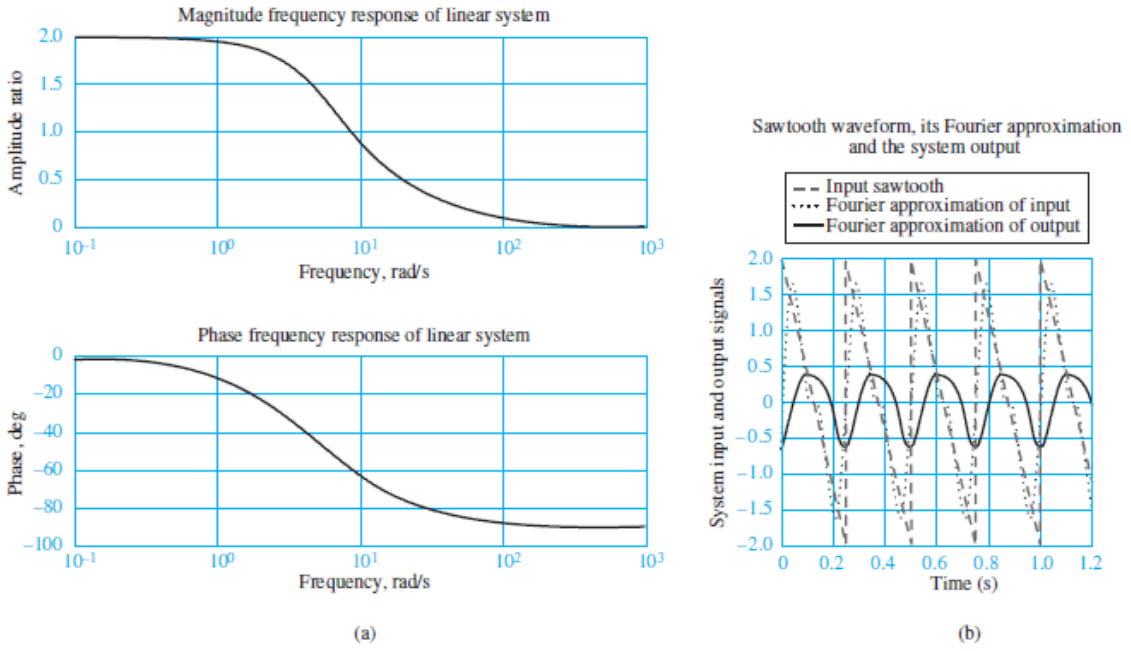


Figure 5.19 (a) Frequency response of linear system; (b) input and output waveforms

Finally, compute the steady-state periodic output of the system:

$$|\mathbf{H}(j\omega_1)| = \frac{2}{\sqrt{(0.2\omega_1)^2 + 1}} = 0.3902 \angle \mathbf{H}(j\omega_1) = -1.37 \text{ rad} = -78.75°$$

$$|\mathbf{H}(j\omega_2)| = \frac{2}{\sqrt{(0.2\omega_2)^2 + 1}} = 0.1980 \angle \mathbf{H}(j\omega_2) = -1.47 \text{ rad} = -84.32°$$

The input and output signals for the system are plotted in Figure 5.19(b). Note how the first two components of the Fourier series of the sawtooth waveform of Example 5.3 provide a coarse approximation of the waveform. Given the frequency response of the system used in this example, would the accuracy of the computed response increase if higher-frequency components ($n > 2$) were included in the Fourier series approximation?

Comments: MatLab has built-in functions to perform calculations of linear systems. It can also be used to approximate any input as a Fourier series.

CHECK YOUR UNDERSTANDING

How would the spectrum plot of Figure 5.17(b) change if the period of the waveform changed from 1 to 0.1 s?

Answer: The fundamental frequency and all harmonics would increase tenfold; the shape is unchanged.

CHECK YOUR UNDERSTANDING

If the *duty cycle* of the pulse train in Figure 5.18(a) is changed to $\tau/T = 0.25$, which of the Fourier coefficients are now zero?

Answer: $n = 4, 8$

CHECK YOUR UNDERSTANDING

Determine the a_n and b_n Fourier coefficients of the signal $y(t) = 1.5 \cos(100t + \pi/4)$. *(Hint:* Use trigonometric identities to expand the cosine function.)

Answer: $a_0 = 0, a_1 = 1.0607, b_1 = 1.0607$. All other coefficients are zero.

CHECK YOUR UNDERSTANDING

Extend the result of Example 5.5 to include three frequency components. What are the amplitude and phase of the components of $y(t)$ at the frequency $3\omega_0$?

Answer: magnitude = 0.0562, phase = −1.505 rad = −86.2°

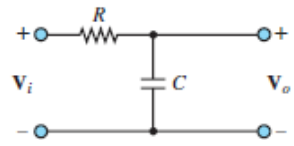
5.3 LOW- AND HIGH-PASS FILTERS

There are many practical applications that involve filters of one kind or another. Modern sunglasses filter out eye-damaging ultraviolet radiation and reduce the intensity of sunlight reaching the eyes. The suspension system of an automobile filters out road noise and reduces the impact of potholes on passengers. An analogous concept applies to electric circuits: It is possible to *attenuate* (i.e., reduce in amplitude) or altogether eliminate signals of unwanted frequencies, such as those that may be caused by electromagnetic interference (EMI).

Low-Pass Filters

Figure 5.20 depicts a simple ***RC* filter** and denotes its input and output voltages, respectively, by $\mathbf{V}_i$ and $\mathbf{V}_o$. The frequency response for the filter may be obtained by considering the function

$$\mathbf{H}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) \tag{5.39}$$



RC low-pass filter. The circuit preserves lower frequencies while attenuating the frequencies above the cutoff frequency $\omega_0 = 1/RC$. The voltages $\mathbf{V}_i$ and $\mathbf{V}_o$ are the filter input and output voltages, respectively.

Figure 5.20 A simple *RC* filter

Page 345

and noting that the output voltage may be expressed as a function of the input voltage by means of a voltage divider, as follows:

$$\mathbf{V}_o(j\omega) = \mathbf{V}_i(j\omega)\frac{1/j\omega C}{R + 1/j\omega C} = \mathbf{V}_i(j\omega)\frac{1}{1 + j\omega CR} \tag{5.40}$$

Thus, the frequency response of the *RC* filter is

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{1}{1 + j\omega CR} \tag{5.41}$$

An immediate observation upon studying this frequency response is that if the signal frequency ω is zero, the value of the frequency response function is 1. That is, the filter is passing all the input. Why? To answer this question, note that at $\omega = 0$, the impedance of the capacitor, $1/j\omega C$, becomes infinite. Thus, the capacitor acts as a DC open-circuit, and the output voltage equals the input:

$$\mathbf{V}_o(j\omega = 0) = \mathbf{V}_i(j\omega = 0) \tag{5.42}$$

As the signal frequency increases, the magnitude of the frequency response decreases since the magnitude and phase angle of the denominator increase with ω.

$$\begin{aligned} \mathbf{H}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) &= \frac{1}{1 + j\omega CR} \\ &= \frac{1}{\sqrt{1 + (\omega CR)^2}} \frac{e^{j0}}{e^{j \arctan(\omega CR/1)}} \\ &= \frac{1}{\sqrt{1 + (\omega CR)^2}} \cdot e^{-j \arctan(\omega CR)} \end{aligned} \tag{5.43}$$

or

$$\mathbf{H}(j\omega) = |\mathbf{H}(j\omega)| e^{j\angle \mathbf{H}(j\omega)} \tag{5.44}$$

with

$$|\mathbf{H}(j\omega)| = \frac{1}{\sqrt{1 + (\omega CR)^2}} = \frac{1}{\sqrt{1 + (\omega/\omega_0)^2}} \tag{5.45}$$

and

$$\angle \mathbf{H}(j\omega) = -\arctan(\omega CR) = -\arctan\frac{\omega}{\omega_0} \tag{5.46}$$

with

$$\omega_0 = \frac{1}{RC} \tag{5.47}$$

The simplest way to envision the effect of the filter is to think of the phasor voltage $\mathbf{V}_i = V_i e^{j\phi_i}$ scaled by a factor of $|\mathbf{H}|$ and shifted by a phase angle $\angle \mathbf{H}$ by the filter

at each frequency, so that the resultant output is given by the phasor $V_o e^{j\phi_o}$, with

$$\begin{aligned} V_o &= |\mathbf{H}| \cdot V_i \\ \phi_o &= \angle\mathbf{H} + \phi_i \end{aligned} \tag{5.48}$$

and where $|\mathbf{H}|$ and $\angle\mathbf{H}$ are functions of frequency. The frequency ω_0 is called the **cutoff** or **critical** or **break** frequency of the filter.

Page 346

It is customary to represent $\mathbf{H}(j\omega)$ in two separate plots, representing $|\mathbf{H}|$ and $\angle\mathbf{H}$ as functions of ω. These are shown in Figure 5.21 in normalized form, that is, with $|\mathbf{H}|$ and $\angle\mathbf{H}$ plotted versus ω/ω_0, corresponding to a cutoff frequency $\omega/\omega_0 = 1$ rad/s. Note that, in the plot, the frequency axis has been scaled logarithmically. Frequency response plots similar to those shown in Figure 5.21 are commonly employed in engineering. For example, the *RC* filter of Figure 5.20 has the property of "passing" signals at low frequencies ($\omega \ll 1/RC$) and of filtering out signals at high frequencies ($\omega \gg 1/RC$). This type of filter is called a **low-pass filter**. The cutoff frequency $\omega = 1/RC$ has a special significance in that it represents—approximately—the boundary between low- and high-frequencies. The value of $|\mathbf{H}(j\omega)|$ at the cutoff frequency is $1/\sqrt{2} = 0.707$. Note how the cutoff frequency depends exclusively on the values of *R* and *C*. Therefore, one can adjust the filter response as desired simply by selecting appropriate values for *C* and *R*.

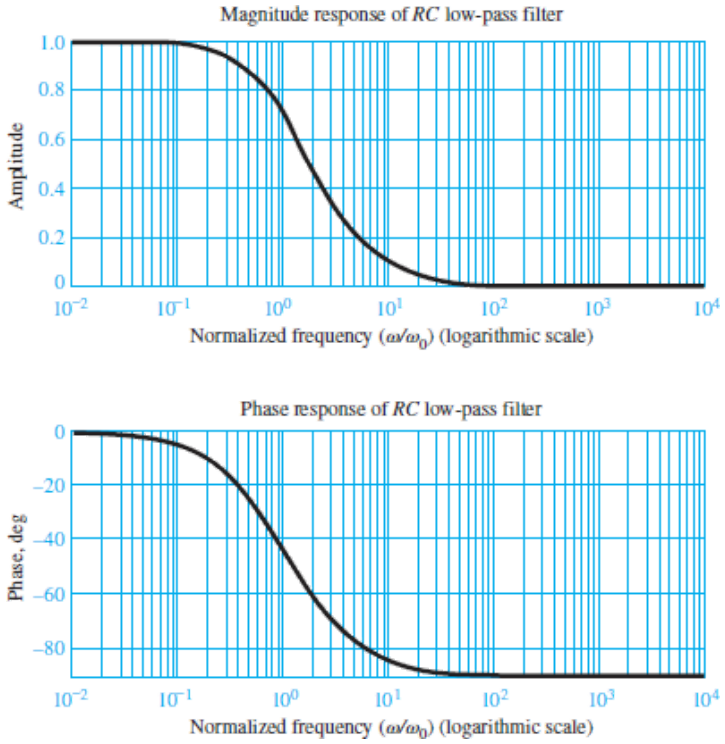


Figure 5.21 Magnitude and phase response plots for *RC* filter

Practical low-pass filters are often much more complex than simple *RC* combinations. The synthesis of such advanced filter networks is beyond the scope of this book; however, the implementation of some commonly used filters is discussed in Chapters 6 and 7, in connection with the operational amplifier.

High-Pass Filters

Just as a low-pass filter preserves low-frequency signals and attenuates those at higher frequencies, a **high-pass filter** attenuates low-frequency signals and preserves Page 347those at frequencies above its cutoff frequency. Consider the high-pass filter circuit shown in Figure 5.22. The frequency response is defined as:

$$\mathbf{H}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega)$$

RC high-pass filter. The circuit preserves higher frequencies while attenuating the frequencies below the cutoff frequency $\omega_0 = 1/RC$.

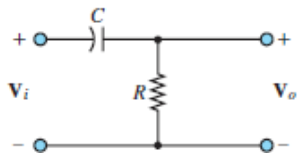


Figure 5.22 High-pass filter

Voltage division yields:

$$\mathbf{V}_o(j\omega) = \mathbf{V}_i(j\omega)\frac{R}{R + 1/j\omega C} = \mathbf{V}_i(j\omega)\frac{j\omega CR}{1 + j\omega CR} \tag{5.49}$$

Thus, the frequency response of the filter is

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega CR}{1 + j\omega CR} \tag{5.50}$$

which can be expressed in magnitude-and-phase form by

$$\mathbf{H}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega CR}{1 + j\omega CR} = \frac{\omega CR\, e^{j\pi/2}}{\sqrt{1 + (\omega CR)^2}\, e^{j \arctan(\omega CR/1)}} \tag{5.51}$$

$$= \frac{\omega CR}{\sqrt{1 + (\omega CR)^2}} \cdot e^{j[\pi/2 - \arctan(\omega CR)]}$$

or

$$\mathbf{H}(j\omega) = |\mathbf{H}|\, e^{j\angle \mathbf{H}}$$

with

$$|\mathbf{H}(j\omega)| = \frac{\omega CR}{\sqrt{1 + (\omega CR)^2}} \tag{5.52}$$

$$\angle \mathbf{H}(j\omega) = 90° - \arctan(\omega CR)$$

You can verify by inspection that the amplitude response of the high-pass filter will be zero at $\omega = 0$ and will asymptotically approach 1 as ω approaches infinity while the phase shift is $\pi/2$ at $\omega = 0$ and tends to zero for increasing ω. Amplitude-and-phase response curves for the high-pass filter are shown in Figure 5.23. These plots have been normalized to have the filter cutoff frequency $\omega/\omega_0 = 1$ rad/s. Note that, once again, it is possible to define a cutoff frequency at $\omega_0 = 1/RC$ in the same way as was done for the low-pass filter.

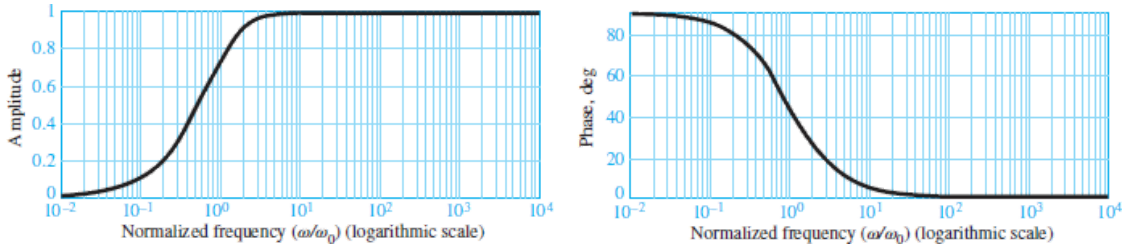


Figure 5.23 Frequency response of a high-pass filter

EXAMPLE 5.6 Frequency Response of *RC* Low-Pass Filter

Problem

Compute the response of the *RC* filter of Figure 5.20 to sinusoidal inputs at the frequencies of 60 and 10,000 Hz. Assume $R = 1\ \text{k}\Omega$, $C = 0.47\ \mu\text{F}$ and $\mathbf{V}_i = 5\angle 0°$ V.

Solution

Known Quantities: $R = 1\ \text{k}\Omega$; $C = 0.47\ \mu\text{F}$; $v_i(t) = 5\cos(\omega t)$ V.

Find: The output voltage $v_o(t)$ at each frequency.

Assumptions: None.

Analysis: In this problem, the input signal voltage and the frequency response of the circuit (equation 5.43) are known, and the output voltage must be found at two different frequencies. If the voltages are represented in phasor form, the frequency response can be used for calculation:

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \mathbf{G}_V(j\omega) = \frac{1}{1 + j\omega CR}$$

$$\mathbf{V}_o(j\omega) = \mathbf{G}_V(j\omega)\mathbf{V}_i(j\omega) = \frac{1}{1 + j\omega CR}\mathbf{V}_i(j\omega)$$

The cutoff frequency of the filter is $\omega_0 = 1/RC = 2{,}128$ rad/s such that the expressions for the frequency response in the form of equations 5.45 and 5.46 are

$$\mathbf{G}_V(j\omega) = \frac{1}{1 + j\omega/\omega_0} \qquad |\mathbf{G}_V(j\omega)| = \frac{1}{\sqrt{1 + (\omega/\omega_0)^2}} \qquad \angle\mathbf{G}(j\omega) = -\arctan\left(\frac{\omega}{\omega_0}\right)$$

Next, recognize that at $\omega = 60$ Hz $= 120\pi$ rad/s, the ratio $\omega/\omega_0 = 0.177$. At $\omega = 10{,}000$ Hz $= 20{,}000\pi$ rad/s, $\omega/\omega_0 = 29.5$. Thus, the output voltage at each frequency can be computed as follows:

$$\mathbf{V}_o(\omega = 2\pi 60) = \frac{1}{1 + j0.177}\mathbf{V}_i(\omega = 2\pi 60) = (0.985 \times 5)\angle -0.175\,\text{V}$$

$$\mathbf{V}_o(\omega = 2\pi 10{,}000) = \frac{1}{1 + j29.5}\mathbf{V}_i(\omega = 2\pi 10{,}000) = (0.0339 \times 5)\angle -1.537\,\text{V}$$

Finally, write the time-domain response for each frequency:

$$v_o(t) = 4.923\cos(2\pi 60t - 0.175)\,\text{V} \qquad \text{at } \omega = 2\pi 60 \text{ rad/s}$$
$$v_o(t) = 0.169\cos(2\pi 10{,}000t - 1.537)\,\text{V} \qquad \text{at } \omega = 2\pi 10{,}000 \text{ rad/s}$$

The magnitude and phase responses of the filter are plotted in Figure 5.24. It should be evident from these plots that only the low-frequency components of the signal are passed by the filter. This low-pass filter would pass only the *bass range* of the audio spectrum.

Comments: Approximate answers can be obtained quickly from the magnitude and phase plots of Figure 5.24. Simply multiply the input voltage amplitude (5 V) by the amplitude response at each frequency, and read off the phase shift at each frequency. The results should be close to the ones computed above.

Page 349

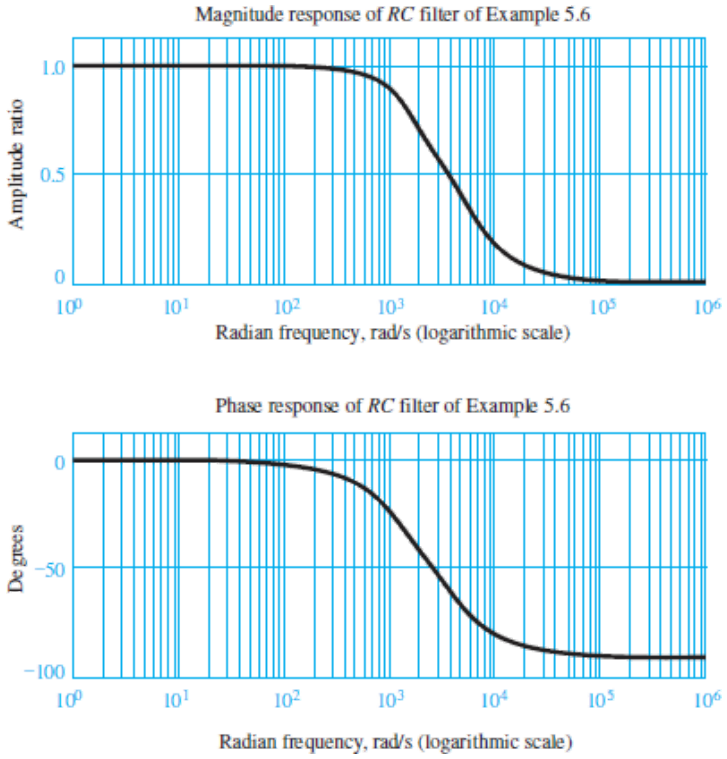


Figure 5.24 Response of *RC* low-pass filter of Example 5.6

EXAMPLE 5.7 A Realistic *RC* Low-Pass Filter Application

Problem

Determine the frequency response function $\mathbf{V}_o/\mathbf{V}_S$ and its frequency response from the network shown in Figure 5.25.

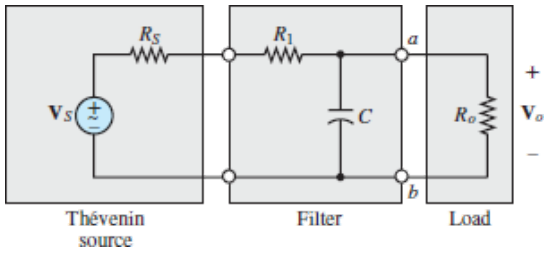


Figure 5.25 *RC* filter inserted in a circuit

Solution

Known Quantities: $R_S = 50\ \Omega$; $R_1 = 200\ \Omega$; $R_o = 500\ \Omega$; $C = 10\ \mu$F.

Find: The frequency response function $\mathbf{V}_o/\mathbf{V}_S$, its frequency response, and the output voltage $v_o(t)$ at given frequencies.

Assumptions: None.

Page 350

Analysis: Figure 5.25 represents a more realistic filtering circuit, in that an *RC* low-pass filter is inserted between the source and load circuits. The Thévenin equivalent impedance seen by the load is

$$\mathbf{Z}_T = \mathbf{Z}_C \| (R_1 + R_S) = \frac{(R_1 + R_S)/j\omega C}{R_1 + R_S + 1/j\omega C}$$

Multiply the numerator and denominator by $j\omega C$ to obtain:

$$\mathbf{Z}_T = \frac{R_1 + R_S}{1 + (j\omega)(R_1 + R_S)C}$$

Apply voltage division to find the Thévenin (open-circuit) voltage $\mathbf{V}_T$ across terminals *a* and *b*.

$$\mathbf{V}_T = \mathbf{V}_S \frac{1/j\omega C}{R_1 + R_S + 1/j\omega C}$$

Again, multiply the numerator and denominator by $j\omega C$ to obtain:

$$\mathbf{V}_T = \mathbf{V}_S \frac{1}{1 + (j\omega)(R_1 + R_S)C}$$

Next, apply voltage division to find $\mathbf{V}_o$.

$$\mathbf{V}_o = \mathbf{V}_T \frac{R_o}{R_o + \mathbf{Z}_T}$$

Substitute for $\mathbf{V}_T$ and $\mathbf{Z}_T$, and multiply the numerator and denominator by $[1 + (j\omega)(R_1 + R_S)C]$ to obtain:

$$\mathbf{V}_o = \mathbf{V}_S \frac{R_o}{R_o[1 + (j\omega)(R_1 + R_S)C] + (R_1 + R_S)}$$

Finally, divide both sides by $\mathbf{V}_S$ and factor $(R_o + R_1 + R_S)$ out of the denominator to find:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_S} = \frac{K}{1 + (j\omega)R_T C}$$

where

$$K = \frac{R_o}{R_o + R_1 + R_S}$$

and

$$R_T = [R_o \| (R_1 + R_S)] = \frac{R_o(R_1 + R_S)}{R_o + R_1 + R_S}$$

Again, K is the DC gain as $\omega \rightarrow 0$ and the capacitor acts as an open-circuit. R_T is the Thévenin equivalent resistance seen by the capacitor. Plug in values for the resistances and capacitance to find:

$$\mathbf{G}_V(j\omega) = \frac{0.667}{1 + j(\omega/600)}$$

Comments: The effect of placing the RC low-pass filter in the midst of the circuit is to shift the cutoff frequency from $1/R_1C$ to $1/R_T C$.

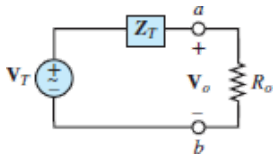


Figure 5.26 Equivalent circuit representation of Figure 5.25

EXAMPLE 5.8 Low-Pass Filter Attenuation

Problem

The frequency response of a particular 2nd-order low-pass filter is described by the following function. At what frequency has the magnitude of the response fallen to 10 percent of its maximum?

$$\mathbf{H}(j\omega) = \frac{K}{(j\omega/\omega_1 + 1)(j\omega/\omega_2 + 1)}$$

Page 351

Solution

Known Quantities: Frequency response function of a filter.

Find: Frequency $\omega_{10\%}$ at which the response amplitude equals 10 percent of its maximum.

Schematics, Diagrams, Circuits, and Given Data: $\omega_1 = 100$; $\omega_2 = 1{,}000$.

Assumptions: None.

Analysis: Notice that the magnitude of each term in the denominator increases with frequency ω. Thus, the maximum amplitude of the frequency response function is K, which is the DC gain as $\omega \rightarrow 0$. As frequency increases, the magnitude of the frequency response function decreases monotonically, which explains why the frequency response function describes a "low-pass" filter. At low frequencies, the

input is "passed" to the output; however, at higher frequencies the output is a filtered (reduced) version of the input. To solve this problem, set the amplitude of the frequency response function equal to 0.1 K and solve for ω, as follows:

$$|\mathbf{H}(j\omega)| = \left|\frac{K}{(j\omega/\omega_1 + 1)(j\omega/\omega_2 + 1)}\right| = 0.1K$$

$$\frac{1}{\sqrt{(1-\omega^2/\omega_1\omega_2)^2 + \omega^2(1/\omega_1 + 1/\omega_2)^2}} = 0.1$$

To simplify this expression introduce the dummy variable $\Omega = \omega^2$, and then invert and square both sides to obtain a quadratic equation in Ω:

$$\left(1 - \frac{\Omega}{\omega_1\omega_2}\right)^2 + \Omega\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right)^2 = 100$$

$$\Omega^2 + \left[(\omega_1\omega_2)^2\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right)^2 - 2\omega_1\omega_2\right]\Omega - 99(\omega_1\omega_2)^2 = 0$$

Plug in values for ω_1 and ω_2 and use the quadratic formula to solve for the two roots $\Omega = -1.6208 \times 10^6$ and $\Omega = 0.6108 \times 10^6$. Only the positive root has a physical meaning; thus, the solution is $\omega = \sqrt{\Omega} = 782$ rad/s.. Figure 5.27(a) depicts the magnitude response of the filter. At a frequency roughly equal to 800 rad/s, the magnitude response is approximately 0.1. At that frequency, the two poles of the frequency response function contribute roughly −82.7° and −38.0° to the overall phase angle for a total of −120.7°, as shown in Figure 5.27(b).

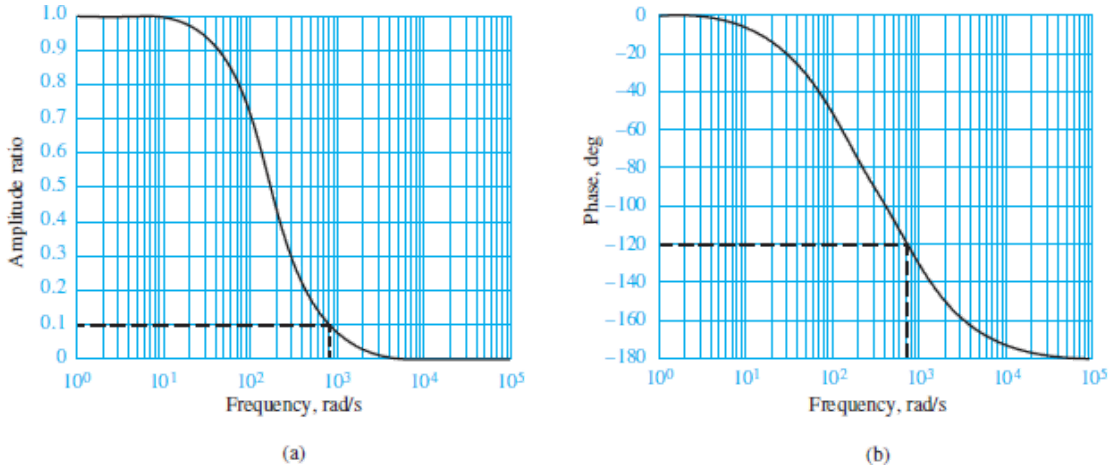


Figure 5.27 Frequency response of filter of Example 5.8. (a) Magnitude response; (b) phase response

EXAMPLE 5.9 Frequency Response of *RC* High-Pass Filter

Problem

Compute the response of the *RC* high-pass filter depicted in Figure 5.28. Evaluate the response of the filter at $\omega_1 = 2\pi \times 100$ and $\omega_2 = 2\pi \times 10{,}000$ rad/s.

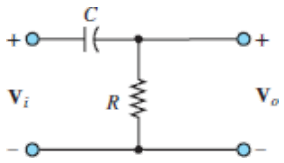


Figure 5.28 High-pass *RC* filter

Solution

Known Quantities: $R = 200\ \Omega$; $C = 0.199\ \mu$F.

Find: The frequency response $\mathbf{G}_V(j\omega)$.

Assumptions: None.

Analysis: The cutoff frequency of the high-pass filter is $\omega_0 = 1/RC \approx 2\pi \times 4{,}000$ rad/s. The frequency response function for the circuit is given by equation 5.50:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega CR}{1 + j\omega CR}$$
$$= \frac{\omega/\omega_0}{\sqrt{1 + (\omega/\omega_0)^2}} \angle \left[\frac{\pi}{2} - \arctan\left(\frac{\omega}{\omega_0}\right)\right]$$

The frequency response function can now be evaluated at ω_1 and ω_2:

$$\mathbf{G}_V(\omega = 2\pi \times 100) = \frac{100/4{,}000}{\sqrt{1 + (100/4{,}000)^2}} \angle \left[\frac{\pi}{2} - \arctan\left(\frac{100}{4{,}000}\right)\right] = 0.025\angle 88.6°$$

$$\mathbf{G}_V(\omega = 2\pi \times 10{,}000) = \frac{10{,}000/4{,}000}{\sqrt{1 + (10{,}000/4{,}000)^2}} \angle \left[\frac{\pi}{2} - \arctan\left(\frac{10{,}000}{4{,}000}\right)\right)$$
$$= 0.929\angle 21.8°$$

These results indicate that the output at $\omega_1 \ll \omega_0$ is very small (2.5 percent) compared to the input while at $\omega_2 \gg \omega_0$ the output is comparable (92.9 percent) to

the input. In general, the input is "passed" to the output at high frequencies ($\omega \gg \omega_0$) while at low frequencies ($\omega \ll \omega_0$) the output is a filtered (reduced) version of the input. The complete frequency response (amplitude and phase) is shown in Figure 5.29.

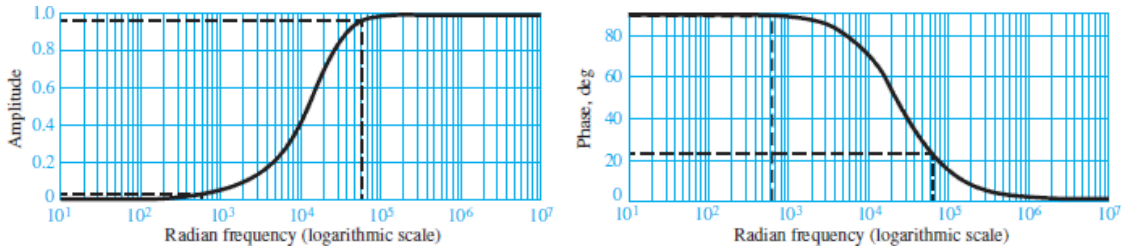


Figure 5.29 Response of high-pass *RC* filter of Example 5.10

Comments: With $\omega_0 = 2\pi \times 4{,}000$ (that is, 4,000 ^{H}z), this filter would pass only the *treble range* of the audio frequency spectrum.

Page 353

CHECK YOUR UNDERSTANDING

A simple *RC* low-pass filter is constructed using a 10-μF capacitor and a 2.2-kΩ resistor. Over what range of frequencies will the output of the filter be within 1 percent of the input signal amplitude (i.e., when will $V_o \leq 0.99\ V_S$)?

Answer: $0 \leq \omega \leq 6.48$ rad/s

CHECK YOUR UNDERSTANDING

In Figure 5.25, let $|\ \mathbf{V}_S\ |\ = 1$ V with an internal resistance $R_S = 50\ \Omega$. Assume $R_1 = 1$ kΩ and $C = 0.47\ \mu$F. Determine the cutoff frequency ω_0 for a load resistance $R_o = 470\ \Omega$.

Answer: $\omega_0 = 6{,}553.3$ rad/s

CHECK YOUR UNDERSTANDING

Use the phase response plot of Figure 5.27(b) to determine at which frequency the phase shift in the output signal (relative to the input signal) is equal to −90°.

Answer: $\omega = 300$ rad/s (approximately)

FOCUS ON MEASUREMENTS

Wheatstone Bridge Filter

Problem:

The Wheatstone bridge circuit of Example 2.2 and Focus on Measurements box, "Wheatstone Bridge and Force Measurements" in Chapter 2 is used in a number of instrumentation applications, including the measurement of force. Figure 5.30 depicts

the bridge circuit. When undesired noise and interference are present in a measurement, it is often appropriate to use a low-pass filter to reduce the effect of the noise. The capacitor that is connected to the output terminals of the bridge in Figure 5.30 constitutes an effective and simple low-pass filter, in conjunction with the bridge resistance. Assume that the average resistance of each leg of the bridge is 350 Ω (a standard value for strain gauges) and that a sinusoidal force is to be measured at a frequency of 30 Hz. From prior measurements, it has been determined that a filter with a cutoff frequency of 300 Hz is sufficient to reduce the effects of noise. Choose a capacitor that matches this filtering requirement.

Page 354

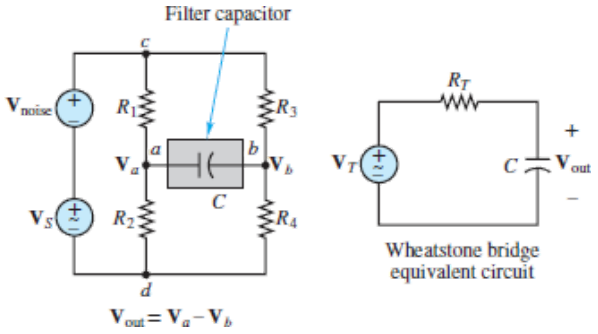


Figure 5.30 Wheatstone bridge with equivalent circuit and simple capacitive filter

Solution:

The Thévenin equivalent network seen by the capacitor can be determined, as illustrated on the right side of Figure 5.30. The Thévenin resistance seen by the capacitor is computed by turning off the two voltage sources and replacing them with short-circuits:

$$R_T = R_1 \| R_2 + R_3 \| R_4 = 350 \| 350 + 350 \| 350 = 350\ \Omega$$

Since the required cutoff frequency is 300 Hz, the capacitor value can be computed from the expression

$$\omega_0 = \frac{1}{R_T C} = 2\pi \times 300$$

or

$$C = \frac{1}{R_T \omega_0} = \frac{1}{350 \times 2\pi \times 300} = 1.51\ \mu\text{F}$$

The frequency response of the bridge circuit is of the same form as equation 5.41:

$$\frac{\mathbf{V}_{out}}{\mathbf{V}_T}(j\omega) = \frac{1}{1 + j\omega C R_T}$$

This response can be evaluated at the frequency of 30 Hz to verify that the attenuation and phase shift at the desired signal frequency are minimal:

$$\frac{\mathbf{V}_{out}}{\mathbf{V}_T}(j\omega = j2\pi \times 30) = \frac{1}{1 + j2\pi \times 30 \times 1.51 \times 10^{-6} \times 350}$$

$$= 0.9951\angle(-5.7°)$$

Page 355

Figure 5.31 depicts the appearance of a 30-Hz sinusoidal signal before and after the addition of the capacitor to the circuit.

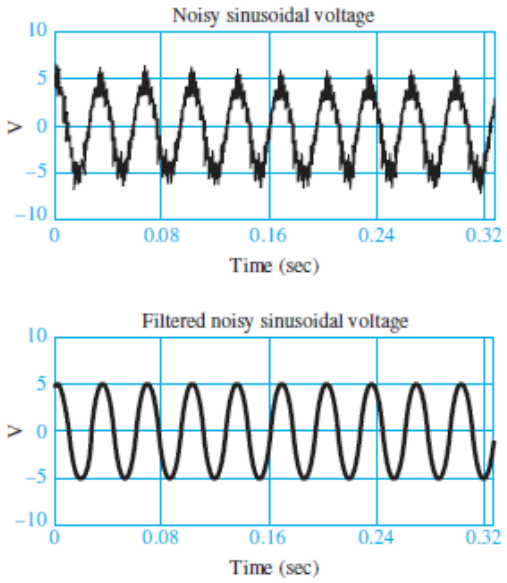


Figure 5.31 Unfiltered and filtered bridge output

5.4 BANDPASS FILTERS, RESONANCE AND QUALITY FACTOR

Using the same principles and procedures as before, it is possible to derive a **bandpass filter** response for particular types of circuits. Such a filter passes the input to the output at frequencies *within* a certain range. The analysis of a simple *second-order* (i.e., two energy storage elements) bandpass filter is similar to that of low- and

high-pass filters. Consider the circuit shown in Figure 5.32 (similar to Figure 5.6) and the designated frequency response function:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega)$$

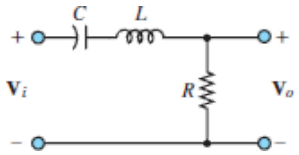


Figure 5.32 *RLC* bandpass filter.

Apply voltage division to find:

$$\begin{aligned}\mathbf{V}_o(j\omega) &= \mathbf{V}_i(j\omega)\frac{R}{R + 1/j\omega C + j\omega L}\frac{j\omega C}{j\omega C} \\ &= \mathbf{V}_i(j\omega)\frac{j\omega CR}{1 + j\omega CR + (j\omega)^2 LC}\end{aligned} \tag{5.53}$$

Thus, the frequency response function is

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega\tau}{1 + j\omega\tau + (j\omega)^2/\omega_n^2} \tag{5.54}$$

where $\tau = R_T C + L/R_N = RC$ and $\omega_n = 1/\sqrt{LC}$..

Page 356

Equation 5.54 contains a constant $K = \tau$, a zero at the origin ($j\omega$) and a quadratic pole $[1 + j\omega\tau + (j\omega/\omega_n)^2]$, which is similar to those present in Equations 5.19 and 5.24. Equation 5.54 can also be expressed in terms of the dimensionless damping coefficient ζ by substituting $2\zeta/\omega_n$ for τ. When $\zeta > 1$, equation 5.54 can be factored to yield the following form:

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{(2\zeta)(j\omega/\omega_n)}{(j\omega/\omega_1 + 1)(j\omega/\omega_2 + 1)} \tag{5.55}$$

where $\omega_1 = \omega_n(\zeta - \sqrt{\zeta^2 - 1})$ and $\omega_2 = \omega_n(\zeta + \sqrt{\zeta^2 - 1})$ are the two **half-power frequencies** that determine the **passband** (or **bandwidth**) of the filter—that is, the frequency range over which the filter "passes" the input signal through to the output. When $\zeta < 1$, the frequency response function of Equation 5.54 still has two well-defined half-power frequencies and an associated bandwidth. However, they must be calculated as described in the following section on resonance and bandwidth.

Notice that as $\omega \rightarrow 0$ the filter response approaches zero since the impedance of the capacitor $1/j\omega C$ approaches infinity (i.e., gets very large.) That is, the capacitor acts as an open-circuit and the output voltage equals zero. Further, as $\omega \rightarrow \infty$, the filter response again approaches zero since the impedance of the inductor $j\omega L$ approaches infinity. That is, the inductor acts as an open-circuit. Thus, the filter does not pass signals at very low or very high frequencies.

In an intermediate band of frequencies, the filter will, to some degree, pass the input signal to the output, the extent of which depends upon the frequency of the input signal. In fact, at $\omega = \omega_n$, $\mathbf{V}_o = \mathbf{V}_i$! At this frequency, $\mathbf{Z}_C = -\mathbf{Z}_L$ such that the impedance seen by $\mathbf{V}_i$ is minimized and equal to R.

The frequency response function shown in Equation 5.55 is a dimensionless gain $\mathbf{G}_V$, where

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega)$$

This gain can be expressed in terms of a magnitude and phase angle $\mathbf{G}_V(j\omega) = |\mathbf{G}_V|e^{j\angle \mathbf{G}_V}$, where

$$|\mathbf{G}_V(j\omega)| = \frac{(2\zeta)(\omega/\omega_n)}{\sqrt{[1+(\omega/\omega_1)^2][1+(\omega/\omega_2)^2]}} \tag{5.56}$$

and

$$\angle \mathbf{G}_V(j\omega) = \frac{\pi}{2} - \arctan\frac{\omega}{\omega_1} - \arctan\frac{\omega}{\omega_2} \tag{5.57}$$

The magnitude and phase plots for the frequency response of the bandpass filter of Figure 5.32 are shown in Figure 5.33. These plots have been normalized to have the filter passband centered at the frequency $\omega = 1$ rad/s.

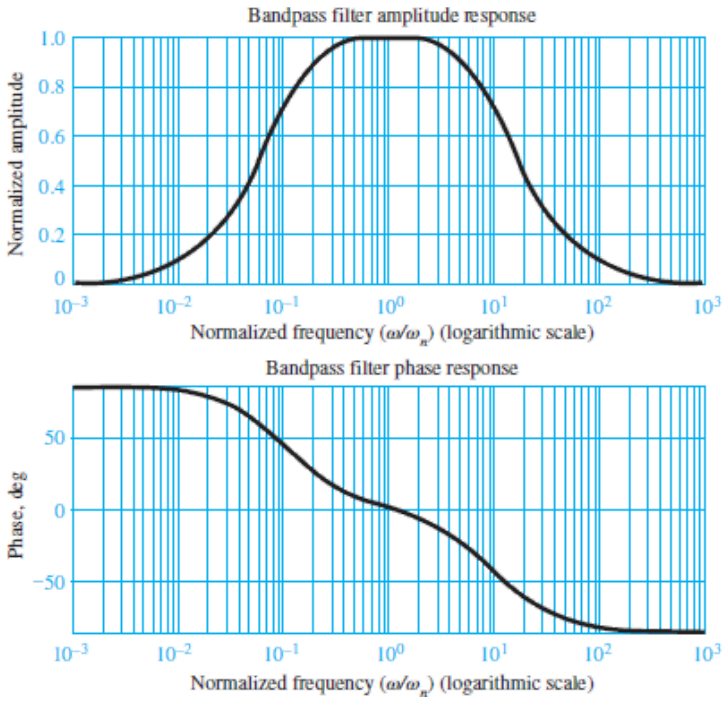


Figure 5.33 Frequency response of *RLC* bandpass filter

The frequency response plots of Figure 5.33 suggest that, in some sense, the bandpass filter acts as a combination of a high-pass and a low-pass filter. As illustrated in the previous cases, it should be evident that one can adjust the filter response as desired simply by selecting appropriate values for L, C, and R.

Resonance and Bandwidth

The response of second-order filters can be explained more generally by way of the series LC bandpass filter of Figure 5.32 in the following forms:

Page 357

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega\tau}{LC(j\omega)^2 + j\omega\tau + 1} = \frac{(2\zeta/\omega_n)j\omega}{(j\omega/\omega_n)^2 + (2\zeta/\omega_n)j\omega + 1} = \frac{(1/Q\omega_n)j\omega}{(j\omega/\omega_n)^2 + (1/Q\omega_n)j\omega + 1} \tag{5.58}$$

where $\tau = R_T C + L/R_N$, R_T is the Thévenin equivalent resistance seen by the capacitor, R_N is the Norton equivalent resistance seen by the inductor, ζ is the dimensionless

damping coefficient, ω_n is the natural frequency and Q is the *quality factor* defined by:[2]

$$\omega_n = \sqrt{\frac{1}{LC}} \quad \text{natural or resonant frequency}$$

$$Q = \frac{1}{2\zeta} = \frac{1}{\omega_n \tau} \quad \text{quality factor}$$

$$\zeta = \frac{\omega_n \tau}{2} = \frac{\omega_n}{2}\left[R_T C + \frac{L}{R_N}\right] \quad \text{dimensionless damping coefficient}$$

(5.59)

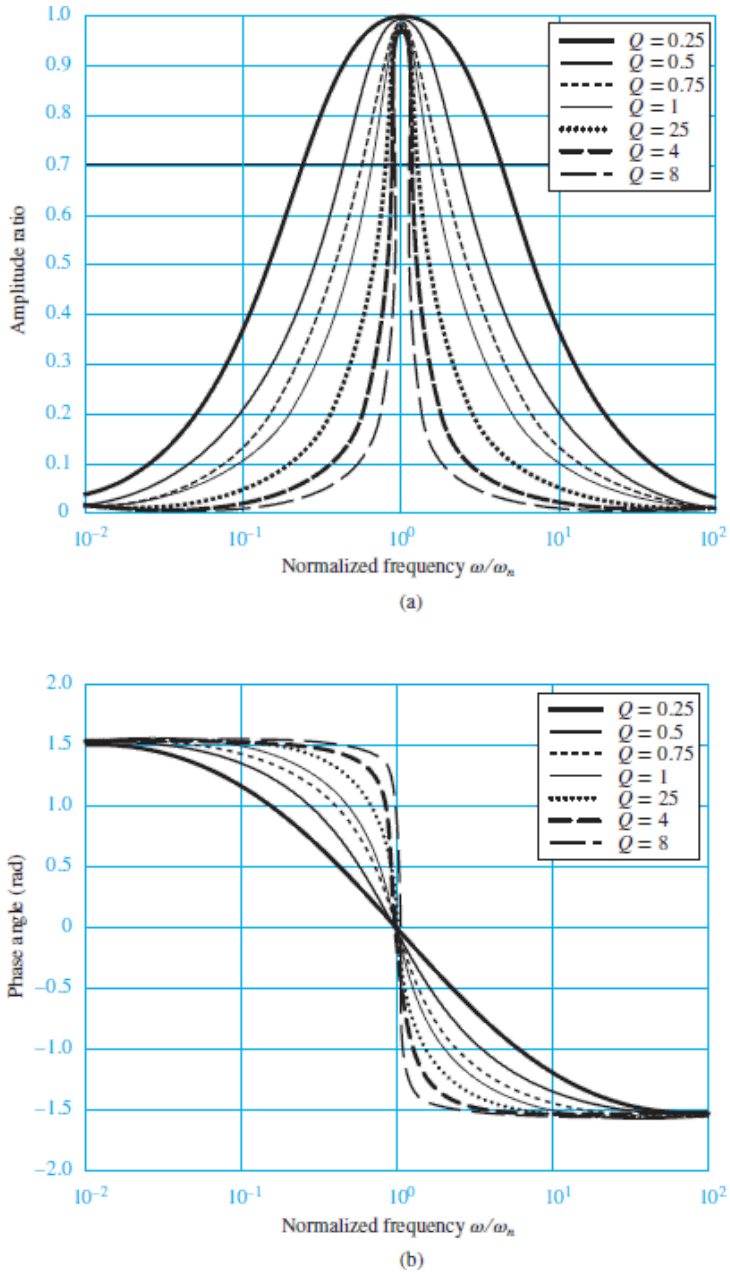


Figure 5.34 (a) Normalized magnitude response of second-order bandpass filter; (b) normalized phase response of second-order bandpass filter

Figure 5.34 depicts the normalized frequency response (magnitude and phase) of the second-order bandpass filter for various values of Q. The peak displayed in the frequency response around the frequency ω_n is called a *resonant peak,* and ω_n is the **resonant frequency**. Note that as the **quality factor** Q increases, the sharpness of the resonance increases and the filter becomes increasingly *selective* (i.e., it has the ability to filter out most frequency components of the input signals Page 359except for a narrow band around the resonant frequency). One measure of the selectivity of

a bandpass filter is its **bandwidth**. The concept of bandwidth can be easily visualized in the plot of Figure 5.34(a) by drawing a horizontal line across the plot at the 0.707 amplitude ratio. The frequency range between (magnitude) frequency response points intersecting this horizontal line is defined as the **half-power bandwidth** of the filter. The name *half-power* stems from the fact that when the voltage or current amplitude ratio is equal to 0.707 (or $1/\sqrt{2}$), the power has been reduced by a factor of $\frac{1}{2}$. The frequencies at which the 0.707 line intersects the frequency response are the **half-power frequencies** ω_1 and ω_2. For the bandpass filter shown in Figure 5.32, when $\zeta > 1$, $\omega_{2,1} = \omega_n(\zeta \pm \sqrt{\zeta^2 - 1})$. For $\zeta < 1$, $\zeta < 1, \omega_{2,1}^2 = \omega_n^2[1 + 2\zeta(\zeta \pm \sqrt{\zeta^2 + 1})]$. These expressions have useful approximations for large and small values of ζ. As $\zeta \to \infty$, $\omega_2 \to 2\zeta$ and $\omega_1 \to (1/2\zeta)$. As $\zeta \to 0$, $\omega_{2,1} \to (1 \pm \zeta)$.

A useful expression relating the bandwidth BW to the natural frequency ω_n and the quality factor Q is shown below. Note that a high-Q filter has a narrow bandwidth and a low-Q filter has a wide bandwidth.

$$BW = \frac{\omega_n}{Q} = \omega_2 - \omega_1 \quad \text{where} \quad \omega_n^2 = \omega_2\omega_1 \qquad \text{Bandwidth} \tag{5.60}$$

EXAMPLE 5.10 Frequency Response of Bandpass Filter

Problem

Compute the frequency response of the bandpass filter of Figure 5.32 for two sets of component values.

Solution

Known Quantities:

a. $R = 1$ kΩ; $C = 10$ μF; $L = 5$ mH.
b. $R = 10$ Ω; $C = 10$ μF; $L = 5$ mH.

Find: The frequency response $\mathbf{H}_V(j\omega)$.

Assumptions: None.

Analysis: The frequency response of the bandpass filter is expressed by equation 5.54.

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega\tau}{1 + j\omega\tau + (j\omega)^2\omega_n}$$

Here, $\tau = R_T C + L/R_N = RC$ and $\omega_n = 1/\sqrt{LC}$. For case a. $\tau = 10^{-2}$ seconds while for case b. $\tau = 10^{-4}$ seconds. For both cases, $\omega_n \approx 4472$ rad/sec. Thus, the dimensionless damping coefficients are $\zeta \approx 22.4$ and $\zeta \approx 0.224$ for cases a. and b., respectively. The frequency response plots for case a. (large series resistance) are shown in Figure 5.35. Those for case b. (small series resistance) are shown in Figure 5.36. (See section 5.5 to learn how to construct, Page 360by hand, straight line asymptotic approximations of these Bode plots.) Since L and C are the same in both cases, the *natural frequency* of the two circuits is the same:

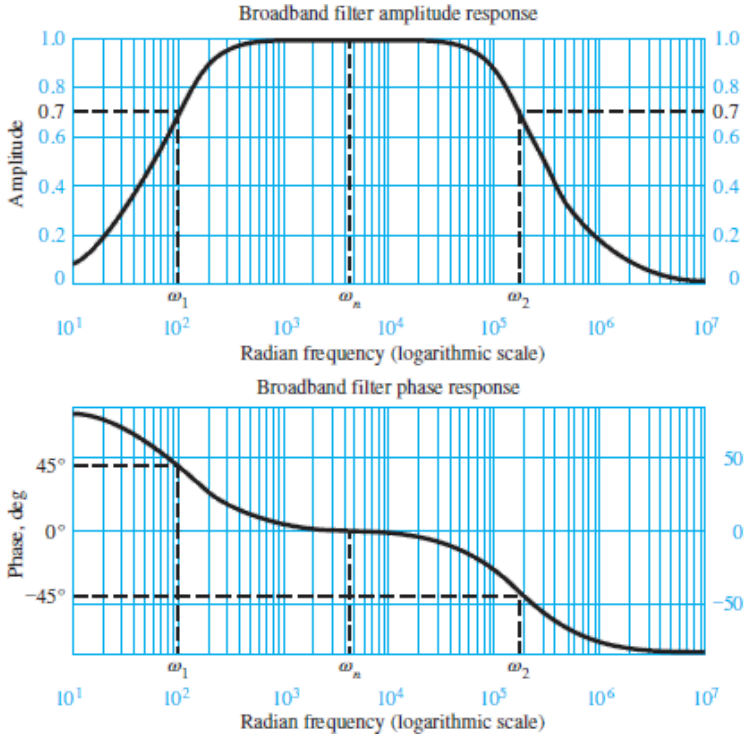


Figure 5.35 Frequency response of broadband (overdamped) bandpass filter of Example 5.10

$$\omega_n = \frac{1}{\sqrt{LC}} = 4.47 \times 10^3 \text{ rad/s}$$

On the other hand, the *quality factor* Q is substantially different:

a. $Q_a = \dfrac{1}{2\zeta_a} = \dfrac{1}{\omega_n \tau_a} \approx 0.022$

b. $Q_b = \dfrac{1}{2\zeta_b} = \dfrac{1}{\omega_n \tau_b} \approx 2.2$

The approximate bandwidths of the two filters are

$$BW_a = \frac{\omega_n}{Q_a} \approx 200{,}000 \text{ rad/s} \qquad \text{case}$$

$$BW_b = \frac{\omega_n}{Q_b} \approx 2{,}000 \text{ rad/s} \qquad \text{case}$$

For case a., the two half-power frequencies are $\omega_1 \approx 0.1$ krad/sec and $\omega_2 \approx 199.9$ krad/sec. It is worth noting that for $\zeta \to \infty$, $\omega_1 \to \omega_n/2\zeta$ and $\omega_2 \to 2\zeta\omega_n$. For case a., $\omega_n/2\zeta = 100.0$ rad/sec and $2\omega_n\zeta = 2.0$ krad/sec, which are very close to the calculated values for ω_1 and ω_2.

For case b., the two half-power frequencies are $\omega_1 \approx 3.6$ krad/sec and $\omega_2 \approx 5.6$ krad/sec. It is worth noting that for $\zeta \to 0$, $\omega_1 \to \omega_n(1 - \zeta)$ and $\omega_2 \to \omega_n(1 + \zeta)$. For case b., $\omega_n(1 - \zeta) \approx 3.47$ krad/sec and $\omega_n(1 - \zeta) \approx 5.47$ krad/sec, which are very close to the calculated values for ω_1 and ω_2.

Page 361

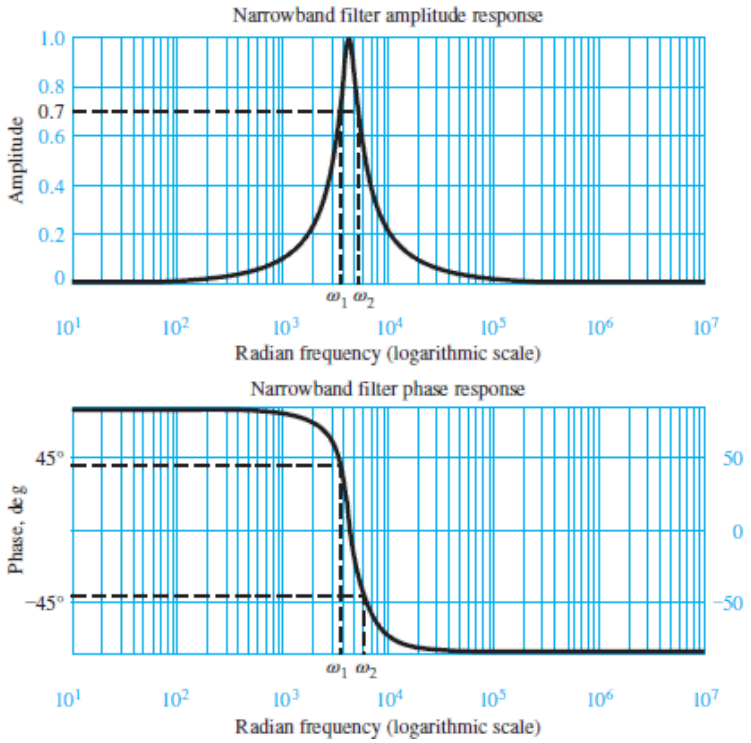


Figure 5.36 Frequency response of narrowband (underdamped) bandpass filter of Example 5.10

Comments: It should be apparent that while at the higher and lower frequencies most of the amplitude of the input signal is filtered from the output, at the midband frequency most of the input signal amplitude passes through the filter. The first bandpass filter analyzed in this example would "pass" the *midband range* of the audio spectrum while the second would pass only a very narrow band of frequencies around the **center frequency**. Such narrowband filters find application in **tuning circuits**, such as those employed in conventional AM radios. In a tuning circuit, a narrowband filter is used to tune in a frequency associated with the **carrier wave** of a radio station (e.g., for a station found at a setting of AM 820, the carrier wave transmitted by the radio station is at a frequency of 820 kHz). By using a variable capacitor, it is possible to tune in a range of carrier frequencies and therefore select the preferred station. Other circuits are then used to decode the actual speech or music signal modulated on the carrier wave.

CHECK YOUR UNDERSTANDING

Compute the half-power frequencies ω_1 and ω_2 for the bandpass filter of Example 5.10 (with R = 1 kΩ) by equating the magnitude of the bandpass filter frequency response to $1/\sqrt{2}$. The result is a quadratic equation in ω.

Answer: $\omega_1 = 99.95$ rad/s; $\omega_2 = 200.1$ krad/s

FOCUS ON MEASUREMENTS

AC Line Interference Filter

Problem:

One application of narrowband filters is seen in rejecting interference due to AC line power. Any undesired 60-Hz signal originating in the AC line can cause serious interference in sensitive instruments. In medical instruments such as the electrocardiograph, 60-Hz notch filters are often provided to reduce the effect of this interference[1] on cardiac measurements. Figure 5.37 depicts a circuit in which the effect of 60-Hz noise is represented by way of a 60-Hz sinusoidal generator connected in series with a signal source ($\mathbf{V}_S$), representing the desired signal. In this example we design a 60-Hz narrowband (or *notch*) filter to remove the unwanted 60-Hz noise.

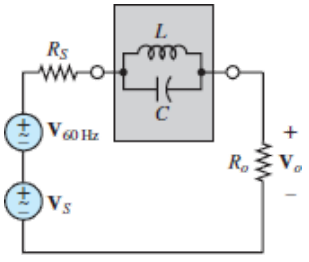


Figure 5.37 60-Hz notch filter

Solution:

Known Quantities—$R_S = 50\ \Omega$.

Find—Appropriate values of L and C for the notch filter.

Assumptions—None.

Analysis—To determine the appropriate capacitor and inductor values, write the expression for the notch filter impedance:

$$\mathbf{Z}_{\|} = \mathbf{Z}_L \| \mathbf{Z}_C = \frac{j\omega L / j\omega C}{j\omega L + 1/j\omega C}$$
$$= \frac{j\omega L}{1 + (j\omega)^2 LC} = \frac{j\omega L}{1 - \omega^2 LC}$$

Note that when $\omega^2 LC = 1$, the impedance of the circuit is infinite! The frequency

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

is the resonant frequency of the LC circuit. If this resonant frequency were selected to be equal to 60 Hz, then the series circuit would show an infinite impedance to 60-Hz currents and would therefore block the interference signal, while passing most of the other frequency components. Select values of L and C that result in $\omega_0 = 2\pi \times 60$. Let $L = 100$ mH. Then

$$C = \frac{1}{\omega_0^2 L} = 70.36\ \mu\text{F}$$

Page 363

The frequency response of the complete circuit is given below:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)} = \frac{R_o}{R_S + R_o + \mathbf{Z}_{\|}}$$
$$= \frac{R_o}{R_S + R_o + j\omega L / [1 + (j\omega)^2 LC]} = \frac{R_o}{R_S + R_o} \frac{1 + (j\omega)^2 LC}{1 + (j\omega)\tau + (j\omega)^2 LC}$$

where $\tau = L/(R_S + R_o)$. This response is plotted in Figure 5.38.

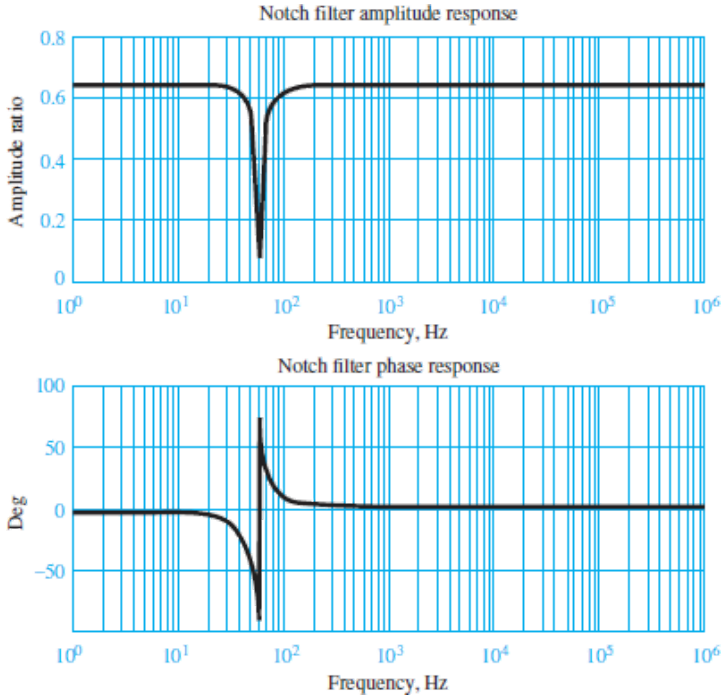


Figure 5.38 Frequency response of 60-Hz notch filter

Comments—It is instructive to calculate the response of the notch filter at frequencies in the immediate neighborhood of 60 Hz, to verify the attenuation effect of the notch filter. See section 5.5 to learn how to construct, by hand, straight line asymptotic approximations of Bode magnitude and phase plots.

FOCUS ON MEASUREMENTS

Seismic Transducer

The configuration of a *seismic displacement transducer* is shown in Figure 5.39. The transducer is housed in a case rigidly affixed to the surface of a body whose motion is to be measured. Thus, the case will experience the same displacement as the body, x_i. Inside the case, a small mass M rests on a spring characterized by stiffness K, placed in parallel with a damper B. The wiper arm of a potentiometer is connected to the floating mass M; the potentiometer is attached to the transducer case, so that the voltage V_o is proportional to the *relative displacement of the mass with respect to the case* x_o.

Page 364

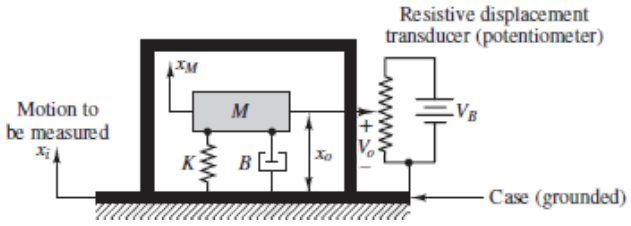


Figure 5.39 Seismic displacement transducer

The equation of motion for the mass-spring-damper system may be obtained by summing all the forces acting on mass M:

$$-Kx_o - B\frac{dx_o}{dt} = M\frac{d^2x_M}{dt^2} = M\left(\frac{d^2x_i}{dt^2} + \frac{d^2x_o}{dt^2}\right)$$

where the motion of the mass is equal to the difference between the motion of the case and the motion of the mass relative to the case itself; that is,

$$x_o = x_M - x_i$$

Assume the relative motion of the mass is periodic such that Fourier analysis can be used to decompose it into a summation of sinusoids. Each of these sinusoids can be expressed in phasor form.

$$\mathbf{X}_i(j\omega) = |X_i|e^{j\phi_i} \quad \text{and} \quad \mathbf{X}_o(j\omega) = |X_o|e^{j\phi_o}$$

Recall from Chapter 3 that taking the derivative of a phasor corresponds to multiplying the phasor by $j\omega$ such that the second-order differential equation can be

expressed in phasor form.

$$M(j\omega)^2\mathbf{X}_o + B(j\omega)\mathbf{X}_o + K\mathbf{X}_o = -M(j\omega)^2\mathbf{X}_i$$
$$(-\omega^2 M + j\omega B + K)\mathbf{X}_o = \omega^2 M\mathbf{X}_i$$

and we can write an expression for the frequency response:

$$\frac{\mathbf{X}_o(j\omega)}{\mathbf{X}_i(j\omega)} = \mathbf{G}(j\omega) = \frac{\omega^2 M}{-\omega^2 M + j\omega B + K}$$

The frequency response of the transducer is plotted in Figure 5.40 for the component values $M = 0.005$ kg and $K = 1{,}000$ N/m and for three values of B:

$B = 10$ N-s/m	($Q \approx 0.22$)	dotted line
$B = 2$ N-s/m	($Q \approx 1.1$)	dashed line
$B = 1$ N-s/m	($Q \approx 2.2$)	solid line

Notice that as the damping B decreases the quality factor Q of the system response increases.

The transducer clearly displays a high-pass response, indicating that for a sufficiently high input signal frequency, the measured displacement x_o (proportional to the voltage V_o) is equal to the input displacement x_i, which is the desired quantity. Note how sensitive the frequency response of the transducer is to changes in damping: as B changes from 2 to 1, a sharp **resonant peak** appears around the frequency $\omega = 316$ rad/s (approximately 50 Hz). As B increases to a value of 10, the amplitude response curve Page 365shifts to the right. Thus, this transducer, with the preferred damping given by $B = 2$, would be capable of correctly measuring displacements at frequencies above a minimum value, about 1,000 rad/s (or 159 Hz). The choice of $B = 2$ as the preferred design may be explained by observing that, ideally, we would like to obtain a constant amplitude response at all frequencies. The magnitude response that most closely approximates the ideal case in Figure 5.40 corresponds to $B = 2$. This concept is commonly applied to a variety of vibration measurements. See section 5.5 to learn how to construct, by hand, straight line asymptotic approximations of these Bode magnitude and phase plots.

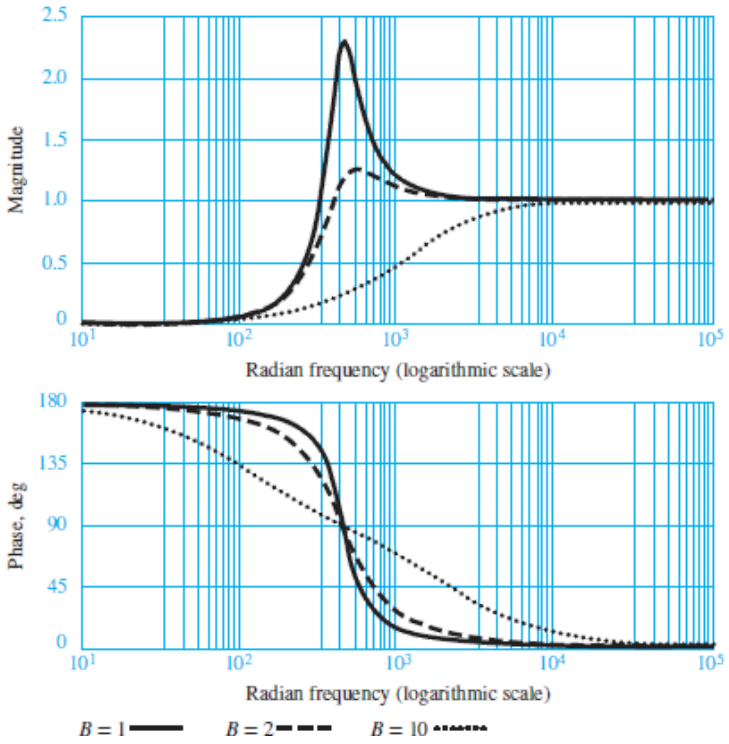


Figure 5.40 Frequency response of seismic transducer

A second-order circuit such as that shown in Figure 5.41 can exhibit the same type of response as the seismic transducer.

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{j\omega L}{R + 1/j\omega C + j\omega L} = \frac{(j\omega L)(j\omega C)}{j\omega CR + 1 + (j\omega L)(j\omega C)}$$
$$= \frac{(j\omega)^2 LC}{(j\omega)^2 LC + (j\omega)RC + 1} = \frac{-\omega^2 LC}{-\omega^2 LC + j\omega RC + 1}$$

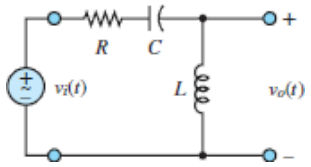


Figure 5.41 Electric circuit analog of the seismic transducer

Page 366

Compare this expression with the frequency response of the seismic transducer. Note that the mass M plays a role analogous to that of the inductance L. The damper B acts in analogy with the resistor R; and the spring K is analogous to the inverse of the capacitance, C. This analogy between the mechanical and electric systems derives from the fact that the equations describing the two systems have the same form.

Engineers often use such analogies to construct electrical *models,* or *analogs,* of mechanical systems. For example, to study the behavior of a large mechanical system, it might be easier and less costly to start by modeling the mechanical system with an inexpensive electric circuit and testing the model, rather than the full-scale mechanical system.

5.5 BODE PLOTS

Frequency response plots of linear systems are often displayed in the form of logarithmic plots, called **Bode plots** after the mathematician Hendrik W. Bode, where the horizontal axis represents frequency on a logarithmic scale (base 10) and the vertical axis represents either the amplitude or phase of the frequency response function. In Bode plots the amplitude is expressed in units of **decibels (dB)**, where

$$\left|\frac{\mathbf{X}_o}{\mathbf{X}_i}\right|_{\text{dB}} = 20 \log_{10}\left|\frac{\mathbf{X}_o}{\mathbf{X}_i}\right| = 20 \log_{10}\frac{|\mathbf{X}_o|}{|\mathbf{X}_i|} \tag{5.61}$$

In general, the argument of the common logarithm shown in Equation 5.61 is a ratio of standard terms, which are known as *zeros* and *poles* when present in the numerator and denominator, respectively. While logarithmic plots may at first seem a daunting complication, they have two significant advantages:

1. The product of terms in a frequency response function becomes a sum of terms because log (ab/c) = log(a) + log(b) − log(c). The advantage here is that Bode (logarithmic) plots can be constructed by summing the plots of individual terms. Moreover, as was discussed in section 5.1, there are only four distinct types of **standard terms** present in any frequency response function:
   a. A constant K.
   b. Poles or zeros "at the origin"$(j\omega)$.
   c. Simple poles or zeros $(1 + j\omega\tau)$ or $(1 + j\omega/\omega_0)$.
   d. Quadratic poles or zeros $[1 + j\omega\tau + (j\omega/\omega_n)^2]$.
2. The individual Bode plots of these four distinct terms are all well approximated by linear segments, which are readily summed to form the overall Bode plot of more complicated frequency response functions.

RC Low-Pass Filter Bode Plots

Consider, for example, the RC low-pass filter of Example 5.6 (Figure 5.20). The frequency response function is

$$\frac{\mathbf{V}_o}{\mathbf{V}_i}(j\omega) = \frac{1}{1 + j\omega/\omega_0} = \frac{1}{\sqrt{1 + (\omega/\omega_0)^2}} \angle -\tan^{-1}\left(\frac{\omega}{\omega_0}\right) \tag{5.62}$$

with a time constant $\tau = RC = 1/\omega_0$, where ω_0 is the break, or half-power, frequency of the filter. This frequency response function contains a constant $K = 1$ and a simple pole with cutoff frequency $\omega_0 = 1/\tau = 1/RC$.

Page 367

Figure 5.42 shows the Bode magnitude and phase plots for the filter. The normalized frequency on the horizontal axis is $\omega\tau$. The magnitude plot is obtained from the logarithmic form of the absolute value of the frequency response function:

$$\left|\frac{\mathbf{V}_o}{\mathbf{V}_i}\right|_{\text{dB}} = 20\log_{10}\frac{|K|}{|1+j\omega\tau|} = 20\log_{10}\frac{|K|}{|1+j\omega/\omega_0|} \tag{5.63}$$

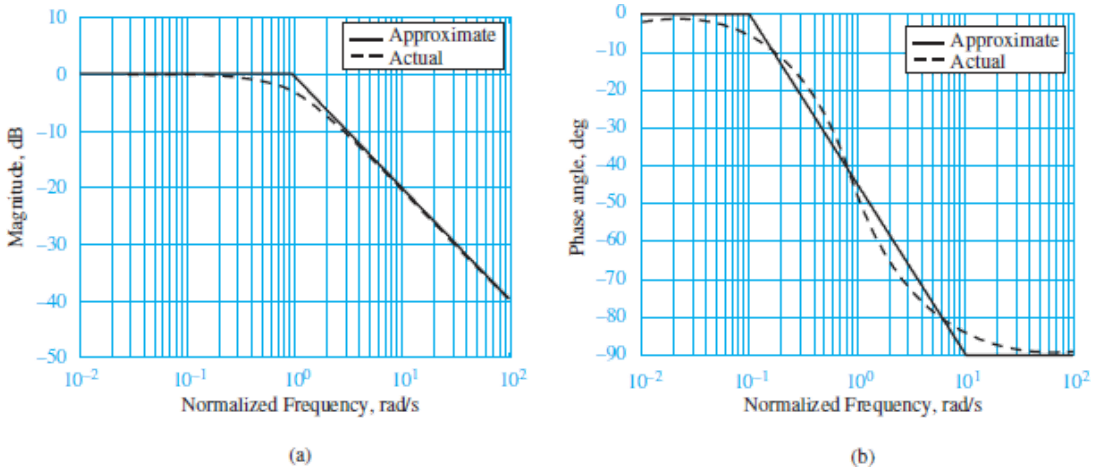


Figure 5.42 Bode plots for a low-pass *RC* filter; the frequency variable is normalized to ω/ω_0. (a) Magnitude response; (b) phase angle response

a

When $\omega \ll \omega_0$, the imaginary part of the simple pole is much smaller than its real part, such that $|\,1 + j\omega/\omega_0\,| \approx 1$. Then:

$$\left|\frac{\mathbf{V}_o}{\mathbf{V}_i}\right|_{\text{dB}} \approx 20\log_{10}K - 20\log_{10}1 = 20\log_{10}K \quad \text{(dB)} \tag{5.64}$$

Thus, at very low frequencies ($\omega \ll \omega_0$), equation 5.63 is well approximated by a straight line of zero slope, which is the *low-frequency asymptote* of the Bode magnitude plot.

When $\omega \gg \omega_0$, the imaginary part of the simple pole is much larger than its real part, such that $| 1 + j\omega/\omega_0 | \approx | j\omega/\omega_0 | = (\omega/\omega_0)$. Then:

$$\left|\frac{\mathbf{V}_o}{\mathbf{V}_i}\right|_{\text{dB}} \approx 20 \log_{10} K - 20 \log_{10} \frac{\omega}{\omega_0} \tag{5.65}$$
$$\approx 20 \log_{10} K - 20 \log_{10} \omega + 20 \log_{10} \omega_0$$

Thus, at very high frequencies ($\omega \gg \omega_0$), equation 5.63 is well approximated by a straight line of −20 dB per **decade** slope that intersects the frequency axis at $\omega = \omega_0$. This line is the *high-frequency asymptote* of the Bode magnitude plot. A decade represents a factor of 10 change in frequency. Thus, a one decade increase in ω is equivalent to a unity change in log ω.

Finally, when $\omega = \omega_0$, the real and imaginary parts of the simple pole are equal, such that $|1 + j\omega/\omega_0| = |1 + j| = \sqrt{2}$. Then equation 5.63 becomes:

$$20 \log_{10} \frac{|K|}{|1 + j\omega/\omega_0|} = 20 \log_{10} K - 20 \log_2 1/2 = 20 \log_{10} K - 3 \text{ dB} \tag{5.66}$$

Page 368

Thus, the Bode magnitude plot of a first-order low-pass filter is approximated by two straight lines intersecting at ω_0. Figure 5.42(a) clearly shows the approximation. The actual Bode magnitude plot is 3 dB lower than the approximate plot at $\omega = \omega_0$, the cutoff frequency.

The phase angle of the frequency response function $\angle(\mathbf{V}_o/\mathbf{V}_i) = -\tan^{-1}(\omega/\omega_0)$ has the following properties:

$$-\tan^{-1}\left(\frac{\omega}{\omega_0}\right) = \begin{cases} 0 & \text{when } \omega \to 0 \\ -\frac{\pi}{4} & \text{when } \omega = \omega_0 \\ -\frac{\pi}{2} & \text{when } \omega \to \infty \end{cases}$$

As a first approximation, the phase angle can be represented by three straight lines:

1. For $\omega < 0.1\omega_0$, $\angle(\mathbf{V}_o/\mathbf{V}_i) \approx 0$.
2. For $0.1\omega_0$ and $10\omega_0$, $\angle(\mathbf{V}o/\mathbf{V}_i) \approx -(\pi/4)\ \log(10\omega/\omega_0)$.
3. For $\omega > 10\omega_0$, $\angle\ (\mathbf{V}_o/\mathbf{V}_i) \approx -\pi/2$.

These straight line approximations are illustrated in Figure 5.42(b).

Table 5.2 lists the differences between the actual and approximate Bode magnitude and phase plots. Note that the maximum difference in magnitude is 3 dB at the cutoff frequency; thus, the cutoff is often called the **3-dB frequency** or the *half-power frequency.*

Table 5.2 Correction factors for asymptotic approximation of first-order filter

ω/ω_0	Magnitude response error, (dB)	Phase response error (deg)
0.1	0	−5.7
0.5	−1	4.9
1	−3	0
2	−1	−4.9
10	0	+5.7

RC High-Pass Filter Bode Plots

The case of an *RC* high-pass filter (see Example 5.9 and Figure 5.28) is analyzed in the same manner as was done for the *RC* low-pass filter. It can be shown that the frequency response function contains a constant $K = RC$, a zero "at the origin" and a simple pole with a 3-dB break frequency $\omega_0 = 1/RC$, which is the same simple pole found previously for the *RC* low-pass filter.

$$\frac{\mathbf{V}_o}{\mathbf{V}_o} = \frac{j\omega CR}{1 + j\omega CR} = \frac{j(\omega/\omega_0)}{1 + j(\omega/\omega_0)} \tag{5.67}$$

$$= \frac{(\omega/\omega_0)\angle(\pi/2)}{\sqrt{1 + (\omega/\omega_0)^2}\angle\arctan(\omega/\omega_0)}$$

$$= \frac{\omega/\omega_0}{\sqrt{1 + (\omega/\omega_0)^2}}\angle\left(\frac{\pi}{2} - \arctan\frac{\omega}{\omega_0}\right)$$

Page 369

Figure 5.43 depicts the Bode plots for equation 5.67, where the horizontal axis indicates the normalized frequency ω/ω_0. Straight line asymptotic approximations may again be determined easily at low and high frequencies. For $\omega \ll \omega_0$, the Bode magnitude approximation intercepts the origin ($\omega/\omega_0 = 1$) with a slope of +20 dB/decade. For $\omega \gg \omega_0$, the Bode magnitude approximation is 0 dB with zero slope. The straight line approximations of the Bode phase plot are

1. For $\omega < 0.1\omega_0$, $\angle(\mathbf{V}_o/\mathbf{V}_i) \approx \pi/2$.
2. For $0.1\omega_0 < \omega < 10\omega_0$, $\angle(\mathbf{V}_o/\mathbf{V}_i) \approx \pi/4 - (\pi/4)\log_{10}(\omega/\omega_0)$

3. For $\omega > 10\omega_0$, $\angle(\mathbf{V}_o/\mathbf{V}_i) \approx 0$.

These straight line approximations are illustrated in Figure 5.43(b).

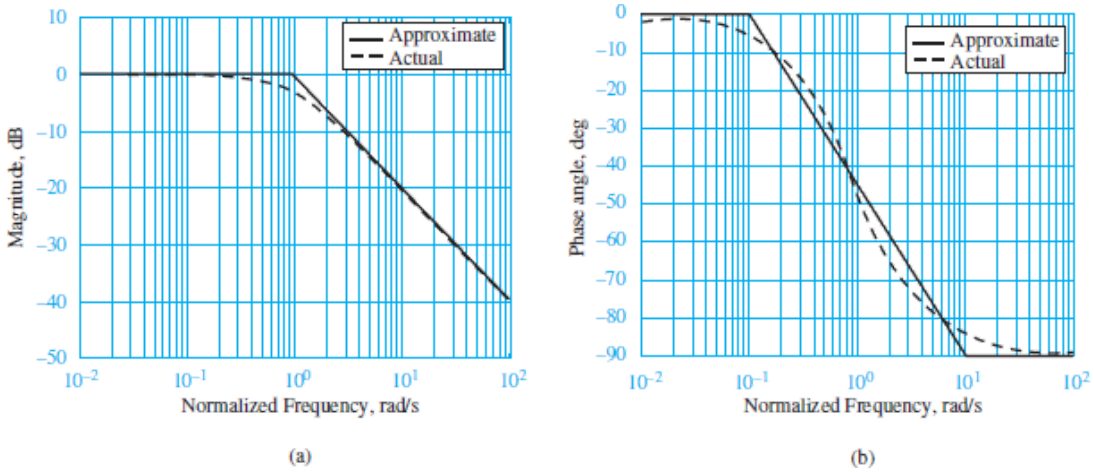


Figure 5.42 Bode plots for *RC* high-pass filter; (a) Magnitude response; (b) phase response

Bode Plots of Higher-Order Filters

Bode plots of high-order systems may be obtained by summing Bode plots of factors of the higher-order frequency response function. Assume, for example,

$$\mathbf{H}(j\omega) = \mathbf{H}_1(j\omega)\mathbf{H}_2(j\omega)\mathbf{H}_3(j\omega) \tag{5.68}$$

which can be expressed, in logarithmic form, as

$$|\mathbf{H}(j\omega)|_{\text{dB}} = |\mathbf{H}_1(j\omega)|_{\text{dB}} + |\mathbf{H}_2(j\omega)|_{\text{dB}} + |\mathbf{H}_3(j\omega)|_{\text{dB}} \tag{5.69}$$

and

$$\angle\mathbf{H}(j\omega) = \angle\mathbf{H}_1(j\omega) + \angle\mathbf{H}_2(j\omega) + \angle\mathbf{H}_3(j\omega) \tag{5.70}$$

Consider as an example the frequency response function

$$\mathbf{H}(j\omega) = \frac{j\omega + 5}{(j\omega + 10)(j\omega + 100)} \tag{5.71}$$

Page 370

The first step in computing the asymptotic approximation consists of factoring each term in the expression so that it appears in the form $a_i(j\omega/\omega_i + 1)$, where the

frequency ω_i corresponds to the appropriate 3-dB frequency, ω_1, ω_2, or ω_3. For example, the function of equation 5.71 is rewritten as:

$$\mathbf{H}(j\omega) = \frac{5(j\omega/5+1)}{10(j\omega/10+1)100(j\omega/100+1)} \tag{5.72}$$

$$= \frac{0.005(j\omega/5+1)}{(j\omega/10+1)(j\omega/100+1)} = \frac{K(j\omega/\omega_1+1)}{(j\omega/\omega_2+1)(j\omega/\omega_3+1)}$$

Equation 5.72 contains a constant K, one simple zero and two simple poles, which can be expressed in logarithmic form:

$$|\mathbf{H}(j\omega)|_{\text{dB}} = |0.005|_{\text{dB}} + \left|\frac{j\omega}{5}+1\right|_{\text{dB}} - \left|\frac{j\omega}{10}+1\right|_{\text{dB}} - \left|\frac{j\omega}{100}+1\right|_{\text{dB}} \tag{5.73}$$

$$\angle\mathbf{H}(j\omega) = \angle 0.005 + \angle\left(\frac{j\omega}{5}+1\right) - \angle\left(\frac{j\omega}{10}+1\right) - \angle\left(\frac{j\omega}{100}+1\right)$$

Each of the terms in the logarithmic magnitude expression can be plotted individually. The constant corresponds to the value −46 dB, plotted in Figure 5.44(a) as a line of zero slope. The simple zero, with a 3-dB frequency $\omega_1 = 5$, departs the zero axis at $\omega_1 = 5$ with a slope of +20 dB/decade; the high-frequency portions of the two simple poles in the denominator are lines of slope −20 dB/decade, departing the zero axis at $\omega_2 = 10$ and $\omega_3 = 100$. With practice, these individual factors are easy to plot by *inspection* once the frequency response function is expressed in standard form, such as shown in equation 5.72.

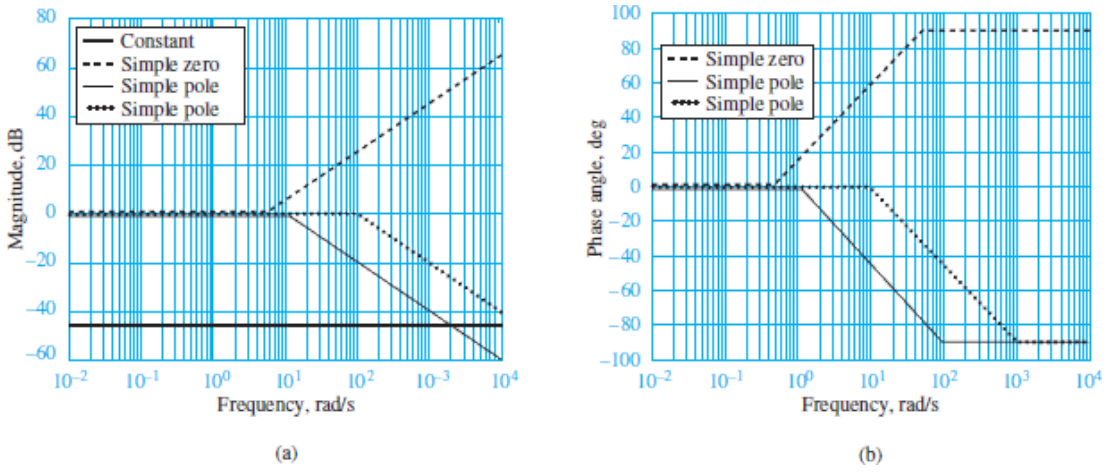


Figure 5.44 Typical Bode plot approximation for a second-order frequency response function; (a) straight line approximation of magnitude response; (b) straight line approximation of phase angle response

Now consider the phase response portion of equation 5.73. First, recognize that the phase angle of the constant is always zero. The phase angle of the simple Page 371zero is approximated as zero for $\omega < 0.1\omega_1$, as a straight line of slope $+\pi/4$ rad/decade for $0.1\omega_1 < \omega < 10\omega_1$, and as $\pi/2$ radians for $\omega > 10\omega_1$. The two simple poles have similar approximations, except that the slopes are $-\pi/4$ radians/decade for $0.1\omega_2 < \omega < 10\omega_2$ and $0.1\omega_3 < \omega < 10\omega_3$ and their high-frequency asymptotes are $-\pi/2$ for $\omega > 10\omega_2$ and $\omega > 10\omega_3$, respectively.

Figure 5.44 depicts the asymptotic approximations of the individual factors in equation 5.73, with the magnitude response shown in Figure 5.44(a) and the phase response shown in Figure 5.44(b). When all the asymptotic approximations are summed, the complete frequency response approximation is obtained. Figure 5.45 depicts the results of the asymptotic Bode approximation when compared with the actual frequency response functions.

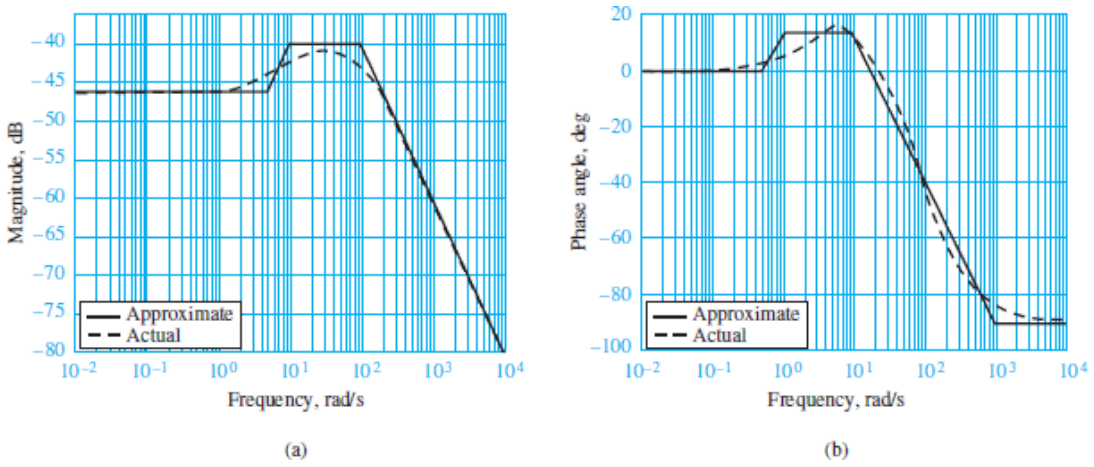


Figure 5.45 Comparison of Bode plot approximation with the actual frequency response function; (a) magnitude response of second-order frequency response function; (b) phase angle response of second-order frequency response function.

FOCUS ON PROBLEM SOLVING

BODE PLOTS

This box illustrates the procedure for constructing the straight line asymp approximations of a Bode plot, such as those shown in Figure 5.45. A zero numerator term; a pole is a denominator term. The method assumes that

frequency response function is in **standard form**, comprised of one or more of three distinct **standard terms** shown below.

1. K Constant
2. $(j\omega)$ Zero/pole "at the origin"
3. $(1 + j\omega\tau) = (1 + j\omega/\omega_0)$ Simple zero/pole

Page 372

First, construct the straight line asymptotic approximations of the Bode magni and phase plots at the low-frequency end. These straight lines are determined ent by the constant K and any zero/pole "at the origin." Then, as frequency increa adjust the *slopes* of these lines to account for the presence of simple zeros/p Refer to Figure 5.44 for examples of how the various terms listed above contribu the overall Bode magnitude and phase plots. Examples 5.11 and 5.12 illustrate details of the method.

1. **Constant** Contributes $20 \log(K)$ dB and $\angle K = 0^\circ$ at all frequencies.
2. **Zero/pole "at the origin"**

- **Zero "at the origin":** Contributes a slope of 20 dB/decade to the E magnitude plot and 90° to the Bode phase plot at all frequencies. The frequency approximations of the magnitude and phase of the entire freque response function are

$$20 \log |\mathbf{H}| \approx 20 \log K + 20 \log(\omega) \text{ dB} \qquad \angle \mathbf{H} \approx 90^\circ$$

  This straight line asymptotic approximation of the magnitude plot has a slop 20 dB/decade and intersects the frequency axis at $\omega = 1/K$. See the frequency portions of Figure 5.43 (a) and (b) for typical examples.

- **Pole "at the origin":** Contributes a slope of −20 dB/decade to the E magnitude plot and -90° to the Bode phase plot at all frequencies. The frequency approximations of the magnitude and phase of the frequency respo function are

$$20 \log |\mathbf{H}| \approx 20 \log K - 20 \log(\omega) \text{ dB} \qquad \angle \mathbf{H} \approx -90^\circ$$

  This straight line asymptotic approximation of the magnitude plot has a slop −20 dB/decade and intersects the frequency axis at $\omega = K$.

3. **Simple zero**

- **Magnitude plot:** At its cutoff frequency ω_0, a simple zero, produces a dB/decade *change in slope.*

- **Phase plot:** One decade *below* its cutoff frequency a simple zero produc 45°/decade *change in slope.* One decade *above* its cutoff frequency a sir zero produces a -45°/decade *change in slope.* The overall effect is that the p plot *increases* by 90° over the span of two decades centered about ω_0.

4. **Simple pole**

- **Magnitude plot:** At its cutoff frequency ω_0, a simple pole, produces a dB/decade *change in slope.* See the changes in slope in Figures 5.42 (a) 5.43 (a) for typical examples.Page 373

- **Phase plot:** One decade *below* its cutoff frequency a simple pole produc -45°/decade *change in slope.* One decade *above* its cutoff frequency a sir pole produces a 45°/decade *change in slope.* The overall effect is that the p plot *decreases* by 90° over the span of two decades centered about ω_0. See changes in slope in Figures 5.42 (b) and 5.43 (b) for typical examples.

Comments:

1. Simple zeros/poles contribute 0 dB, no slope, and 0° at the low- frequency en
2. A zero slope at the low-frequency end indicates no zero/pole "at the origin."
3. A frequency response function can contain zeros/poles raised to a power, suc $(1 + j\omega/\omega_0)^2$ or$(j\omega)^3$. Since $x^2 = x \cdot x$ the exponent indicates the numbe repetitions of the zero/pole.
4. Since 20 log $|(j\omega)^3| = 20 \log |(j\omega)|^3 = 60 \log(\omega)$ the exponent multip the *slope* of a zero/pole "at the origin" in a Bode magnitude plot. Likewise exponent multiplies the phase angle of a zero/pole "at the origin" in a B phase plot.
5. Since 20 log $|(1 + j\omega/\omega_0)^2| = 20 \log |(1 + j\omega/\omega_0)|^2 = 40 \log |(1 + \omega_0)|$ the exponent multiplies the *change of slope* at the cutoff frequency simple zero/pole in a Bode magnitude plot. Likewise, the exponent multip the *change of slope* one decade below and above the cutoff frequency simple zero/pole in a Bode phase plot.
6. The correction factors shown in Table 5.2 can be used to improve the p angle approximation plot.

EXAMPLE 5.11 Bode Plot Approximation

Problem

Sketch the asymptotic approximation of the Bode plot for the frequency response function

$$\mathbf{H}(j\omega) = \frac{0.1j\omega + 20}{2 \times 10^{-5}(j\omega)^3 + 0.1002(j\omega)^2 + j\omega}$$

Solution

Known Quantities: Frequency response function of a circuit.

Find: Bode plot approximation of given frequency response function.

Page 374

Assumptions: None

Analysis: Following the Focus on Problem Solving box "Bode Plots," factor the function into the standard form. Notice that the quadratic pole can be factored into two simple poles. It is important to realize that not all quadratic terms can be so factored, but in this example it is possible. The result is

$$\mathbf{H}(j\omega) = \frac{20(j\omega/200 + 1)}{j\omega(j\omega/10 + 1)(j\omega/5{,}000 + 1)}$$

Notice the constant 20 and the $j\omega$ term in the denominator, which is a pole "at the origin." The low end of the frequency response is determined by these two terms. At low frequencies:

$$20 \log|\mathbf{H}| \approx 20 \log(20) - 20 \log(\omega)$$
$$\angle\mathbf{H} \approx 0° - \angle(j\omega) = -90°$$

That is, the magnitude of the denominator factor $j\omega$ is represented by a line with slope of −20 dB/decade intersecting the frequency (horizontal) axis at $\omega = 1$. Its phase response is a constant equal to $-\pi/2$ radians or −90 degrees.

The asymptotic straight line approximations of the magnitude and phase responses of the constant, the pole "at the origin," the simple zero, and the two simple poles are shown in Figure 5.46. In Figure 5.46(a), notice the changes in slope at the cutoff frequencies of the simple zero (ω = 200 rad/sec) and the two simple poles (ω = 10 rad/sec and ω = 5000 rad/sec). Also notice in Figure 5.46(b) the changes in slope one decade below and one decade above the same cutoff frequencies. These straight line approximations are summed to produce the approximate overall frequency response plots, which are shown in Figure 5.47.

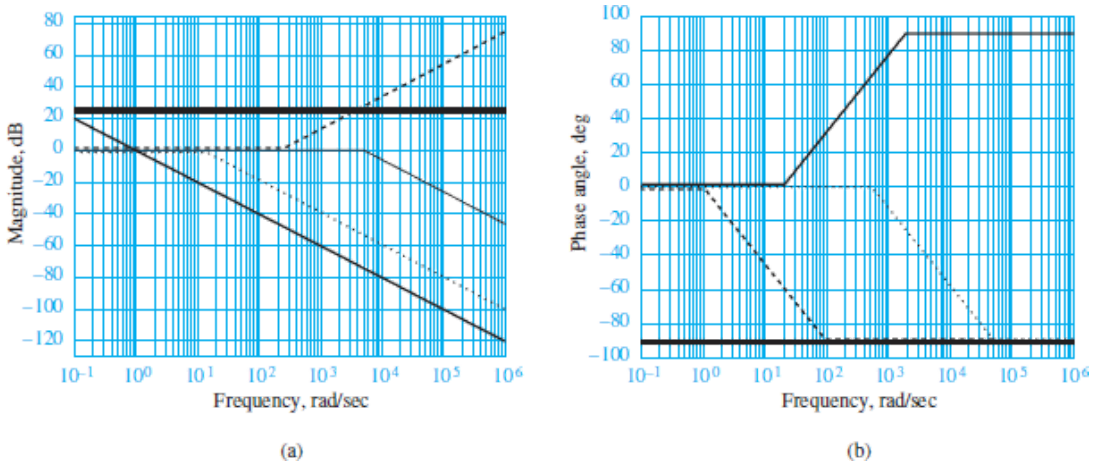


Figure 5.46 Approximate (asymptotic) frequency responses of individual first-order terms; (a) straight line approximation of magnitude response; (b) straight line approximation of phase angle response.

Comments: A computer program such as MatLab can be used to generate the Bode plot approximation shown in Figures 5.46 and 5.47. However, the process of creating approximate Bode plots is a necessary practice when learning to understand and interpret actual Bode plots.

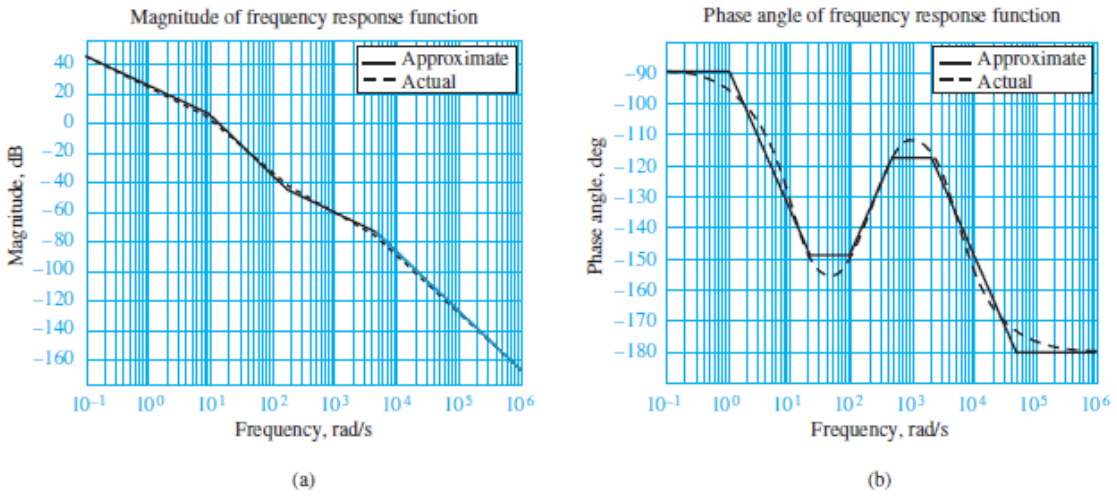


Figure 5.47 Comparison of approximate and exact frequency response; (a) magnitude of frequency response function; (b) phase angle of frequency response function.

EXAMPLE 5.12 Bode Plot Approximation

Problem

Sketch the asymptotic approximation of the Bode plot for the frequency response function

$$\mathbf{H}(j\omega) = \frac{10^{-3}(j\omega)^2 + 0.1j\omega}{[1/(9\times 10^4)](j\omega)^2 + (3{,}030/90{,}000)j\omega + 1}$$

Solution

Known Quantities: Frequency response function of a circuit.

Find: Bode plot approximation of given frequency response function.

Assumptions: None

Analysis: Following the Focus on Problem Solving box "Bode Plots," factor the function into standard form. Notice that the quadratic pole can be factored into two

simple poles. It is important to realize that not all quadratic terms can be so factored, but in this example it is possible. The result is

$$\mathbf{H}(j\omega) = \frac{0.1j\omega(j\omega/100 + 1)}{(j\omega/30 + 1)(j\omega/3{,}000 + 1)}$$

Notice the constant 0.1 and the $j\omega$ term in the numerator, which is a zero "at the origin." The low end of the frequency response is determined by these two terms. At low frequencies:

$$20 \log|\mathbf{H}| \approx 20 \log(0.1) + 20 \log(\omega)$$

$$\angle\mathbf{H} \approx 0° + \angle(j\omega) = 90°$$

Page 376

That is, the magnitude of $j\omega$ is represented by a line with slope +20 dB/decade intersecting the frequency (horizontal) axis at $\omega = 1$. The phase of the factor $j\omega$ is a constant and equal to $\pi/2$ radians or 90 degrees.

The asymptotic straight line approximations of the magnitude and phase responses of the constant, the zero "at the origin," the simple zero, and the two simple poles are shown in Figure 5.48. In Figure 5.48(a), notice the changes in slope at the cutoff frequencies of the simple zero (ω = 100 rad/sec) and the two simple poles (ω = 30 rad/sec and ω = 3000 rad/sec). Also notice in Figure 5.48(b) the changes in slope one decade below and one decade above the same cutoff frequencies. These straight line approximations are summed to produce the approximate overall frequency response plots, which are shown in Figure 5.49.

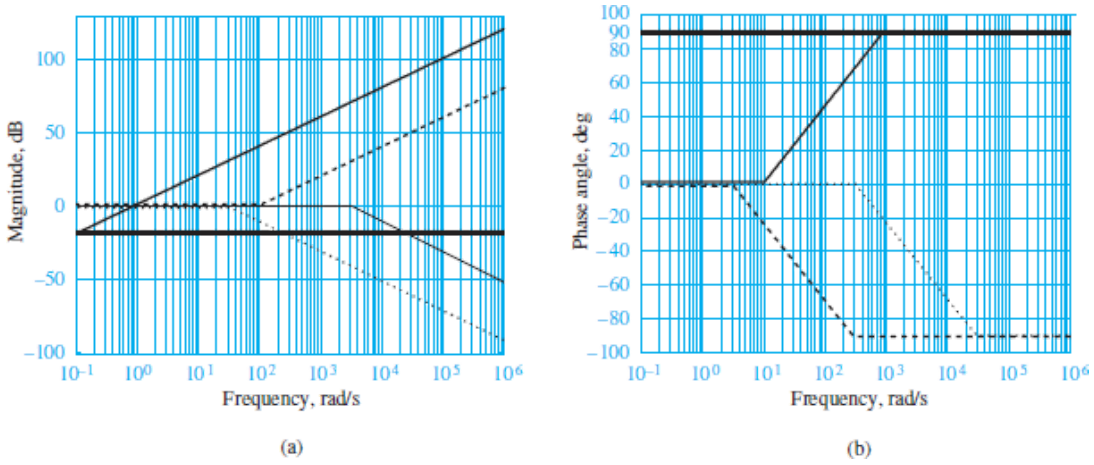


Figure 5.48 Approximate (asymptotic) frequency responses of individual first-order terms; (a) straight line approximation of magnitude response; (b) straight line approximation of phase angle response.

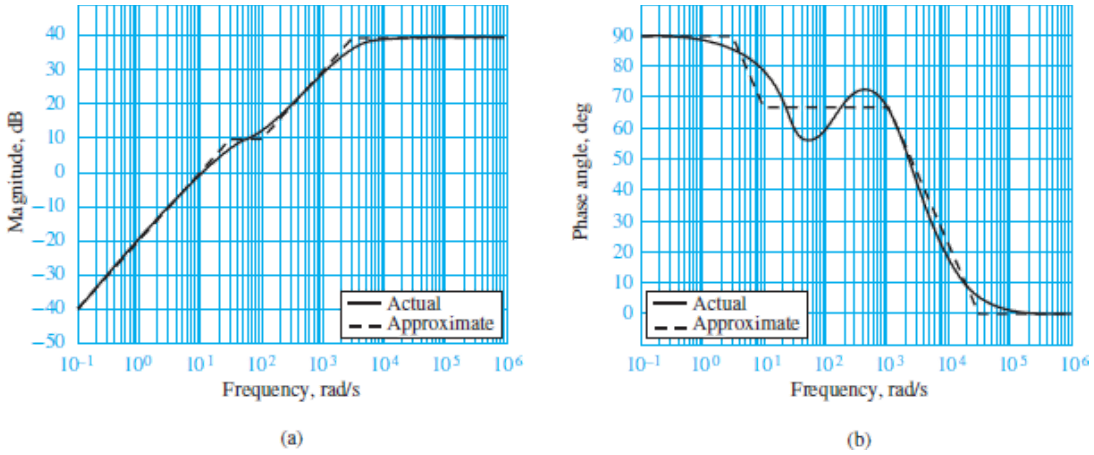


Figure 5.49 Comparison of approximate and exact frequency responses; (a) magnitude of frequency response function; (b) phase angle of frequency response function.

Page 377

Comments: Bode plots can be generated using MatLab. Circuit simulation programs, such as TINA®, can also generate Bode plots.

FOCUS ON PROBLEM SOLVING

QUADRATIC ZEROS/POLES

This box illustrates how quadratic zeros/poles impact the straight line asymp approximations of Bode magnitude and phase plots. A quadratic zero/ contributes to the overall frequency response at frequencies near and abov natural frequency ω_n. When possible, factor a quadratic zero/pole into two si zeros/poles and refer to the previous Focus on Problem Solving box.

$$1 + (2\zeta/\omega_n)j\omega + (j\omega/\omega_n)^2 = 1 + (1/Q\omega_n)j\omega + (j\omega/\omega_n)^2 \qquad \text{Quadratic zero/pole}$$

1. **Quadratic zero**

- **Magnitude plot:** A quadratic zero produces a 40 dB/decade *change in slop* frequency increases across the natural frequency ω_n.

- **Phase plot:** A quadratic zero produces a 90° increase in phase angl frequency approaches the natural frequency ω_n and another 90° increas phase angle as frequency increases beyond the natural frequency. In tot quadratic zero produces a 180° increase in phase angle as frequency incre across the natural frequency.

- **When $Q \gg 1$:** Contributes approximately −20 log (Q) dB at $\omega = \omega_n$ such the magnitude plot traverses a *local minimum.* The bandwidth of this l minimum decreases as Q increases.

2. **Quadratic pole**

- **Magnitude plot:** A quadratic pole produces a −40 dB/decade *change in slop* frequency increases across the natural frequency ω_n.

- **Phase plot:** A quadratic pole produces a 90° decrease in phase angl frequency approaches the natural frequency ω_n and another 90° decreas phase angle as frequency increases beyond the natural frequency. In tot quadratic pole produces a 180° decrease in phase angle as frequency incre across the natural frequency.

- **When $Q \gg 1$:** Contributes approximately 20 log (Q) dB at $\omega = \omega_n$ such tha magnitude plot traverses a *local maximum.* This effect is known as a **resona** The bandwidth of this local maximum decreases as Q increases. See Fig 5.34 and 5.50.

Page 378

Comment: The impact of a quadratic pole at its natural frequency dep significantly upon the quality factor $Q = 1/(2\zeta)$ as suggested by Figures 5.34 5.50.

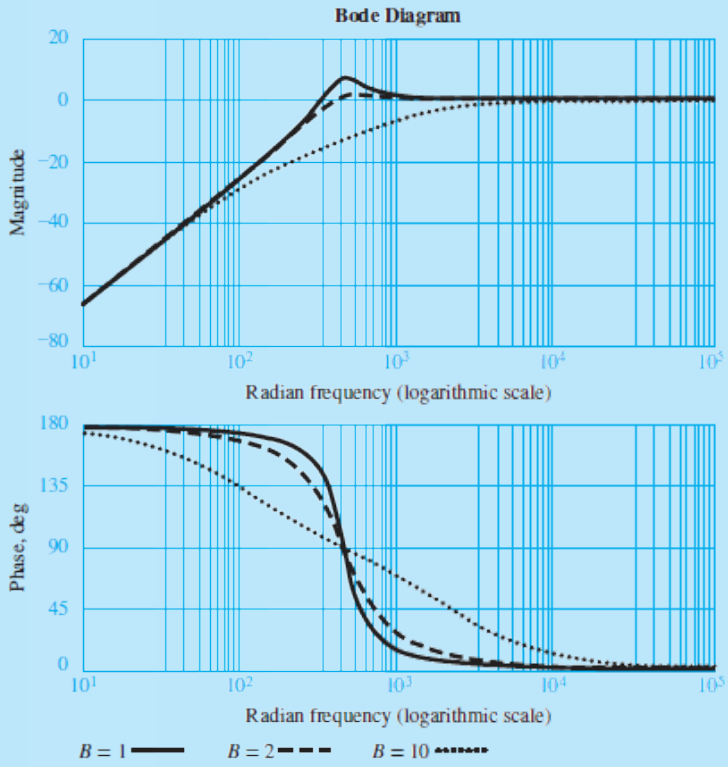
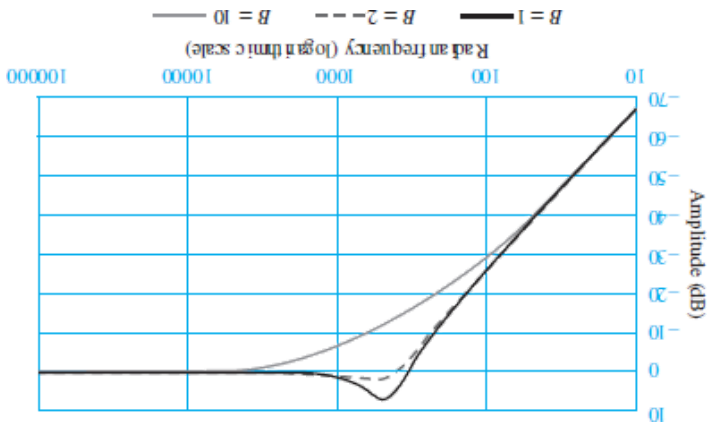


Figure 5.50 Quadratic pole with $Q \approx 0.22, 1.1, 2.2$.

CHECK YOUR UNDERSTANDING

Use the information in the Focus on Measurements box, "Seismic Transducer," to create the Bode magnitude plot of the frequency response of the seismic displacement transducer. Compare the Bode magnitude plot to the linear magnitude plot shown in Figure 5.40. Is the slope at the low-frequency end as expected? Is the change in slope across the natural frequency as expected? Is the resonant peak at the natural frequency for $B = 1$ as expected? The resonant peak estimate for $Q \gg 1$ is 20 log (Q).



Conclusion

Chapter 5 focuses on the frequency response of linear circuits, and it is a natural extension of the material covered in Chapter 3. The concepts of the spectrum of a signal, obtained through the Fourier series representation for periodic signals, and of the frequency response of a filter are very useful ideas that extend well beyond electrical engineering. For example, civil, mechanical, and aeronautical engineering students who study the vibrations of structures and machinery will find that the same methods are employed in those fields.

Upon completing this chapter, you should have mastered the following learning objectives:

1. *Understand the physical significance of frequency domain analysis, and compute the frequency response of circuits by using AC circuit analysis tools.* You had already acquired the necessary tools (phasor analysis and impedance) to compute the frequency response of circuits in Chapter 3; in the material presented in section 5.1, these tools are put to use to determine the frequency response functions of linear circuits.
2. *Compute the Fourier spectrum of periodic signals by using the Fourier series representation, and use this representation in connection with frequency response ideas to compute the response of circuits to periodic inputs.* The concept of spectrum is very important in many engineering applications; in section 5.2 you learned to compute the Fourier spectrum of an important class of functions: those that repeat periodically. The frequency spectrum of signals makes frequency domain analysis (i.e., computing the response of circuits using the phasor domain representation of signals) very easy, even for relatively complex signals because it allows you to decompose the signals into a summation of sinusoidal components, which can then be easily handled one at a time.

3. *Analyze simple first- and second-order electrical filters, and determine their frequency response and filtering properties.* With the concept of *frequency response* firmly in hand, now you can analyze the behavior of electrical filters and study the frequency response characteristics of the most common types, that is, low-pass, high-pass, and bandpass filters. Filters are very useful devices and are explored in greater depth in Chapters 6 and 7.
4. *Compute the frequency response of a circuit and its graphical representation in the form of a Bode plot.* Graphical approximations of Bode plots can be very useful to develop a quick understanding of the frequency response characteristics of a linear system, almost by inspection. Bode plots find use in the discipline of automatic control systems, a subject that is likely to be encountered by most engineering majors.

Page 380

HOMEWORK PROBLEMS

Section 5.2: Fourier Analysis

5.1 Use trigonometric identities to show that the equalities in equations 5.27 and 5.28 hold.

5.2 Derive a general expression for the Fourier series coefficients of the square wave of Figure 5.15(a) in the text.

5.3 Compute the Fourier series coefficient of the periodic function shown in Figure P5.3 and defined as:

$$x(t) = \begin{cases} 0 & 0 \le t \le \frac{T}{3} \\ A & \frac{T}{3} \le t \le T \end{cases}$$

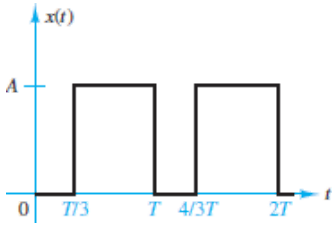


Figure P5.3

5.4 Compute the Fourier series coefficient of the periodic function shown in Figure P5.4 and defined as

$$x(t) = \begin{cases} \cos\left(\frac{2\pi}{T}t\right) & -\frac{T}{4} \le t < \frac{T}{4} \\ 0 & \text{else} \end{cases}$$

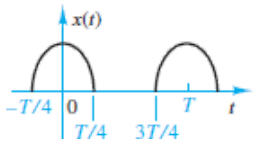


Figure P5.4

5.5 Compute the Fourier series expansion of the function shown in Figure P5.5, and express it in sine-cosine (a_n, b_n coefficients) form.

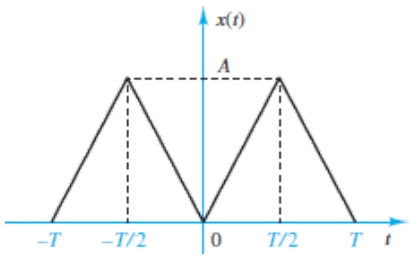


Figure P5.5

5.6 Compute the Fourier series expansion of the function shown in Figure P5.6, and express it in sine-cosine (a_n, b_n coefficients) form.

$$x(t) = \begin{cases} \sin\left(\frac{2\pi}{T}t\right) & 0 \le t < \frac{T}{2} \\ 0 & \frac{T}{2} \le t < T \end{cases}$$

Figure P5.6

5.7 Write an expression for the signal shown in Figure P5.7, and derive a complete expression for its Fourier series.

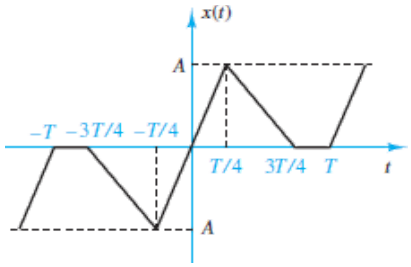


Figure P5.7

5.8 Write an expression for the signal shown in Figure P5.8 and derive its Fourier series.

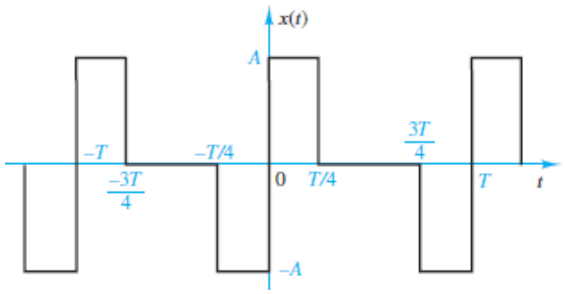


Figure P5.8

5.9 Find the Fourier series for the periodic function shown in Figure P5.9. Determine integral expressions for the Fourier coefficients.

Page 381

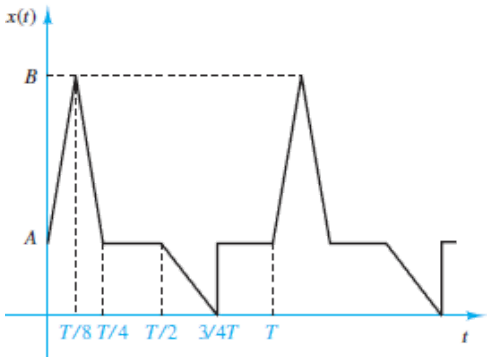


Figure P5.9

5.10 Find the Fourier series for the periodic function shown in Figure P5.10. Determine integral expressions for the Fourier coefficients.

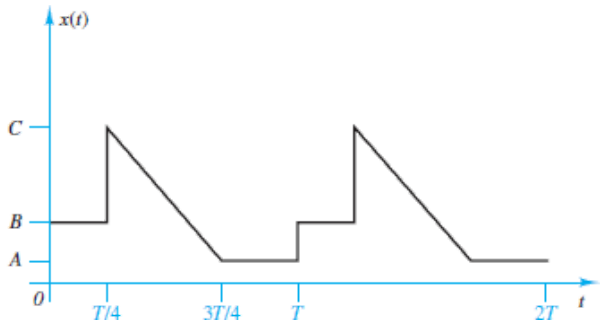


Figure P5.10

Section 5.3: Low- and High-Pass Filters

5.11 a. Determine the frequency response $V_{out}(j\omega)/V_{in}(j\omega)$ for the circuit of Figure P5.11. Assume $L = 0.5$ H and $R = 200$ kΩ.

b. Plot the magnitude and phase of the circuit for frequencies between 10 and 10^7 rad/s on graph paper, with a linear scale for frequency.

c. Repeat part b, using semilog paper. (Place the frequency on the logarithmic axis.)

d. Plot the magnitude response on semilog paper with magnitude in decibels.

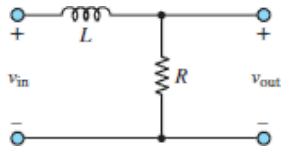


Figure P5.11

5.12 Repeat the instructions of Problem 5.11 for the circuit of Figure P5.12.

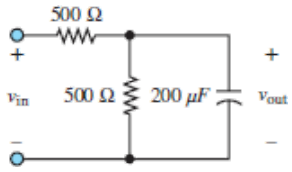


Figure P5.12

5.13 Repeat the instructions of Problem 5.11 for the circuit of Figure 5.13.

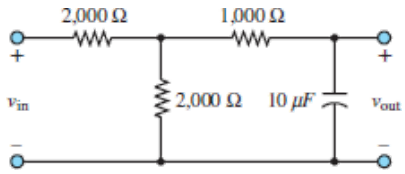


Figure P5.13

5.14 In the circuit shown in Figure P5.14, where $C = 0.5\ \mu$F and $R = 2$ kΩ,

a. Determine how the input impedance $\mathbf{Z}(j\omega) = \mathbf{V}_i(j\omega)/\mathbf{I}_i(j\omega)$ behaves at extremely high and low frequencies.

b. Find an expression for that impedance.

c. Show that this expression can be manipulated into the form $\mathbf{Z}(j\omega) = R[1 - j(1/\omega RC)]$.

d. Determine the frequency $\omega = \omega_C$ for which the imaginary part of the expression in part c is equal to 1.

e. Estimate (without computing it) the magnitude and phase angle of $\mathbf{Z}(j\ \omega)$ at $\omega = 10$ rad/s and $\omega = 10^5$ rad/s.

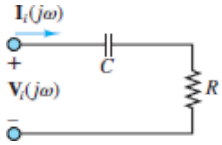


Figure P5.14

5.15 In the circuit shown in Figure P5.15, where $L = 2$ mH and $R = 2$ kΩ,

a. Determine how the input impedance $\mathbf{Z}(j\omega) = \mathbf{V}_i(j\omega)/\mathbf{I}_i(j\omega)$ behaves at extremely high and low frequencies.

Page 382

b. Find an expression for the impedance.

c. Show that this expression can be manipulated into the form $\mathbf{Z}(j\omega) = R[1 + j\ (\omega L/R)]$.

d. Determine the frequency $\omega = \omega_C$ for which the imaginary part of the expression in part c is equal to 1.

e. Estimate (without computing it) the magnitude and phase angle of $\mathbf{Z}(j\omega)$ at $\omega = 10^5$ rad/s, 10^6 rad/s, and 10^7 rad/s.

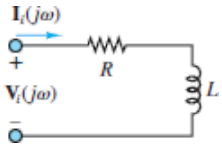


Figure P5.15

5.16 In the circuit of Figure P5.16:

$$R_1 = 1.3\ \text{k}\Omega \qquad R_2 = 1.9\ \text{k}\Omega$$
$$C = 0.5182\ \mu\text{F}$$

Determine:

a. How the voltage frequency response function

$$\mathbf{H}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

behaves at extremes of high and low frequencies.

b. An expression for the voltage frequency response function and show that it can be manipulated into the form

$$\mathbf{H}_v(j\omega) = \frac{H_o}{1 + jf(\omega)}$$

where

$$H_o = \frac{R_2}{R_1 + R_2} \qquad f(\omega) = \omega R_T C \qquad R_T = \frac{R_1 R_2}{R_1 + R_2}$$

c. The frequency at which $f(\omega) = 1$ and the value of H_o in decibels.

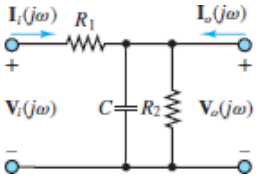


Figure P5.16

5.17 In the circuit shown in Figure P5.17, determine the frequency response function in the form:

$$\mathbf{H}_v(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)} = \frac{H_o}{1 \pm jf(\omega)}$$

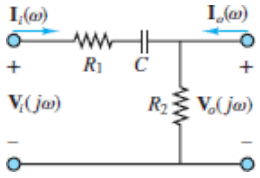


Figure P5.17

5.18 The circuit shown in Figure P5.18 has

$$R_1 = 100\ \Omega \qquad R_o = 100\ \Omega$$
$$R_2 = 50\ \Omega \qquad C = 80\ \text{nF}$$

Determine the frequency response $\mathbf{V}_o(j\omega)/\mathbf{V}_{\text{in}}(j\omega)$.

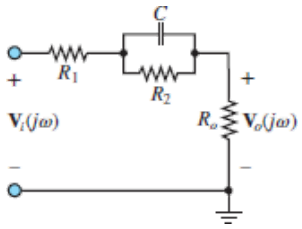


Figure P5.18

5.19 a. Determine the frequency response $\mathbf{V}_{\text{out}}(j\omega)/\mathbf{V}_{\text{in}}(j\omega)$ for the circuit of Figure P5.19.

b. Plot the magnitude and phase of the circuit for frequencies between 1 and 100 rad/s on graph paper, with a linear scale for frequency.

c. Repeat part b, using semilog paper. (Place the frequency on the logarithmic axis.)

d. Plot the magnitude response on semilog paper with magnitude in dB.

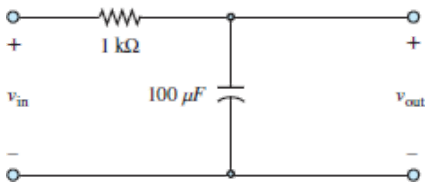


Figure P5.19

5.20 Consider the circuit shown in Figure P5.20. Use the values for R and L given in Problem 5.15.

a. Sketch the amplitude response of $\boldsymbol{Y} = \mathbf{I}/\boldsymbol{V}_S$.

b. Sketch the amplitude response of $\boldsymbol{V}_1/\boldsymbol{V}_S$.

c. Sketch the amplitude response of $\boldsymbol{V}_2/\boldsymbol{V}_S$.

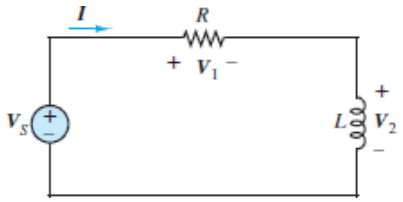


Figure P5.20

5.21 Using a 15-kΩ resistance, design an RC high-pass filter with a breakpoint at 200 kHz.

5.22 Using a 500-Ω resistance, design an RC low-pass filter that would attenuate a 120-Hz sinusoidal voltage by 20 dB with respect to the DC gain.

5.23 At what frequency is the phase shift introduced by the circuit of Example 5.6 equal to −10°?

5.24 At what frequency is the output of the circuit of Example 5.6 attenuated by 10 percent (that is, $V_o = 0.9\ V_i$)?

5.25 Assume the filter shown in Figure 5.11 is excited by the first two Fourier components of the sawtooth waveform in Example 5.3. Determine the output of the filter, and plot the input and output waveforms on the same graph. Assume the period $T = 10\ \mu$s and the peak amplitude $A = 1$ for the sawtooth waveform.

5.26 Repeat Problem 5.25 with the square wave of Figure 5.15(a) as the input.

5.27 Repeat Problem 5.25 for the pulse train of Example 5.4 as the input.

5.28 Assume the circuit shown in Figure P5.12 is excited by the first three Fourier components of the sawtooth waveform in Example 5.3. Determine the output of the filter, and plot the input and output waveforms on the same graph. Assume $T = 0.5$ s and $A = 2$ for the sawtooth waveform.

5.29 Repeat Problem 5.28 with the square wave of Figure 5.15(a) as the input.

5.30 Repeat Problem 5.28 with the pulse train of Example 5.4 as the input.

5.31 Assume the filter shown in Figure P5.13 is excited by the first four Fourier components of the sawtooth waveform in Example 5.3. Determine the output of

the filter, and plot the input and output waveforms on the same graph. Assume $T = 0.1$ s and $A = 1$ for the sawtooth waveform.

5.32 Repeat Problem 5.31 with the square wave of Figure 5.15(a) as the input.

5.33 Repeat Problem 5.31 with the pulse train of Example 5.4 as the input.

Section 5.4: Bandpass Filters, Resonance and Quality Factor

5.34 Repeat Problem 5.11 for the circuit of Figure P5.34. $R_1 = 300\ \Omega$, $R_2 = R_3 = 500\ \Omega$, $L = 4$ H, $C_1 = 40\ \mu$F, $C_2 = 160\ \mu$F.

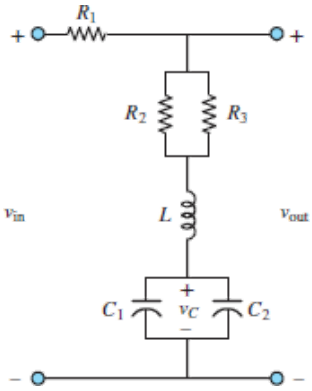


Figure P5.34

5.35 Determine the frequency response of the circuit of Figure P5.35, and generate frequency response plots. $R_1 = 20\ \text{k}\Omega$, $R_2 = 100\ \text{k}\Omega$, $L = 1$ H, $C = 100\ \mu$F.

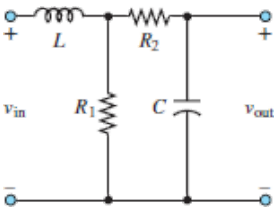


Figure P5.35

5.36 In the circuit shown in Figure P5.36, if

$L = 190$ mH $\quad R_1 = 2.3\ \text{k}\Omega$

$C = 55$ nF $\quad R_2 = 1.1\ \text{k}\Omega$

a. Determine how the input impedance behaves at extremely high or low frequencies.

b. Find an expression for the input impedance in the form

$$\mathbf{Z}(j\omega) = Z_o\left[\frac{1 + jf_1(\omega)}{1 + jf_2(\omega)}\right]$$

$$Z_o = R_1 + \frac{L}{R_2 C}$$

$$f_1(\omega) = \frac{\omega^2 R_1 LC - R_1 - R_2}{\omega(R_1 R_2 C + L)}$$

$$f_2(\omega) = \frac{\omega^2 LC - 1}{\omega C R_2}$$

c. Determine the four frequencies at which $f_1(\omega) = +1$ or -1 and $f_2(\omega) = +1$ or -1.

d. Plot the impedance (magnitude and phase) versus frequency.

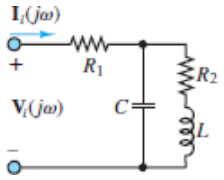


Figure P5.36

5.37 The circuit shown in Figure P5.37 is a second-order circuit because it has two reactive components (L and *C*). A complete solution will not be attempted. However, determine:

a. The behavior of the voltage frequency response at extremely high and low frequencies.

b. The output voltage $\mathbf{V}_o$ if the input voltage has a frequency where:

$$\mathbf{V}_i = 7.07\angle\frac{\pi}{4}\mathbf{V} \qquad R_1 = 2.2\ \text{k}\Omega$$
$$R_2 = 3.8\ \text{k}\Omega \qquad X_c = 5\ \text{k}\Omega \qquad X_L = 1.25\ \text{k}\Omega$$

c. The output voltage if the frequency of the input voltage doubles so that

$$X_C = 2.5\ \text{k}\Omega \qquad X_L = 2.5\ \text{k}\Omega$$

d. The output voltage if the frequency of the input voltage again doubles so that

$$X_C = 1.25\ \text{k}\Omega \qquad X_L = 5\ \text{k}\Omega$$

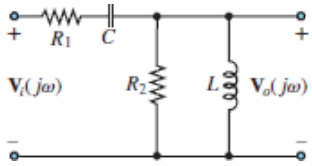


Figure P5.37

5.38 In an *RLC* circuit, assume ω_1 and ω_2 such that $\mathbf{I}(j\omega_1) = \mathbf{I}(j\omega_2) = I_{max}/\sqrt{2}$ and $\Delta\omega$ such that $\Delta\omega = \omega_2 - \omega_1$. In other words, $\Delta\omega$ is the bandwidth of the current curve where the current has fallen to $1/\sqrt{2} = 0.707$ of its maximum value at the resonance frequency. At these frequencies, the power dissipated in a resistance becomes one-half of the dissipated power at the resonance frequency. In an *RLC* circuit with a high-quality factor, show that $Q = \omega_0/\Delta\omega$.

5.39 In an *RLC* circuit with a high quality factor:

a. Show that the impedance at the resonance frequency becomes a value of Q times the inductive resistance at the resonance frequency.

b. Determine the impedance at the resonance frequency, assuming L = 280 mH, C = 0.1 μF, R = 25 Ω.

5.40 At what frequencies is the output of the circuit of Example 5.10 attenuated by 10 percent (that is, $V_o = 0.9\ V_i$)?

5.41 At what frequencies is the phase shift introduced by the circuit of Example 5.10 equal to 20°?

5.42 Assume the filter shown in Figure P5.34 is excited by the first two Fourier components of the sawtooth waveform in Example 5.3. Determine the output of the filter, and plot the input and output waveforms on the same graph. Assume T = 50 ms and A = 2 for the sawtooth waveform.

5.43 Repeat Problem 5.42 for T = 0.5 s and 5 ms, and compare the results with T = 50 ms.

5.44 Repeat Problem 5.42 for the square wave of Figure 5.15(a).

5.45 Repeat Problem 5.42 with the pulse train of Example 5.4 as the input.

5.46 Assume the filter shown in Figure P5.35 is excited by the first three Fourier components of the sawtooth waveform in Example 5.3. Determine the output of the filter, and plot the input and output waveforms on the same graph. Assume T = 5 s and A = 1 for the sawtooth waveform.

5.47 Repeat Problem 5.46 for $T = 50$ s, and compare the results.

5.48 Repeat Problem 5.46 with the square wave of Figure 5.15(a) as the input.

5.49 Repeat Problem 5.46 with the pulse train of Example 5.4 as the input.

5.50 Consider the circuit shown in Figure P5.50. Determine the resonant frequency and the bandwidth for the circuit.

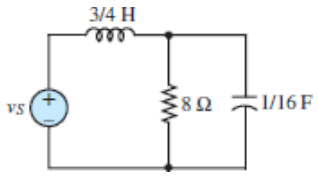


Figure P5.50

5.51 Are the filters shown in Figure P5.51 low-pass, high-pass, bandpass, or bandstop (notch) filters?

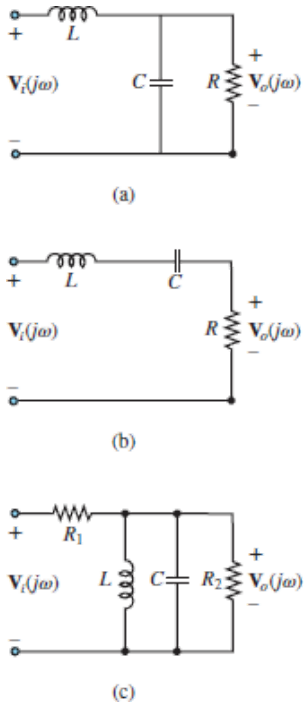


Figure P5.51

5.52 Determine if each of the circuits shown in Figure P5.52 is a low-pass, high-pass, bandpass, or bandstop (notch) filter.

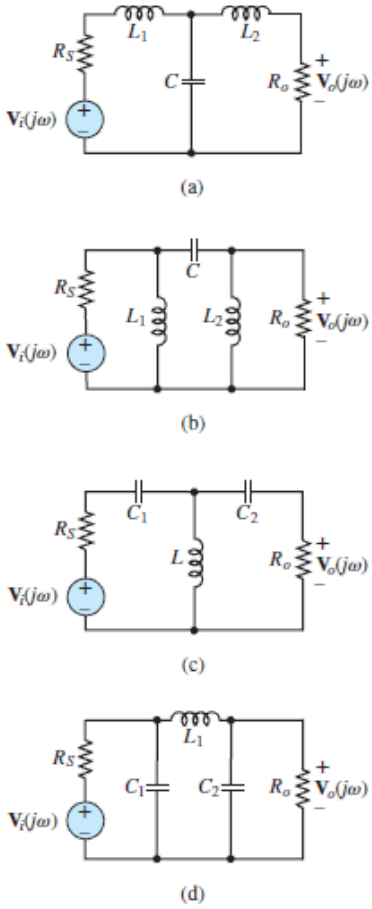


Figure P5.52

5.53 For the filter circuit shown in Figure P5.53:

a. Determine if this is a low-pass, high-pass, bandpass, or bandstop filter.

b. Determine the frequency response $\mathbf{V}_o(j\omega)/\mathbf{V}_i(j\omega)$assuming L = 10 mH, C = 1 nF, R_1 = 50 Ω, R_2 = 2.5 kΩ

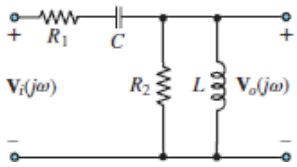


Figure P5.53

Page 386

5.54 In the filter circuit shown in Figure P5.54: L = 10 H, C = 1 nF, R_S = 20 Ω, R_C = 100 Ω, R_o = 5 kΩ Determine the frequency response $\mathbf{V}_o(j\omega)/\mathbf{V}_i(j\omega)$. What type

of filter does this frequency response represent?

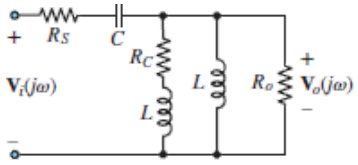


Figure P5.54

5.55 In the filter circuit shown in Figure P5.54: $L = 0.1$ mH, $C = 8$ nF, $R_S = 300\ \Omega$, $R_C = 10\ \Omega$, $R_o = 500\ \Omega$. Determine the frequency response $\mathbf{V}_o(j\omega)/\mathbf{V}_i(j\omega)$. What type of filter does this frequency response represent?

5.56 In the filter circuit shown in Figure P5.56:

$$R_S = 5\ \text{k}\Omega \qquad C = 56\ \text{nF}$$
$$R_o = 100\ \text{k}\Omega \qquad L = 9\ \mu\text{H}$$

Determine:

a. The voltage frequency response

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

b. The resonant frequency.

c. The half-power frequencies.

d. The bandwidth and Q.

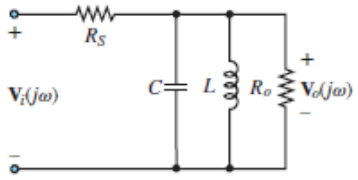


Figure P5.56

5.57 In the filter circuit shown in Figure P5.56:

$$R_S = 5\ \text{k}\Omega \qquad C = 0.5\ \text{nF}$$
$$R_o = 100\ \text{k}\Omega \qquad L = 1\ \text{mH}$$

Determine:

a. The voltage frequency response

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

b. The resonant frequency.

c. The half-power frequencies.

d. The bandwidth and Q.

5.58 In the filter circuit shown in Figure P5.58:

$$R_S = 500\ \Omega \qquad R_o = 5\ \text{k}\Omega$$
$$R_C = 4\ \text{k}\Omega \qquad L = 1\ \text{mH}$$
$$C = 5\ \text{pF}$$

Determine the frequency response $\mathbf{G}_V(j\omega)$, where:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

What type of filter does this frequency response represent?

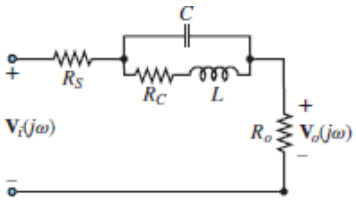


Figure P5.58

5.59 In the notch filter circuit shown in Figure P5.59, derive the voltage frequency response $\mathbf{G}_V(j\omega)$ in standard form, where:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

Assume:

$$R_S = 500\ \Omega \qquad R_o = 5\ \text{k}\Omega$$
$$C = 5\ \text{pF} \qquad L = 1\ \text{mH}$$

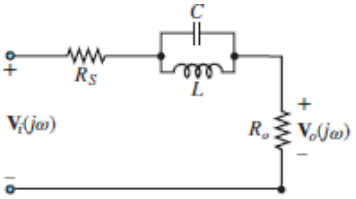


Figure P5.59

5.60 In the notch filter circuit shown in Figure P5.59, derive the voltage frequency response $\mathbf{G}_V(j\omega)$ in standard form, where:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

Page 387

Assume:

$$R_S = 500\ \Omega \qquad R_o = 5\ \text{k}\Omega$$
$$\omega_n = 12.13\ \text{Mrad/s} \qquad C = 68\ \text{nF}$$
$$L = 0.1\ \mu\text{H}$$

Also, determine the half-power frequencies, bandwidth, and Q.

5.61 In the notch filter circuit shown in Figure P5.59, derive the voltage frequency response $\mathbf{G}_V(j\omega)$ in standard form, where:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

Assume:

$$R_S = 4.4\ \text{k}\Omega \quad R_o = 60\ \text{k}\Omega \quad \omega_n = 25\ \text{Mrad/s}$$
$$C = 0.8\ \text{nF} \quad L = 2\ \mu\text{H}$$

Also, determine the half-power frequencies, bandwidth, and Q.

5.62 In the bandstop (notch) filter shown in Figure P5.62:

$$L = 0.4\ \text{mH} \qquad R_c = 100\ \Omega$$
$$C = 1\ \text{pF} \qquad R_S = R_o = 3.8\ \text{k}\Omega$$

Determine:

a. An expression for the voltage frequency response:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)} = G_0 \frac{1 + jf_1(\omega)}{1 + jf_2(\omega)}$$

b. The magnitude of the frequency response at very high and very low frequencies and at the resonant frequency.

c. The magnitude of the frequency response at the resonant frequency.

d. The resonant and half-power frequencies.

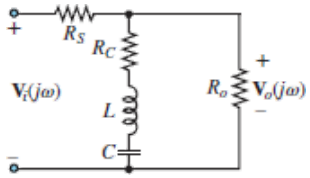


Figure P5.62

5.63 In the filter circuit shown in Figure P5.56, assume:

$$R_S = 5\,\text{k}\Omega \qquad C = 5\,\text{nF}$$
$$R_o = 50\,\text{k}\Omega \qquad L = 2\,\text{mH}$$

Determine:

a. An expression for the voltage frequency response function

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

b. The resonant frequency.

c. The half-power frequencies.

d. The bandwidth and Q.

5.64 Many stereo speakers are two-way speaker systems; that is, they have a woofer for low-frequency sounds and a tweeter for high-frequency sounds. To get the proper separation of frequencies going to the woofer and to the tweeter, *crossover* circuitry is used. A crossover circuit is effectively a bandpass, high-pass, or low-pass filter. The system model is shown in Figure P5.64. The function of the *crossover* circuitry is to channel frequencies below a given crossover frequency, f_c, into the woofer and frequencies higher than f_c into the tweeter. Assume an ideal amplifier such that $R_S = 0$ and that the desired crossover frequency is 1,200 Hz. Find C and L when $R_1 = R_2 = 8\ \Omega$. [*Hint:* Set the break frequency of the network seen by the amplifier equal to the desired crossover frequency.]

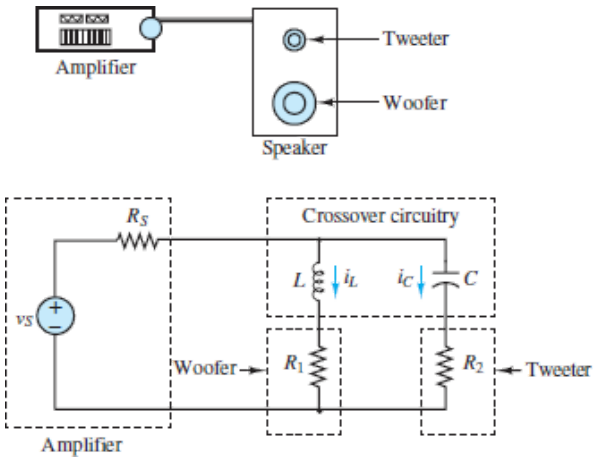


Figure P5.64

Section 5.5: Bode Plots

5.65 Determine the frequency response $\mathbf{V}_{out}(\omega)/\mathbf{V}_S(\omega)$ for the network in Figure P5.65. Generate the Bode magnitude and phase plots when $R_S = R_o = 5$ kΩ, $L = 10$ μH, and $C = 0.1$ μF.

Page 388

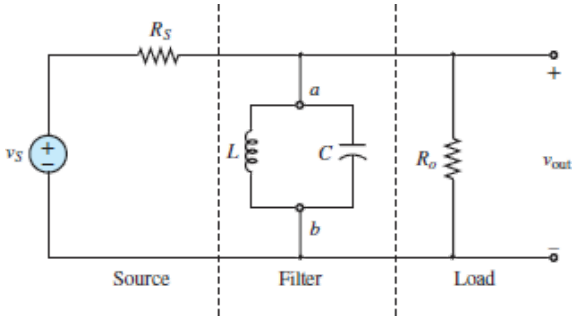


Figure P5.65

5.66 Refer to Problem 5.64 but assume that $L = 2$ mH, $C = 125$ μF, and $R_S = R_1 = R_2 = 4$ Ω in Figure P5.64.

a. Determine the impedance seen by the amplifier as a function of frequency. At what frequency is maximum power transferred by the amplifier?

b. Generate the Bode magnitude and phase plots of the currents through the woofer and tweeter.

5.67 For the notch filter shown in Figure P5.67 assume that $R_S = R_0 = 500\ \Omega$, $L = 10$ mH, and $C = 0.1\ \mu$F.

a. Determine the frequency response $\mathbf{V}_{out}(j\omega)/\mathbf{V}_S(j\omega)$.

b. Generate the associated Bode magnitude and phase plots.

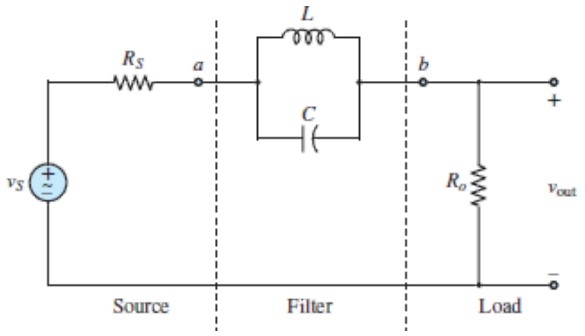


Figure P5.67

5.68 It is very common to see interference caused by power lines, at a frequency of 60 Hz. This problem outlines the design of the notch filter shown in Figure P5.68 to reject a band of frequencies around 60 Hz.

a. Determine the impedance $\mathbf{Z}_{ab}(j\omega)$ between nodes a and b for the filter of Figure P5.68. r_L represents the resistance of a practical inductor.

b. For what value of C will the center frequency of $\mathbf{Z}_{ab}(j\omega)$ equal 60 Hz when $L = 100$ mH and $r_L = 5\ \Omega$?

c. Would the "sharpness," or selectivity, of the filter increase or decrease if r_L were increased?

d. Assume that the filter is used to eliminate the 60-Hz noise from a 1-kHz sine wave. Evaluate the frequency response $\mathbf{V}_o/\mathbf{V}_{in}(j\omega)$ at both frequencies when:

$$v_g(t) = \sin(2\pi\, 1{,}000\, t)\ \text{V} \qquad r_g = 50\ \Omega$$
$$v_n(t) = 3\sin(2\pi\, 60\, t) \qquad R_o = 300\ \Omega$$

Assume $L = 100$ mH and $r_L = 5\ \Omega$. Use the value of C found in part b.

e. Generate the Bode magnitude and phase plots for $\mathbf{V}_o/\mathbf{V}_{in}$. Mark the plots at 60 Hz and 1,000 Hz.

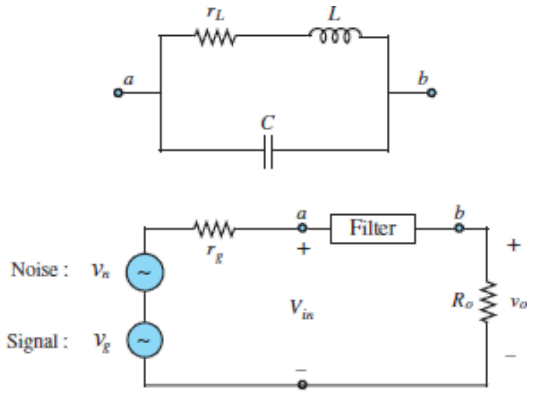


Figure P5.68

5.69 The circuit of Figure P5.69 is representative of an amplifier-speaker connection. The crossover filter allows low-frequency signals to pass to the woofer. The filter's topography is known as a π network.

a. Find the frequency response $\mathbf{V}_o(j\omega)/\mathbf{V}_S(j\omega)$.

b. If $C_1 = C_2 = C$, $R_S = R_o = 600\ \Omega$, and $1/\sqrt{LC} = R/L = 1/RC = 2{,}000\pi$, generate the Bode magnitude and phase plots in the range $100\ \text{Hz} \le f \le 10\ \text{kHz}$.

Page 389

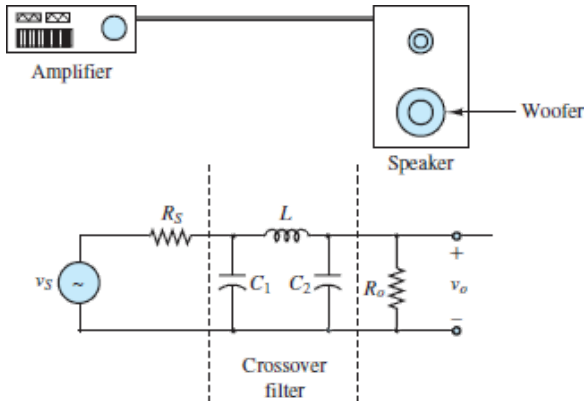


Figure P5.69

5.70 For the circuit shown in Figure P5.70:

a. Determine the frequency response:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_{\text{out}}(j\omega)}{\mathbf{V}_{\text{in}}(j\omega)}$$

b. Sketch, by hand, the associated Bode magnitude and phase plots. List all the steps in constructing the plot. Clearly show the break frequencies on the frequency axis. (*Hint:* Use the MatLab command "roots" or a calculator to quickly determine the polynomial roots.)

c. Use the MatLab command "Bode" to generate the same plots. Verify your sketch. Assume $R_1 = R_2 = 2\text{k}\Omega$, $L = 2$ H, $C_1 = C_2 = 2$ mF.

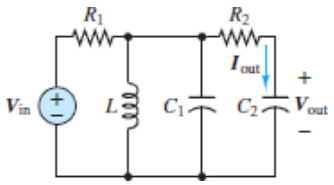


Figure P5.70

5.71 Repeat all parts of Problem 5.70 for the frequency response:

$$\mathbf{H}(j\omega) = \frac{\mathbf{I}_{out}(j\omega)}{\mathbf{V}_{in}(j\omega)}$$

Use the same component values as in Problem 5.70.

5.72 Repeat all parts of Problem 5.70 for the circuit of Figure P5.72 and the frequency response:

$$\mathbf{H}(j\omega) = \frac{\mathbf{V}_{out}(j\omega)}{\mathbf{I}_{in}(j\omega)}$$

Let $R_1 = R_2 = 1$ kΩ, $C = 1$ μF, $L = 1$ H.

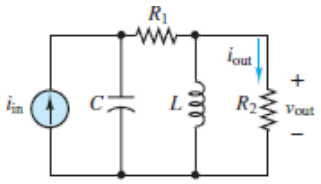


Figure P5.72

5.73 Repeat all parts of Problem 5.70 for the circuit of Figure P5.72 and the frequency response:

$$\mathbf{G}_I(j\omega) = \frac{\mathbf{I}_{out}(j\omega)}{\mathbf{I}_{in}(j\omega)}$$

Use the same values as in Problem 5.72.

5.74 For the circuit of Figure P5.74 determine the frequency response $\mathbf{H}(j\omega) = \mathbf{V}_{out}/\mathbf{I}_{in}$. Repeat all parts of Problem 5.70. Assume $R_1 = R_2 = 2$ kΩ, $C_1 = C_2 = 1$ mF.

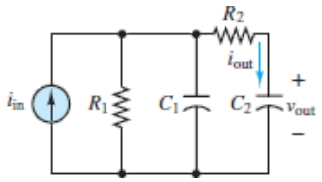


Figure P5.74

5.75 Repeat all parts of Problem 5.70 for the circuit of Figure P5.74 and the frequency response:

$$\mathbf{G}_I(j\omega) = \frac{\mathbf{I}_{out}(j\omega)}{\mathbf{I}_{in}(j\omega)}$$

Use the same component values as in Problem 5.74.

5.76 Refer to Figure P5.34 and assume $R_1 = 300\ \Omega$, $R_2 = R_3 = 500\ \Omega$, $L = 4$H, $C_1 = 40\ \mu$F, $C_2 = 160\ \mu$F.

a. Determine the frequency response:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_{out}(j\omega)}{\mathbf{V}_{in}(j\omega)}$$

b. Sketch, by hand, the associated Bode magnitude and phase plots. List all the steps in constructing the plot. Clearly show the break frequencies on the frequency axis. (*Hint:* Use the MatLab command "roots" or a calculator to quickly determine the polynomial roots.)

c. Use the MatLab command "Bode" to generate the same plots. Verify your sketch.

Page 390

5.77 Refer to Figure P5.34 and the parameter values listed in Problem 5.76.

a. Determine for the frequency response:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_C(j\omega)}{\mathbf{V}_{in}(j\omega)}$$

b. Repeat parts b and c of Problem 5.76 for this frequency response.

5.78 Refer to Figure P5.35 and repeat the instructions of parts b and c of Problem 5.76. Assume $R_1 = 20\ \text{k}\Omega$, $R_2 = 100\ \text{k}\Omega$, $L = 1\ \text{H}$, $C = 100\ \mu\text{F}$.

5.79 Assume in a certain frequency range that the ratio of output amplitude to input amplitude is proportional to $1/\omega^3$. What is the slope of the Bode magnitude plot in this frequency range, expressed in dB/decade?

5.80 Assume that the amplitude of an output voltage depends on frequency according to:

$$\mathbf{V}(j\omega) = \frac{A\omega + B}{\sqrt{C + D\omega^2}}$$

Find:

a. Each break frequency in terms of A, B, C and D.

b. The slope (in dB/decade) of the Bode magnitude plot at the high-frequency end.

c. The slope (in dB/decade) of the Bode plot at the low-frequency end.

5.81 Determine the equivalent impedance $\mathbf{Z}_{eq}$ in standard form as defined in Figure P5.81(a). Choose the Bode plot from Figure P5.81(b) that best describes the behavior of the impedance as a function of frequency. Describe how to find the resonant and cutoff frequencies, and the magnitude of the impedance for those ranges where it is constant. Label the Bode plot to indicate which feature you are discussing.

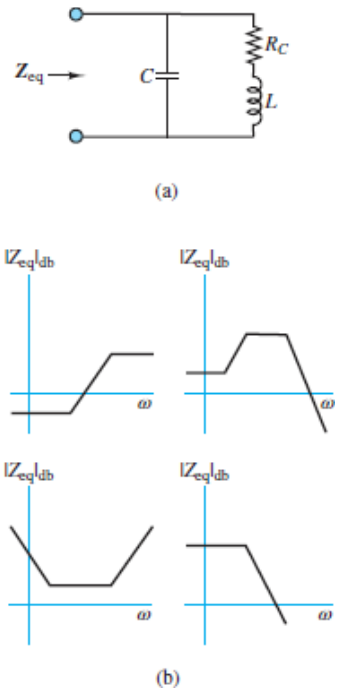


Figure P5.81

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]The circuitry in a high-fidelity audio system is far more complex than the circuits discussed in this chapter. However, from the standpoint of intuition and everyday experience, the audio analogy provides a useful example. The audio spectrum terms *bass*, *midrange*, and *treble* are well known, but not well understood. The material presented in this chapter provides a technical basis for understanding these concepts.

[2]These definitions are the same as those introduced in the section on second-order transient response in Chapter 4.

[3]See the Focus on Measurements box: "Electrocardiogram Amplifier" in Chapter 6 and section 7.2 for further information on electrocardiograms and line noise, respectively.

PART II
SYSTEMS AND INSTRUMENTATION

krithnarong Raknagn/Shutterstock

C H A P T E R

6

OPERATIONAL AMPLIFIERS

Amplification and switching are the two fundamental operations carried out by diodes and transistors, which are themselves the two fundamental electronic components. Of course, many specialized electronic devices have been developed from diodes and transistors. One of these is the *operational amplifier*, or op-amp, the mastery of which is essential to any practical application of electronics. This chapter presents the general features of an ideal amplifier and the specific features of the operational amplifier and various popular and powerful circuits based upon it. The effects of feedback in amplifier circuits is discussed as well as the gain and frequency response of the operational amplifier. The models presented in this chapter are based on concepts that have already been explored at length in earlier chapters, namely, Thévenin and Norton equivalent networks, impedance, transient response and frequency response. The chapter is designed to provide both a thorough analytical and practical understanding of the operational amplifier so that a student can successfully use it in practical amplifier circuits found in many engineering applications.

Learning Objectives

Students will learn to...

1. Understand the properties of ideal amplifiers and the concepts of gain, inpu impedance, output impedance, and feedback. *Section 6.1*.
2. Understand the difference between open-loop and closed-loop op-amp configurations; and compute the gain of (or complete the design of) simple inverting, non-inverting, summing, and differential amplifiers using ideal op-amp analysis. Analyze more advanced op-amp circuits, using ideal op-amp analysis; and identify important performance parameters in op-amp data sheets. *Section 6.2*.
3. Analyze and design simple active filters. Analyze and design ideal integrator and differentiator circuits. *Sections 6.3–6.5*.
4. Understand the principal physical limitations of an op-amp. *Section 6.6*.

6.1 IDEAL AMPLIFIERS

Amplifiers are an essential aspect of many electronic applications. Perhaps the most familiar use of an amplifier is to convert the low-voltage, low-power signal from a digital audio player (e.g., iPhone) to a level suitable for driving a pair of earbuds or headphones, as shown in Figure 6.1. Amplifiers have important applications in practically every field of engineering because the vast majority of transducers and sensors used for measurement produce electrical signals, which are then amplified, filtered, sampled, and processed by analog and digital electronic instrumentation. For example, mechanical engineers use thermistors, accelerometers, and strain gauges to convert temperature, acceleration, and strain into electrical signals. These signals must be amplified prior to transmission and then filtered (a function carried out by amplifiers) prior to sampling the data in preparation for producing a digital version of the original analog signal. Other, less obvious, functions such as impedance isolation are also performed by amplifiers. It should now be clear that amplifiers do more than simply produce an enlarged replica of a signal although that function is certainly very important. This chapter explores the general features of amplifiers and focuses on the characteristics and applications of a particularly important integrated-circuit, the **operational amplifier**.

Figure 6.1 Typical digital audio player (*fad82/Shutterstock*)

Ideal Amplifier Characteristics

The simplest model for an amplifier is depicted in Figure 6.2, where a signal v_S is amplified by a factor G, called the *voltage gain* of the amplifier. Ideally, the *input impedance* of the amplifier is infinite such that $v_{in} = v_S$; if its *output impedance* is zero, v_o will be determined by the amplifier independent of R such that:

$$v_o = Gv_{in} = Gv_S \qquad \text{Ideal amplifier} \tag{6.1}$$

Note that the input seen by the amplifier is a Thévenin source (v_S in series with R_S), while the output seen by the amplifier is a single equivalent resistance R.

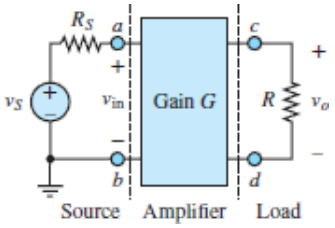


Figure 6.2 Amplifier between source and load

A more realistic (but still quite simple) amplifier model is shown in Figure 6.3. In this figure the concepts of input and output impedance of the amplifier are incorporated as single resistances R_{in} and R_{out}, respectively. That is, from the perspective Page 395of the load R the amplifier acts as a Thévenin source (Av_{in} in series with R_{out}), while from the perspective of the external source (v_S in series with R_S) the amplifier acts as an equivalent resistance R_{in}. The constant A is the

multiplier associated with the dependent (controlled) voltage source and is known as the *open-loop gain.*[1]

Using the amplifier model of Figure 6.3 and applying voltage division, the input voltage to the amplifier is now:

$$v_{ab} = v_{\text{in}} = \frac{R_{\text{in}}}{R_S + R_{\text{in}}} v_S \tag{6.2}$$

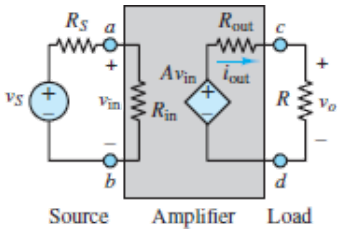


Figure 6.3 Simple voltage amplifier model

The output voltage of the amplifier can also be found by applying voltage division, where:

$$v_o = Av_{\text{in}} \frac{R}{R_{\text{out}} + R} \tag{6.3}$$

Substitute for v_{in} and divide both sides by v_S to obtain:

$$\frac{v_o}{v_S} = A \frac{R_{\text{in}}}{R_{\text{in}} + R_S} \frac{R}{R + R_{\text{out}}} \tag{6.4}$$

which is the overall voltage gain from v_S to v_o. The voltage gain G of the amplifier itself is

$$G \equiv \frac{v_o}{v_{\text{in}}} = A \frac{R}{R_{\text{out}} + R} \tag{6.5}$$

For this model, the voltage gain G is dependent upon the external resistance R, which means that the amplifier performs differently for different loads. Moreover, the input voltage v_{in} to the amplifier is a modified version of v_S. Neither of these results seem desirable. Rather, it stands to reason that the gain of a "quality" amplifier would be independent of its load and would not impact its source signal. These attributes are achieved when $R_{\text{out}} \ll R$ and $R_{\text{in}} \gg R_S$. In the limit that $R_{\text{out}} \to 0$:

$$\lim_{R_{out}\to 0} \frac{R}{R + R_{out}} = 1 \tag{6.6}$$

such that:

$$G \equiv \frac{v_o}{v_{in}} \approx A \qquad \text{when} \qquad R_{out} \to 0 \tag{6.7}$$

Also, in the limit that $R_{in} \to \infty$:

$$\lim_{R_{in}\to\infty} \frac{R_{in}}{R_{in} + R_S} = 1 \tag{6.8}$$

such that

$$v_{in} \approx v_S \qquad \text{when} \qquad R_{in} \to \infty \tag{6.9}$$

In general, a "quality" voltage amplifier will have a very small output impedance and a very large input impedance.

Page 396

Input and Output Impedance

In general, the input impedance R_{in} and the output impedance R_{out} of an amplifier are defined as:

$$R_{in} = \frac{v_{in}}{i_{in}} \qquad \text{and} \qquad R_{out} = \frac{v_{OC}}{i_{SC}} \tag{6.10}$$

where v_{OC} is the open-circuit voltage and i_{SC} is the short-circuit current at the output of the amplifier. An ideal voltage amplifier has zero output impedance and infinite input impedance so that the amplifier does not suffer from loading effects at its input or output terminals. In practice, voltage amplifiers are designed to have large input impedance and small output impedance.

It is a worthwhile exercise to show that an ideal *current amplifier* has zero input impedance and infinite output impedance. Also, an ideal *power amplifier* is designed so that its input impedance matches its source network and its output impedance matches its load impedance.

Feedback

Feedback, which is the process of using the output of an amplifier to reinforce or inhibit its input, plays a critical role in many amplifier applications. Without feedback an amplifier is said to be in *open-loop* mode; with feedback an amplifier is said to be in *closed-loop* mode. The output of the amplifier model shown in Figure 6.3 does not affect its input (because there is no path from output to input), so feedback is not present, and the model is open loop. As suggested earlier, the most basic characteristic of an amplifier is its *gain*, which is simply the ratio of the output to the input. The open-loop gain A of a practical amplifier (e.g., an operational amplifier) is usually very large, whereas the closed-loop gain G is a reduced version of the open-loop gain. The relationship between A and G is developed and explored in the rest of this chapter.

There are two types of feedback possible in closed-loop mode: *positive feedback*, which tends to reinforce the amplifier input, and *negative feedback*, which tends to inhibit the amplifier input. Both positive and negative feedback have useful applications; however, negative feedback is by far the most common type of feedback found in applications. In general, negative feedback causes the large open-loop gain A of an amplifier to be exchanged for a smaller closed-loop gain G. While this exchange may seem undesirable at first glance, several key benefits accompany the exchange. These benefits to the amplifier are

1. Decreased sensitivity to variations in circuit and environmental parameters, most notably temperature.
2. Increased bandwidth.
3. Increased linearity.
4. Increased signal-to-noise ratio.

Negative feedback is implemented by establishing one or more paths from the output to the input of the amplifier. The impedance of each feedback path can be adjusted to produce improved input and output impedances of the overall amplifier circuit. These input and output impedances are key characteristics for understanding the *loading effects* of other circuits attached to an amplifier.

Figure 6.4 shows a *signal-flow diagram* of an amplifier situated between a source and a load. The arrows indicate the direction of signal flow. The signals shown are u_s , u_f , e, and y. The output signal of each rectangle is a multiple of its input signal, where the two constants, A and β, are both positive such that:

$$y = Ae \quad \text{and} \quad u_f = \beta y \tag{6.11}$$

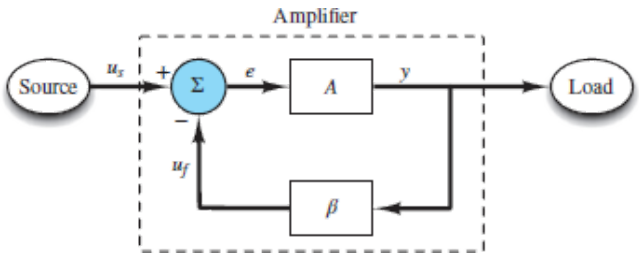


Figure 6.4 Signal-flow diagram of generic amplifier

The circle sums its inputs, u_s and u_f, to produce one output, e. The polarity signs (±) indicate that u_s and u_f make positive and negative contributions to the sum, respectively. That is

$$e = u_s - u_f = u_s - \beta y \tag{6.12}$$

Because the feedback signal u_f makes a negative contribution to the sum, the signal flow diagram of Figure 6.4 is said to employ negative feedback.

Equations 6.11 and 6.12 can be combined to yield:

$$y = Ae = A(u_s - u_f) = A(u_s - \beta y) \tag{6.13}$$

which can be rearranged to solve for y. Then, the closed-loop gain of the amplifier is

$$G \equiv \frac{y}{u_s} = \frac{A}{1 + A\beta} \tag{6.14}$$

The quantity $A\beta$ is known as the *loop gain*. Implicit in the derivation of equation 6.14 is that the behavior of each block within the amplifier is not affected by the other blocks nor by the external source and load. In other words, the blocks are *ideal* such that *loading effects* are zero.

Two important observations can be made at this point:

1. The closed-loop gain G depends upon β, which is known as the *feedback factor*.
2. Since $A\beta$ is positive, the closed-loop gain G is smaller than the open-loop gain A.

Furthermore, for most practical amplifiers, $A\beta$ is quite large such that:

$$G \approx \frac{1}{\beta} \tag{6.15}$$

This result is particularly important (and probably surprising!) because it indicates that the closed-loop gain G of the amplifier is largely *independent* of the open-loop gain A, as long as $A\beta \gg 1$, and that G is, in turn, determined largely by the feedback factor, β.

Page 398

> When $A\beta \gg 1$, the closed-loop gain G of an amplifier is determined largely by the feedback factor, β.

Furthermore, equation 6.14 can be used to find the ratio of the two inputs, u_s and u_f.

$$\frac{u_f}{u_s} = \frac{y}{u_s}\frac{u_f}{y} = \frac{A}{1 + A\beta}\beta = \frac{A\beta}{1 + A\beta} \tag{6.16}$$

Thus, when $A\beta \gg 1$, another important result is

$$\frac{u_f}{u_s} \to 1 \qquad \text{or} \qquad u_s - u_f \to 0 \tag{6.17}$$

This result indicates that when the loop gain $A\beta$ is large, the *difference* between the input signal u_s and the feedback signal u_f is driven toward zero.

> When $A\beta \gg 1$, the *difference* between the input signal u_s and the feedback signal u_f is driven toward zero.

Both of the results of equations 6.15 and 6.17 will show up repeatedly in the analysis of operational amplifier circuits in closed-loop mode.

Benefits of Negative Feedback

As mentioned in the previous section, negative feedback provides several benefits in exchange for a reduced gain. For example, take the derivative of both sides of equation 6.14 to find:

$$dG = \frac{dA}{1 + A\beta} - \frac{A\beta dA}{(1 + A\beta)^2} = \frac{dA}{(1 + A\beta)^2} \tag{6.18}$$

Divide the left side by G and the right side by $A/(1 + A\beta)$ to obtain:

$$\frac{dG}{G} = \frac{1}{1 + A\beta}\frac{dA}{A} \tag{6.19}$$

When $A\beta \gg 1$, this result indicates that the percentage change in G due to a percentage change in A is relatively small. In other words, the closed-loop gain G is relatively insensitive to changes in the open-loop gain A.

> When $A\beta \gg 1$, the closed-loop gain G is relatively insensitive to changes in the open-loop gain A.

For any amplifier, the open-loop gain A is a function of frequency. For example, the open-loop gain $A(\omega)$ of an op-amp is characterized by a simple pole such that:

$$A(\omega) = \frac{A_0}{1 + j\omega/\omega_o} \tag{6.20}$$

Page 399

where ω_o is its 3-dB break frequency. The Bode magnitude characteristic plot is shown in Figure 6.5. Equation 6.20 can be substituted into equation 6.14 to obtain:

$$G(\omega) = \frac{A(\omega)}{1 + A(\omega)\beta} = \frac{A_0/(1 + j\omega/\omega_o)}{1 + A_0\beta/(1 + j\omega/\omega_o)} \tag{6.21}$$

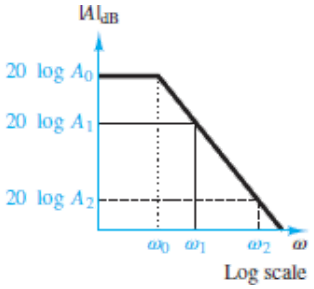


Figure 6.5 Typical amplifier Bode magnitude characteristic

Multiply the numerator and denominator on the right side of equation 6.21 by $1 + j\omega/\omega_o$ and then factor out $1 + A_0\beta$ from the denominator to obtain:

$$G(\omega) = \frac{A_0}{1 + A_0\beta}\frac{1}{1 + j\omega/\omega_g} = G_o\frac{1}{1 + j\omega/\omega_g} \tag{6.22}$$

where $\omega_g = \omega_o(1 + A_0\beta)$. Thus, the closed-loop 3-dB break frequency is $(1 + A_0\beta)$ *larger* than the open-loop 3-dB break frequency.

> The closed-loop 3-dB break frequency is $(1 + A_0\beta)$ larger than the open-loop 3-dB break frequency.

Likewise, if the amplifier is characterized by a simple zero, its 3-dB break frequency will be $(1 + A_0\beta)$ *smaller* than the open-loop 3-dB break frequency. It is a worthwhile exercise to derive this result.

Similar analyses can be performed to show the increased linearity and increased signal-to-noise ratio resulting from negative feedback. All these benefits are acquired at the expense of amplifier gain.

6.2 THE OPERATIONAL AMPLIFIER

An **operational amplifier** (op-amp) is an **integrated circuit (IC)** that contains a large number of microscopic electrical and electronic components integrated on a single silicon wafer. An op-amp can be used in conjunction with other common components to create circuits that perform amplification and filtering, as well as mathematical operations, such as addition, subtraction, multiplication, differentiation, and integration, on electrical signals. Op-amps are found in most measurement and instrumentation systems, serving as a versatile building block for many applications.

The behavior of an op-amp is well described by fairly simple models, which permit an understanding of its effects and applications without delving into its internal details. Its simplicity and versatility make the op-amp an appealing electronic device with which to begin understanding electronics and integrated circuits. Figure 6.6(d) shows a standard single op-amp IC chip pin layout. It has two input pins (2 and 3) and one output pin (6). Also notice the two DC power supply pins (4 and 7) that provide external power to the chip and thus *enable* the op-amp. Operational amplifiers are *active* devices; that is, they need an external power source to function. Pin 4 is held at a low DC voltage V_S^-, while pin 7 is held at a high DC voltage V_S^+. These two DC voltages are well below and above, respectively, the op-amp's reference voltage and bound the output of the op-amp.

Figure 6.6(a) shows the so-called small-signal, low-frequency model of an op-amp, which is exactly the same amplifier model shown in Figure 6.3. For this model, the input impedance is R_{in} and the output impedance is R_{out}. The op-amp itself is a *difference amplifier* because its output is a function of the difference between two input voltages, v^+ and v^-, which are known as the *noninverting* and *inverting* inputs, respectively. Notice that the value of the internal dependent voltage source is $A(v^+ - v^-)$, where A is the *open-loop gain* of the op-amp. In a practical op-amp, A is quite large by design, typically on the order of 10^5 to 10^7. As discussed in the previous section, this large open-loop gain can be exchanged, by design, for a smaller *closed-loop gain* G to acquire various beneficial characteristics for an amplifier circuit, of which the op-amp is just one component.[2]

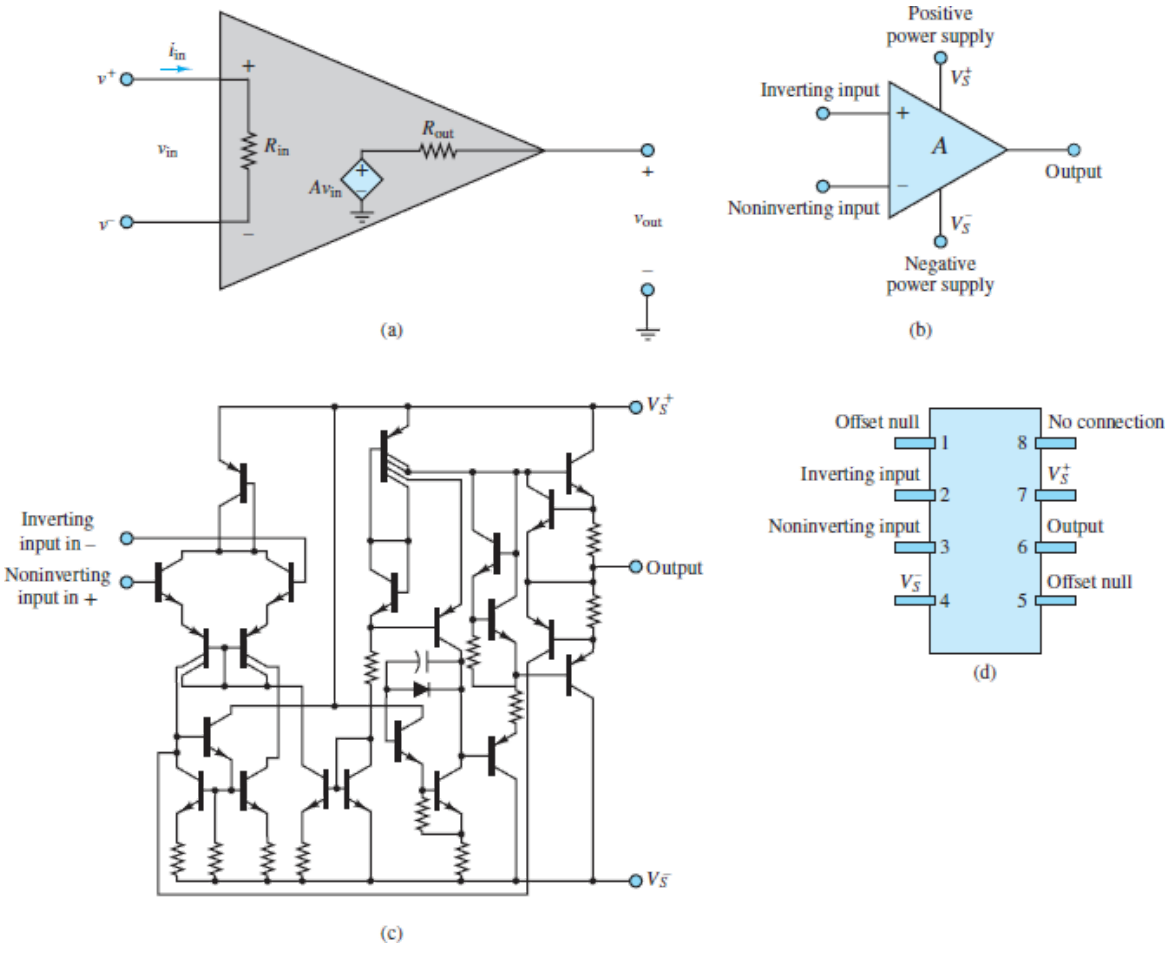


Figure 6.6 (a) Small-signal op-amp model; (b) simplified op-amp circuit symbol; (c) generic op-amp IC schematic; (d) single op-amp IC

chip pin layout

Page 401

The Ideal Op-Amp

Practical op-amps have a large open-loop gain A, as noted previously. The input impedance R_{in} is also large, typically on the order of 10^6 to 10^{12} Ω, while the output impedance R_{out} is small, typically on the order of 10^0 or 10^1 Ω. In the ideal case, the open-loop gain and the input impedance would be infinite, while the output impedance would be zero. When the output impedance is zero, the output voltage of an ideal op-amp is simply

$$v_{out} = A(v^+ - v^-) = A\Delta v \tag{6.23}$$

The implication for a practical op-amp with a very large open-loop gain A is that one of the two following possibilities will hold:

1. In the case that $\Delta v \neq 0$, the output voltage *saturates* near either the positive or negative DC power supply value, $VS+$ or $VS-$, as shown in Figure 6.7. These external DC power supply *rails* enable a practical op-amp but also bound the op-amp output voltage v_{out}. This case applies to all practical applications of an op-amp where there is no feedback from v_{out} to v^-. *Open-loop mode* applications such as a simple *comparator* operate in this fashion. Other circuits that employ positive feedback only, such as a *Schmitt trigger*, are also included in this class of applications.
2. In the case that $\Delta v = 0$, the product $A\Delta v$ is finite and the output voltage is determined by the external circuitry attached to the op-amp. Recall from Section 6.1 that when $A\beta \gg 1$ the closed-loop gain of an amplifier is approximately equal to $1/\beta$ and largely independent of A itself. Thus, this case applies to all practical applications of an op-amp in *closed-loop mode;* that is, when negative feedback is present from v_{out} to v^-.

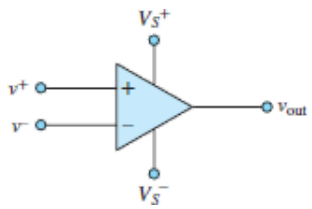


Figure 6.7 The absence of the open-loop gain symbol A in the triangle indicates an ideal operational amplifier. When $\Delta v \neq 0$ the output voltage saturates near one of the two power supply rails, v_S^+ or v_S^-.

The infinite impedances of the input terminals of an ideal op-amp imply that the current into or out of those terminals is zero. This result is known as the *first golden rule* of ideal op-amps:

$$i^+ = i^- = 0 \qquad \text{First golden rule} \tag{6.24}$$

Also recall from the discussion of negative feedback in Section 6.1 that when $A\beta \gg 1$ the difference between the two amplifier inputs, u_s and u_f, approaches zero. In the context of ideal op-amps, where $A \to \infty$, the difference between the two amplifier inputs, v^+ and v^-, will be zero *as long as there is a feedback path from* v_{out} *to* v^-.

$$v^+ = v^- \qquad \text{Second golden rule; negative feedback required} \tag{6.25}$$

The Golden Rules of Ideal Op-Amps:

1. $i^+ = i^- = 0$.
2. $v^+ = v^-$ (when negative feedback is present).

Page 402

Amplifier Archetypes

There are three fundamental amplifiers that utilize the operational amplifier and employ *negative feedback.* They are:

- The inverting amplifier.
- The noninverting amplifier.
- The unity-gain isolation buffer (or voltage follower).

These archetypes have many important applications and are the building blocks for other important amplifiers. Understanding and recognizing these archetypes is

an essential first step in the study of amplifiers based upon the op-amp. It is worth emphasizing that the op-amp is rarely used as a stand-alone amplifier; rather it is used along with other components to form specialized amplifiers.

The Inverting Amplifier

Figure 6.8 shows a basic *inverting amplifier* circuit. The name derives from the fact that the input signal v_S "sees" the inverting terminal (−) and that, as is shown below, the output signal v_o is an inverted (negative) version of the input signal. To determine the relationship between the output and the input signals, assume the op-amp is ideal and apply KCL at the inverting node marked v^-.

$$i_S = i_F + i_{\text{in}} \tag{6.26}$$

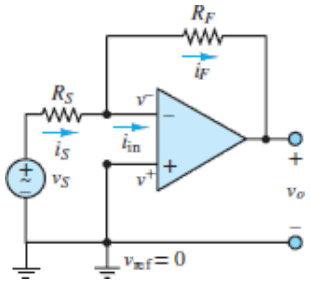


Figure 6.8 Inverting amplifier

However, the first golden rule of ideal op-amps requires that $i_{\text{in}} = 0$. Thus, $i_S = i_F$ such that R_S and R_F form a *virtual* series connection. Ohm's law can be applied to each resistor to yield:

$$i_S = \frac{v_S - v^-}{R_S} \qquad i_F = \frac{v^- - v_o}{R_F} \tag{6.27}$$

These expressions can be simplified by noting that $v^+ = 0$ and thus by the second golden rule of ideal op-amps $v^- = v^+ = 0$. Thus:

$$\frac{v_S}{R_S} = \frac{-v_o}{R_F} \tag{6.28}$$

Cross-multiply to find the closed-loop gain G:

$$G = \frac{v_o}{v_S} = -\frac{R_F}{R_S} \quad \text{Inverting amplifier} \tag{6.29}$$

Note that the magnitude of G can be greater or less than 1.

An alternate approach is to apply voltage division across the virtual series connection of R_S and R_F.

$$\frac{v_S - v_o}{v_S - 0} = \frac{R_S + R_F}{R_S}$$

or

$$1 - \frac{v_o}{v_S} = 1 + \frac{R_F}{R_S} \tag{6.30}$$

Subtract 1 from each side of this expression to find the same result as equation 6.29.

Page 403

Notice that the closed-loop gain G of an inverting amplifier is determined solely by the choice of resistors. This result was derived for an ideal op-amp. For a practical op-amp the result is only slightly different as long as the open-loop gain A is large. It is important to remember that this result depends upon both golden rules of ideal op-amps and that, in particular, the second golden rule is valid only when negative feedback is present.

As long as the open-loop gain A is large, the presence of negative feedback from the output to the inverting input drives the voltage difference between the two input terminals to zero.

The input impedance of the inverting amplifier is simply:

$$R_{\text{in}} = \frac{v_S}{i_S} = \frac{v_S - 0}{i_S} = R_S \tag{6.31}$$

Notice the important role played by the *virtual* ground at the inverting terminal in making this calculation so easy. This result also reveals a shortcoming of the inverting amplifier. In general, an ideal amplifier would have an infinite input impedance so as to not load the source network. It is tempting to correct this problem by choosing R_S to be very large; however, in so doing, the closed-loop gain (equation 6.29) will be reduced. Thus, it is not possible to design an inverting amplifier to have a large gain and also a large input impedance. Alas, there is no such thing as a free lunch!

The Noninverting Amplifier

Figure 6.9 shows a basic *noninverting amplifier* circuit. The name derives from the fact that the input signal v_S "sees" the noninverting terminal (+) and that, as is shown below, the output signal v_o is a noninverted (positive) version of the input signal. To determine the relationship between the output and the input signals assume the op-amp is ideal and apply KCL at the inverting node marked v^-.

$$i_F = i_1 + i_{in} \tag{6.32}$$

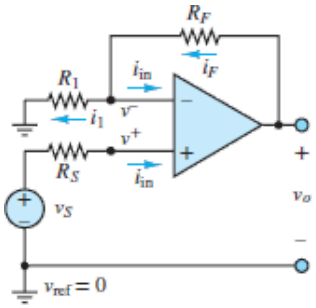


Figure 6.9 Noninverting amplifier

However, the first golden rule of ideal op-amps states that $i_{in} = i^- = i^+ = 0$. Thus, $i_F = i_1$ such that R_F and R_1 form a *virtual* series connection. Ohm's law can be applied to each resistor to yield:

$$i_1 = \frac{v^- - 0}{R_1} \qquad i_F = \frac{v_o - v^-}{R_F}$$

or

$$\frac{v^-}{R_1} = \frac{v_o - v^-}{R_F} \tag{6.33}$$

Since there is negative feedback present, the second golden rule of ideal op-amps can be applied such that $v^- = v^+$. Notice that because $i_{in} = 0$, the voltage drop Page 404across R_S is zero with the result that $v^- = v^+ = v_S$. Substitute this result into equation 6.33 and rearrange terms to yield the closed-loop gain G:

$$G = \frac{v_o}{v_S} = 1 + \frac{R_F}{R_1} \qquad \text{Noninverting amplifier} \tag{6.34}$$

Note that $G \leq 1$.

An alternate approach is to apply voltage division across the virtual series connection of R_1 and R_F.

$$\frac{v_o - 0}{v^- - 0} = \frac{R_1 + R_F}{R_1} \tag{6.35}$$

Since $v^- = v^+ = v_S$:

$$\frac{v_o}{v_S} = 1 + \frac{R_F}{R_1} \tag{6.36}$$

which is the same result as that found in equation 6.34.

Notice that the closed-loop gain G of a noninverting amplifier is determined solely by the choice of resistors. This result was derived for an ideal op-amp. For a practical op-amp the result is only slightly different as long as the open-loop gain A is large. It is important to remember that this result depended upon both golden rules of ideal op-amps and that, in particular, the second golden rule is valid only when negative feedback is present.

As long as the open-loop gain A is large, the presence of negative feedback from the output to the inverting input drives the voltage difference between the two input terminals to zero.

The input impedance of the noninverting amplifier is simply

$$R_{\text{in}} = \frac{v^+}{i_{\text{in}}} \rightarrow \infty \tag{6.37}$$

In practice, the input impedance of a noninverting amplifier is very large due to the very large input impedance at the noninverting terminal, which limits i_{in} to very small values. Notice that the closed-loop gain of the noninverting amplifier is independent of its input impedance. Thus, the noninverting amplifier does not suffer from a trade-off between gain and input impedance, as does the inverting amplifier. However, the gain of a noninverting amplifier is limited to values greater than one, whereas the gain of the inverting amplifier can take on any value. Alas, again there is no such thing as a free lunch!

Unity-Gain Isolation Buffer or Voltage Follower

Figure 6.10 shows a *unity-gain isolation buffer*, which is also known as a *voltage follower*. Notice that the input signal v_S "sees" the noninverting terminal (+) such that the output signal v_o should be a noninverted (positive) version of v_S. Page 405The analysis of this circuit is as simple as the circuit itself. Assume that the op-amp is ideal. Since negative feedback is present, both golden rules are valid. That is

$$i^+ = i^- = 0 \qquad \text{and} \qquad v^+ = v^- \tag{6.38}$$

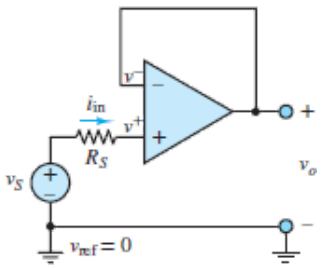


Figure 6.10 Unity-gain isolation buffer or voltage follower

By observation, $v^+ = v_S$ and $v^- = v_o$ with the result that the closed-loop gain G is

$$G = \frac{v_o}{v_S} = 1 \qquad \text{Unity-gain isolation buffer (voltage follower)} \tag{6.39}$$

The reason this circuit is called a voltage follower should now be obvious: The output voltage v_o "follows" (equals) the input voltage v_S. The reason this circuit is also known as a unity-gain isolation buffer is due to the infinite *input impedance* of the ideal op-amp, such that the voltage source v_S experiences no loading effect. In general, the output of the op-amp itself will be attached to a load. However, the output terminal will supply whatever current is necessary to maintain the output voltage at v_S . Thus, the source v_S is *isolated* or *buffered* from the output.

The input impedance of an isolation buffer is simply

$$R_{\text{in}} = \frac{v_S}{i_{\text{in}}} \to \infty \tag{6.40}$$

In practice, the input impedance of an isolation buffer is very large due to the very large input impedance of the op-amp, which limits i_{in} to very small values. The closed-loop gain is fixed at unity as long as the open-loop gain A is large such that v^- will be driven to v^+ by negative feedback.

Application of Thévenin's theorem

Notice in Figures 6.8 and 6.9 that the input is represented as a Thévenin source. The implication is that the previous results for inverting and noninverting amplifiers can be applied to any case where the input of the amplifier circuit is linear and can be simplified to an equivalent Thévenin source. In other words, R_S and v_S are the Thévenin equivalent resistance and the open-circuit voltage, respectively, of any arbitrary linear input circuit.

For example, consider the inverting amplifier circuit shown in Figure 6.11. It does not have the same form as the archetype of Figure 6.8. However, the voltage source v_{in} "sees" the inverting terminal; therefore, the output voltage v_o will be an inverted version of v_{in}. The circuit is an inverting amplifier. To solve for v_o replace the entire linear network to the left of terminals a and b with its Thévenin equivalent.

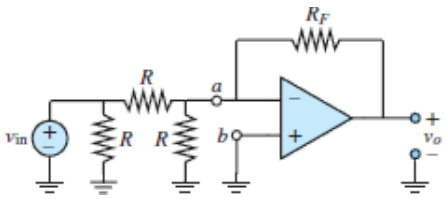


Figure 6.11 Inverting amplifier before simplification to archetype

Page 406

Figure 6.12 shows the source network detached at terminals *a* and *b*. To find the Thévenin equivalent resistance between those terminals turn off the independent voltage source (set $v_{in} = 0$) and replace it with a short-circuit. Then:

$$R_T = R_{ab} = R \| R = \frac{R}{2} \tag{6.41}$$

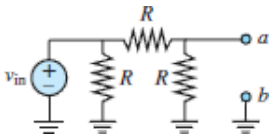


Figure 6.12 Source network detached at terminals *a* and *b*

The Thévenin (open-circuit) voltage across terminals *a* and *b* can be found from voltage division:

$$v_T = v_{ab} = \frac{R}{R+R} v_{in} = \frac{v_{in}}{2} \tag{6.42}$$

The Thévenin equivalent source network attached to the rest of the amplifier circuit is shown in Figure 6.13. Notice that the simplified amplifier is now identical in form to the inverting amplifier archetype of Figure 6.8. Thus, using equation 6.29:

$$\frac{v_o}{v_T} = -\frac{R_F}{R_T} = -\frac{2R_F}{R} \tag{6.43}$$

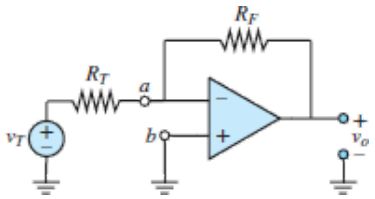


Figure 6.13 Inverting amplifier after simplification to archetype

The closed-loop gain G of the original amplifier circuit shown in Figure 6.11 is

$$G = \frac{v_o}{v_{\text{in}}} = \frac{v_o}{v_T} \cdot \frac{v_T}{v_{\text{in}}} = -\frac{2R_F}{R} \cdot \frac{1}{2} = -\frac{R_F}{R} \tag{6.44}$$

Figure 6.13 generalizes Figure 6.8 by representing explicitly the source network as the Thévenin equivalent network of any linear input source network. The same approach can be taken to generalize the noninverting amplifier and isolation buffer circuits shown in Figures 6.9 and 6.10, respectively, where v_S and R_S are now the Thévenin (open-circuit) voltage and the Thévenin equivalent resistance, respectively, of the input source network.

Multiple Sources and the Principle of Superposition

There are many situations that call for an amplifier to accommodate multiple input source networks. The analysis of these amplifiers can be accomplished using basic principles, such as KCL, KVL, and Ohm's law. However, it is often useful to apply the principle of superposition to simplify the overall amplifier circuit into multiple component amplifiers, each with only one independent source still turned on. Page 407Thévenin's theorem can often be used to transform these component amplifiers into one of the amplifier archetypes: the inverting amplifier, the noninverting amplifier, or the isolation buffer. Two important examples of amplifiers with multiple input sources are the summing amplifier and the difference amplifier.

The Summing Amplifier

A useful op-amp circuit that is based on the inverting amplifier is the **op-amp summer**, or **summing amplifier**, shown in Figure 6.14. Assume the op-amp is ideal. The first golden rule of op-amps states that $i^+ = i^- = 0$. Thus, when KCL is applied at the inverting node, the result is

$$\sum_{n=1}^{N} i_n = i_1 + i_2 + \cdots + i_N = i_F \tag{6.45}$$

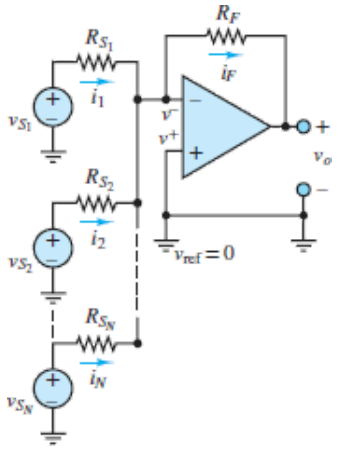


Figure 6.14 Summing amplifier

Since negative feedback is present, the second golden rule is also valid such that $v^- = v^+ = 0$. Ohm's law can then be applied at each resistor to obtain:

$$i_n = \frac{v_{S_n} - 0}{R_{S_n}} \qquad n = 1, 2, \ldots, N$$

and

$$i_F = \frac{0 - v_o}{R_F} \tag{6.46}$$

The results of equations 6.46 can be plugged into equation 6.45 to find:

$$\sum_{n=1}^{N} \frac{v_{S_n}}{R_{S_n}} = -\frac{v_o}{R_F}$$

or

$$v_o = -\sum_{n=1}^{N} \frac{R_F}{R_{S_n}} v_{S_n} \tag{6.47}$$

The output of the summing amplifier is the weighted sum of N input sources, where the weighting factor for each source v_{S_n} is the ratio of the feedback resistance R_F to the source resistance R_{S_n}. Notice that if $R_S = R_{S_1} = R_{S_2} = \cdots = R_{S_N}$, then:

$$v_o = -\frac{R_F}{R_S}\sum_{n=1}^{N} v_{S_n} \qquad \text{Summing amplifier} \tag{6.48}$$

The summing amplifier can also be analyzed using the principle of superposition. Consider the case where all the voltage sources except $R_2, \ldots, R_N$ are turned off. Then the voltage drops across the resistors $R_2, \ldots, R_N$ are all zero since $v^- = 0$ and a zero voltage source is equivalent to a short-circuit. Thus, for this Page 408case, $i_2 = i_3 = \cdots = i_N = 0$ as shown in Figure 6.15. Since $i^+ = i^- = 0$ for an ideal op-amp, KCL applied at the inverting terminal node yields simply:

$$i_1 = i_F \tag{6.49}$$

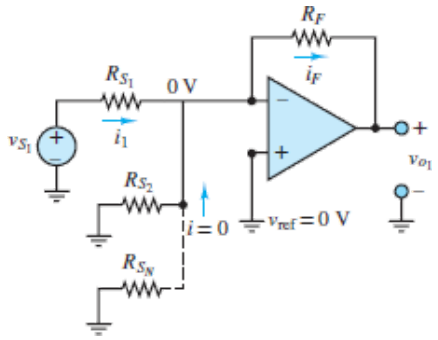


Figure 6.15 Summing amplifier with only one source turned on

Again, because negative feedback is present, the second golden rule is valid such that $v^- = v^+ = 0$. Ohm's law can then be applied to R_{S_1} and R_F to obtain:

$$i_1 = \frac{v_{S_1} - 0}{R_{S_1}} \qquad \text{and} \qquad i_F = \frac{0 - v_{o_1}}{R_F} \tag{6.50}$$

Plug these two results into equation 6.49 and rearrange to yield:

$$v_{o_1} = -\frac{R_F}{R_{S_1}} v_{S_1} \tag{6.51}$$

where v_{o_1} is the component of v_o due to the voltage source v_{S_1}. It is worth noting that this result is equivalent to what would be obtained for the inverting amplifier archetype shown in Figure 6.16. This equivalence is due to the fact that the currents $i_2, i_3, \ldots, i_N$ are all zero such that R_{S_1} and R_F are in a *virtual* series connection as in the inverting amplifier archetype.

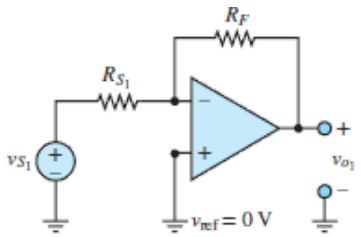


Figure 6.16 Equivalent inverting amplifier circuit for summing amplifier with only one source turned on

Since the Thévenin source pairs v_{S_n} and R_{S_n} in Figure 6.14 are all in parallel, the component of v_o due to v_{S_n} is:

$$v_{o_n} = -\frac{R_F}{R_{S_n}} v_{S_n} \qquad n = 1, 2, \ldots, N \tag{6.52}$$

Summing all these component contributions yields:

$$v_o = -\sum_{n=1}^{N} \frac{R_F}{R_{S_n}} v_{S_n} \tag{6.53}$$

which is the same result as that found in equation 6.47.

The Difference Amplifier

A useful op-amp circuit that is based on the inverting and noninverting amplifier archetypes is the **difference** or **differential amplifier** shown in Figure 6.17. This amplifier is frequently used to subtract one signal from another and perhaps amplify that difference as well, as is done in the example discussed in the Focus on Measurements box, "Electrocardiogram (EKG) Amplifier."

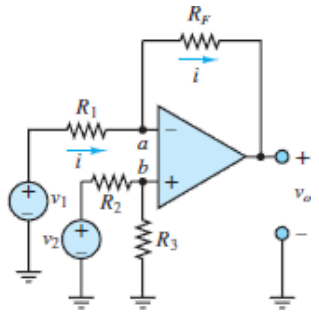


Figure 6.17 Amplifier with input sources at the inverting and noninverting terminals

The analysis of the difference amplifier can be accomplished by applying basic principles (e.g., KCL, Ohm's law) or by applying the principle of superposition. Both Page 409approaches will assume an ideal op-amp, and since negative feedback is present, both golden rules are valid. The former approach begins by noting that $i^+ = i^- = 0$ such that R_1 and R_F are in a *virtual* series connection as are R_2 and R_3 . Thus, the voltage at the noninverting terminal v^+ can be computed from voltage division.

$$v^+ = \frac{R_3}{R_3 + R_2} v_2 = \frac{R_3/R_2}{1 + (R_3/R_2)} v_2 \tag{6.54}$$

Likewise, voltage division along the other virtual series connection yields:

$$i = \frac{v_1 - v_o}{R_1 + R_F} = \frac{v^- - v_o}{R_F} \tag{6.55}$$

Solving for v^- yields:

$$v^- = \frac{R_F v_1 + R_1 v_o}{R_1 + R_F} = \frac{(R_F/R_1) v_1 + v_o}{1 + (R_F/R_1)} \tag{6.56}$$

The second golden rule is $v^+ = v^-$ such that:

$$\frac{R_3/R_2}{1 + (R_3/R_2)} v_2 = \frac{(R_F/R_1) v_1 + v_o}{1 + (R_F/R_1)}$$

or

$$v_o = \frac{1 + (R_F/R_1)}{1 + (R_3/R_2)} \frac{R_3}{R_2} v_2 - \frac{R_F}{R_1} v_1 \tag{6.57}$$

In this form the expression for v_o is too complicated to leave much of an impression. However, it is greatly simplified by choosing the resistor values to satisfy:

$$\frac{R_F}{R_1} = \frac{R_3}{R_2} \tag{6.58}$$

such that:

$$v_o = \frac{R_F}{R_1}(v_2 - v_1) \qquad \text{Difference amplifier} \tag{6.59}$$

Figure 6.18 shows one particular version of a difference amplifier where equation 6.58 is satisfied by setting $R_3 = R_F$ and $R_2 = R_1$.

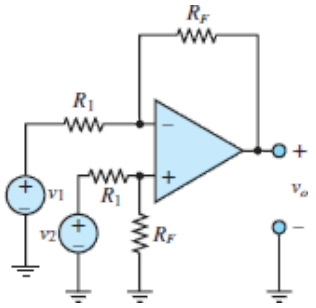


Figure 6.18 Difference amplifier

The circuit in Figure 6.17 can also be analyzed using the principle of superposition. The op-amp is still assumed to be ideal, and since negative feedback is present, both golden rules are valid. To begin, set $v_2 = 0$ and find the component of v_o due to v_1 as shown in Figure 6.19. Since $i^+ = 0$ the voltage drops across R_2 and R_3 are zero such that $v_1^+ = 0$. Thus, the circuit is equivalent to the inverting amplifier shown in Figure 6.20 with the output given by:

$$v_{o_1} = -\frac{R_F}{R_1}v_1 \tag{6.60}$$

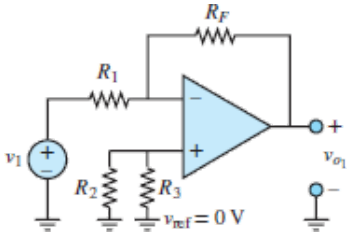


Figure 6.19 Amplifier of Figure 6.17 when $v_2 = 0$

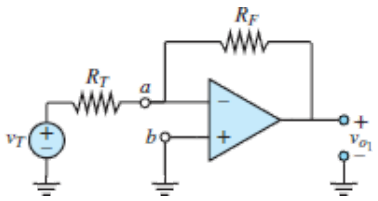


Figure 6.20 An inverting amplifier results when $v_2 = 0$, where $V_T = v_1$ and $R_T = R_1$.

Now set $v_1 = 0$ in Figure 6.17 and find the component of v_o due to v_2 as shown in Figure 6.21. Since $i^+ = 0$, R_2 and R_3 are in a *virtual* series connection. Apply voltage division to yield:

$$v_2^+ = \frac{R_3}{R_3 + R_2} v_2 \tag{6.61}$$

Page 410

Thus, the circuit is equivalent to the noninverting amplifier shown in Figure 6.22 with the output given by:

$$v_{o_2} = \left(1 + \frac{R_F}{R_1}\right) v_2^+ = \left(1 + \frac{R_F}{R_1}\right) \frac{R_3}{R_3 + R_2} v_2 \tag{6.22}$$

Finally, apply the principle of superposition to obtain:

$$v_o = v_{o_1} + v_{o_2} = -\frac{R_F}{R_1} v_1 + \left(1 + \frac{R_F}{R_1}\right) \frac{R_3}{R_3 + R_2} v_2 \tag{6.63}$$

As before, this expression is greatly simplified by choosing the resistor values such that:

$$\frac{R_F}{R_1} = \frac{R_3}{R_2} \tag{6.64}$$

The result of this choice when applied to equation 6.63 is (of course!) equation 6.59.

Both of the solution methods shown above are completely valid. However, the principle of superposition has the added appeal of determining the individual contributions of each input source and therefore allows for a quick recalculation of the solution when only one of the input sources is changed.

When the linear source networks seen by the input terminals are more complicated than those shown in Figure 6.17 it is possible to simplify those networks using Thévenin's theorem. For example, the source network seen by the noninverting terminal in Figure 6.21 can be replaced with that shown in Figure 6.22, where:

$$v_T = \frac{R_3}{R_3 + R_2} v_2 \qquad \text{and} \qquad R_T = R_2 \| R_3 \tag{6.65}$$

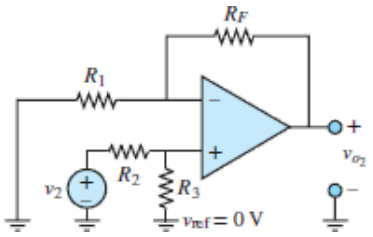


Figure 6.21 A noninverting amplifier results when $v_1 = 0$.

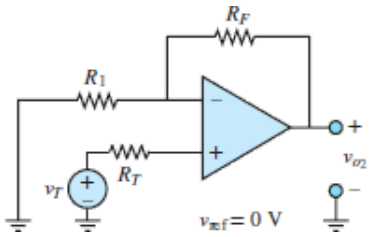


Figure 6.22 A noninverting amplifier results when $v_1 = 0$, where v_T and R_T are computed as shown in equation 6.65.

Page 411

Common and Difference Modes

It is often necessary to amplify the difference between two signals that may both be corrupted by noise or interference. The two input signals v_1 and v_2 can be decomposed into two parts: the **common mode** (CM) and the **difference** or **differential mode** (DM). These two modes are defined mathematically as:

$$v_{CM} = \frac{v_1 + v_2}{2} \qquad \text{and} \qquad v_{DM} = v_2 - v_1 \tag{6.66}$$

where the common mode v_{CM} is the average value of v_1 and v_2.

$$v_1 = v_{CM} - \frac{v_{DM}}{2} \qquad \text{and} \qquad v_2 = v_{CM} + \frac{v_{DM}}{2} \tag{6.67}$$

With these definitions, the output of an ideal difference amplifier is

$$v_o = \frac{R_F}{R_1}(v_2 - v_1) = \frac{R_F}{R_1} v_{DM} \tag{6.68}$$

In other words, the common mode of the two input signals is *rejected* by the difference amplifier. In many situations, the noise and interference of one input is identical to (or nearly the same as) that of the other input. Thus, a difference amplifier can be used to eliminate noise and interference that is common to both inputs. In practice, the output of a difference amplifier is given by:

$$v_o = A_{\text{DM}}(v_2 - v_1) + A_{\text{CM}}\left(\frac{v_2 + v_1}{2}\right) \tag{6.69}$$

where A_{DM} and A_{CM} are the difference-mode and common-mode gains, respectively. Ideally, $A_{CM} = 0$, such as for the ideal op-amp circuit of Figure 6.18. In practice, the extent to which a practical difference amplifier rejects the common mode is known as the **common-mode rejection ratio (CMRR)**:

$$\text{CMRR} = 20 \log \left|\frac{A_{\text{DM}}}{A_{\text{CM}}}\right| \quad \text{(in dB)} \tag{6.70}$$

For example, op-amps themselves are difference amplifiers. A particular op-amp known as the 741 has a typical CMRR of 90 dB. The Focus on Measurements box, "Electrocardiogram (EKG) Amplifier," examines a typical application of a difference amplifier.

Table 6.1 summarizes the basic op-amp circuits presented in this section.

Table 6.1 Summary of basic amplifiers

Configuration	Circuit diagram	Output voltage (ideal op-amp)
Inverting amplifier	Figure 6.8	$-\frac{R_F}{R_S}v_S$
Noninverting amplifier	Figure 6.9	$\left(1+\frac{R_F}{R_S}\right)v_S$
Unity-gain isolation buffer	Figure 6.10	v_S
Summing amplifier	Figure 6.14	$-\frac{R_F}{R_S}\sum_{n=1}^{N} v_{S_n}$
Difference amplifier	Figure 6.18	$=\frac{R_F}{R_1}(v_2 - v_1)$

FOCUS ON MEASUREMENTS

Electrocardiogram (EKG) Amplifier

This example illustrates the principle behind a two-lead **electrocardiogram (EKG) measurement**. The desired cardiac waveform is given by the difference between the potentials measured by two electrodes suitably placed on the patient's chest, as shown in Figure 6.23. A healthy, noise-free EKG waveform $v_1 - v_2$ is shown in Figure 6.24.

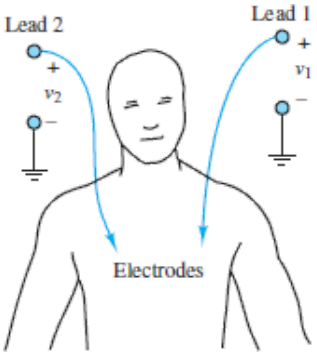


Figure 6.23 Two-lead electrocardiogram

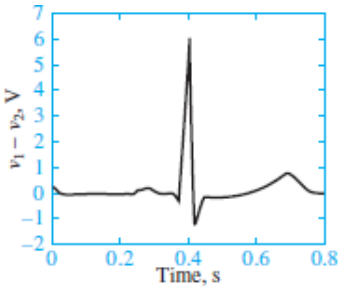


Figure 6.24 EKG waveform

Unfortunately, noise present on the 60-Hz, 110-V AC line used to power the equipment may appear in the EKG itself, due to capacitive coupling. Ambient electromagnetic interference can also interact with the closed-loop formed by the lead wires to generate another source of noise. Other sources of noise include changes at the electrode-skin interface due to respiration, muscle contractions, and other displacements. In addition, different DC offsets due to the electrodes complicate the signals. The signal processing associated with an actual EKG involves instrumentation amplifiers (see Example 6.2) and active filters (see Section 6.3). In this example, the focus is limited to the role of a difference amplifier in rejecting common-mode 60-Hz noise found in a typical EKG. That noise can be represented as a cosine function with angular frequency 377 rad/sec (which is equivalent to 60 Hz) as shown here.

Lead 1:

$$v_1(t) + v_n(t) = v_1(t) + V_n \cos(377t + \phi_n)$$

Lead 2:

$$v_2(t) + v_n(t) = v_2(t) + V_n \cos(377t + \phi_n)$$

As shown in Figure 6.25, the interference signal $V_n \cos(377t + \phi_n)$ is approximately the same at both leads because the electrodes are designed to be identical and are used in close proximity to each other. If the resistors of the difference amplifier are properly matched, the voltage output will be:

$$v_o = \frac{R_2}{R_1}[(v_1 + v_n) - (v_2 + v_n)] = \frac{R_2}{R_1}(v_1 - v_2)$$

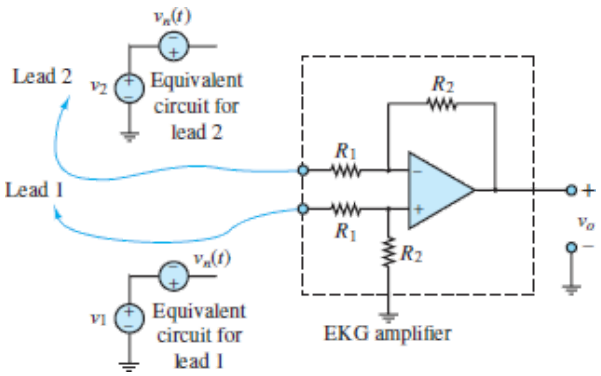


Figure 6.25 EKG amplifier

Thus, common-mode 60-Hz noise is eliminated, or greatly reduced, while the desired EKG waveform is amplified. Great! In practice, the common-mode rejection ratio is not infinite but can be made quite large to satisfy the design specifications required for a proper diagnosis.

Page 413

FOCUS ON MEASUREMENTS

Sensor Calibration Circuit

In many practical instances, the output of a sensor is related to the physical variable we wish to measure in a form that requires some signal conditioning. The most desirable form of a sensor output is one in which the electrical output of the sensor (e.g., voltage) is related to the physical variable by a constant factor. Such a relationship is depicted in Figure 6.26(a), where k is the calibration constant relating voltage to temperature. Note that k is a positive number, and that the *calibration curve* passes through the (0, 0) point. On the other hand, the sensor characteristic of Figure 6.26(b) is described by:

$$v_{sensor} = -\beta T + V_0$$

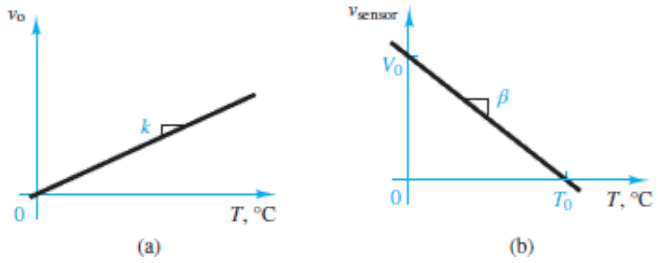


Figure 6.26 Sensor calibration curves

It is possible to modify the sensor calibration curve of Figure 6.26(b) to the more desirable one of Figure 6.26(a) by means of the simple circuit displayed in Figure 6.27. The inverting gain R_F/R_S of this circuit is used to convert the negative temperature coefficient (NTC) β to the desired positive calibration constant k. The zero (or bias) offset is adjusted by means of a potentiometer connected to the voltage supplies to produce a constant reference voltage V_{ref} at the noninverting terminal.

Page 414

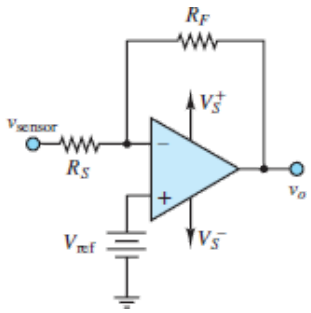


Figure 6.27 Sensor calibration circuit

When $V_{ref} = 0$, the sensor sees an inverting amplifier such that:

$$(v_o)_{sensor} = -\frac{R_F}{R_S} v_{sensor}$$

Likewise, when $v_{sensor} = 0$, the battery sees a noninverting amplifier such that:

$$(v_o)_{ref} = 1 + \frac{R_F}{R_S} V_{ref}$$

The total output is given by the principle of superposition.

$$\begin{aligned} v_o &= (v_o)_{\text{sensor}} + (v_o)_{\text{ref}} \\ &= -\frac{R_F}{F_S} v_{\text{sensor}} + \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}} \\ &= -\frac{R_F}{R_S}(-\beta T + V_0) + \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}} \end{aligned}$$

The requirement for a linear response such as that shown in Figure 6.26(a) is $v_o = kT$, which is satisfied by suitable choices of R_F, R_S, and V_{ref} such that:

$$kT = -\frac{R_F}{R_S}(-\beta T + V_0) + \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}}$$

For this equation to hold, the coefficients of T on both sides must be equal and the sum of the constant terms on the right side must equal zero. That is

$$k = \frac{R_F}{R_S}\beta$$

and

$$\frac{R_F}{R_S} V_0 = \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}}$$

or

$$V_{\text{ref}} = \frac{R_F}{R_S + R_F} V_0$$

Page 415

It is worth noting that $V_{\text{ref}} \approx V_0$ when $R_F \gg R_S$. Thus, when this condition holds, the appropriate battery voltage for the sensor calibration circuit can be determined directly from the sensor calibration curve of Figure 6.26(b). To avoid loading the sensor output pick R_S large enough that the amplifier input resistance seen by the sensor is much larger than the output resistance of the sensor.

It is also worth noting that the effect of the battery voltage is to raise or lower the calibration curve. For this reason, this circuit is known more generally as a *level shifter.* See Example 6.3 for further discussion.

EXAMPLE 6.1 Inverting Amplifier Circuit

Problem

Determine the voltage gain and output voltage for the inverting amplifier circuit of Figure 6.8. What will the uncertainty in the gain be if 5 and 10 percent tolerance resistors are used, respectively?

Solution

Known Quantities: Feedback and source resistances, source voltage.

Find: $G = v_o/v_S$; maximum percent change in G for 5 and 10 percent tolerance resistors.

Schematics, Diagrams, Circuits, and Given Data: $R_S = 1\ \text{k}\Omega$; $R_F = 10\ \text{k}\Omega$; $v_S(t) = A\cos(\omega t)$; $A = 0.015$ V; $\omega = 50$ rad/s.

Assumptions: The amplifier behaves ideally; that is, the input current into the op-amp is zero, and negative feedback forces $v^+ = v^-$.

Analysis: Using equation 6.29, the output voltage is

$$v_o(t) = G \times v_S(t) = -\frac{R_F}{R_S} \times v_S(t) = -10 \times 0.015\cos(\omega t) = -0.15\cos(\omega t)$$

The input and output waveforms are sketched in Figure 6.28.

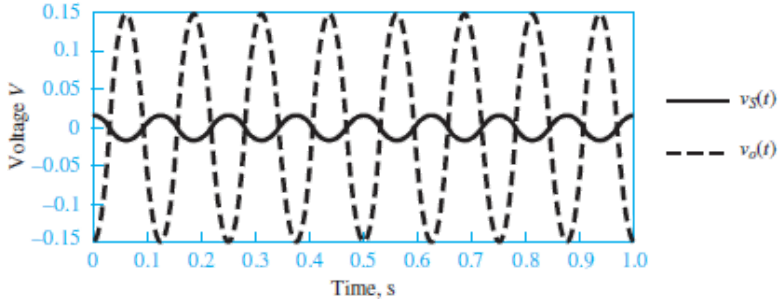


Figure 6.28

The nominal gain of the amplifier is $G_{\text{nom}} = -10$. If 5 percent tolerance resistors are employed, the worst-case error will occur at the extremes:

$$G_{\min} = -\frac{R_{F\min}}{R_{S\max}} = -\frac{9{,}500}{1{,}050} = 9.05 \qquad G_{\max} = -\frac{R_{F\max}}{R_{S\min}} = -\frac{10{,}500}{950} = 11.05$$

The percentage error is therefore computed as

$$100 \times \frac{G_{\text{nom}} - G_{\min}}{G_{\text{nom}}} = 100 \times \frac{10 - 9.05}{10} = 9.5\%$$

$$100 \times \frac{G_{\text{nom}} - G_{\max}}{G_{\text{nom}}} = 100 \times \frac{10 - 11.05}{10} = -10.5\%$$

Thus, the amplifier gain could vary by as much as ±10 percent (approximately) when 5 percent resistors are used. If 10 percent resistors were used, calculate a percent error of approximately ±20 percent, as shown below.

$$G_{\min} = -\frac{R_{F\min}}{R_{S\max}} = -\frac{9{,}000}{1{,}100} = 8.18 \qquad G_{\max} = -\frac{R_{F\max}}{R_{S\min}} = -\frac{11{,}000}{900} = 12.2$$

$$100 \times \frac{G_{\text{nom}} - G_{\min}}{G_{\text{nom}}} = 100 \times \frac{10 - 8.18}{10} = 18.2\%$$

$$100 \times \frac{G_{\text{nom}} - G_{\max}}{G_{\text{nom}}} = 100 \times \frac{10 - 12.2}{10} = -22.2\%$$

Comments: Note that the worst-case percent error in the closed-loop gain G is approximately double the resistor tolerance. This result can be calculated by assuming a resistor tolerance x and noting that the worst case is

$$|\Delta G| = \frac{R_F(1+x)}{R_S(1-x)} - \frac{R_F}{R_S}$$

Let $G_{\text{nom}} = -R_F/R_S$ such that:

$$\frac{|\Delta G|}{|G_{\text{nom}}|} = \frac{1+x}{1-x} - 1 = \frac{2x}{1-x}$$

$$= 2x(1 + x + x^2 + \cdots) \approx 2x \qquad (x \ll 1)$$

EXAMPLE 6.2 Instrumentation Amplifier

Problem

Determine the closed-loop voltage gain of the instrumentation amplifier circuit of Figure 6.29.

Solution

Known Quantities: Feedback and source resistances.

Find:

$$G = \frac{v_o}{v_1 - v_2}$$

Assumptions: Assume ideal op-amps.

Page 417

Analysis: Separate high impedance input stages are often used to isolate the sensor input signals v_1 and v_2 from the finite input impedance of a difference amplifier stage. This technique is present in the **instrumentation amplifier (IA)** circuit shown in Figure 6.29.

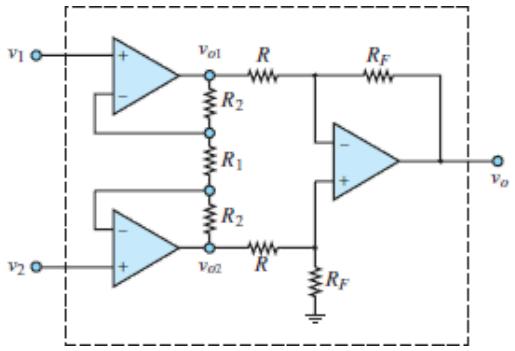


Figure 6.29 Instrumentation amplifier

Because the instrumentation amplifier has widespread application—and to ensure the best possible match between resistors—the entire circuit of Figure 6.29 is often packaged as a single integrated circuit. The advantage is that resistors R, R_F and R_2 can be matched much more precisely in an integrated circuit than would be possible using discrete components.

The principle of superposition can be applied to solve for the output voltage of the instrumentation amplifier. First, turn off the voltage source v_1 (set $v_1 = 0$)

such that the upper and lower parts of the left-half of Figure 6.29 are equivalent to the circuits shown in Figure 6.30(a) and (b), respectively.

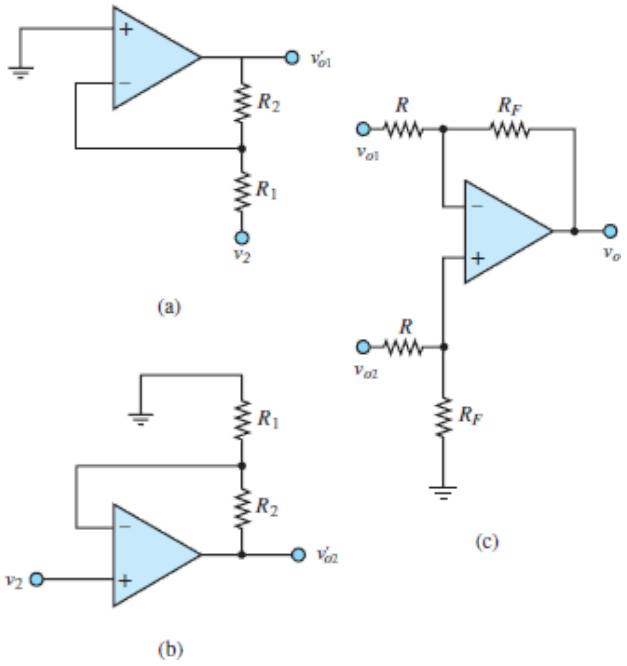


Figure 6.30 The principle of superposition can be applied to the upper (a) and lower (b) input stages of the instrumentation amplifier. The output stage of the instrumentation amplifier (c) is a difference amplifier

Page 418

From the perspective of v_2, the circuit in Figure 6.30(a) is an inverting amplifier (see Figure 6.8) and the circuit in Figure 6.30(b) is a noninverting amplifier (see Figure 6.9). The result is that the output voltages v'_{o1} and v'_{o2} are

$$v'_{o1} = -\frac{R_2}{R_1}v_2 \qquad v'_{o2} = \left(1+\frac{R_2}{R_1}\right)v_2$$

Second, turn off the voltage source v_2 (set $v_2 = 0$) and turn v_1 back on. Due to the symmetry of the left-half of the instrumentation amplifier the result is that the output voltages v''_{o1} and v''_{o2} are

$$v''_{o1} = \left(1+\frac{R_2}{R_1}\right)v_1 \qquad v'_{o2} = -\frac{R_2}{R_1}v_1$$

The results of summing the component voltages are

$$v_{o1} = \left(1 + \frac{R_2}{R_1}\right) v_1 - \frac{R_2}{R_1} v_2$$

and

$$v_{o2} = \left(1 + \frac{R_2}{R_1}\right) v_2 - \frac{R_2}{R_1} v_1$$

These voltages are the inputs to the right-half of the instrumentation amplifier as shown in Figure 6.30(c). That circuit is a difference amplifier (see Figure 6.18). The output of a difference amplifier is given by equation 6.59. Thus, the overall output of the instrumentation amplifier is

$$v_o = \frac{R_F}{R}(v_{o2} - v_{o1}) = \frac{R_F}{R}\left(1 + \frac{2R_2}{R_1}\right)(v_2 - v_1)$$

The closed-loop voltage gain of the instrumentation amplifier is shown below. The overall gain is the product of the gain of the input and difference stages, which are $1+(2R_2/R_1)$ and R_F/R, respectively.

$$G = \frac{v_o}{v_2 - v_1} = \frac{R_F}{R}\left(1 + \frac{2R_2}{R_1}\right) \qquad \text{Instrumentation amplifier}$$

EXAMPLE 6.3 Level Shifter

Problem

The level shifter of Figure 6.31 has the ability to add or subtract a DC offset to or from a signal. Analyze the circuit, and design it so that it can remove a 1.8-V DC offset from a sensor signal.

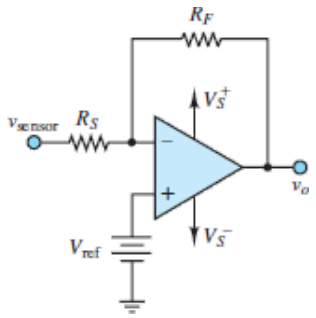


Figure 6.31 Level shifter

Solution

Known Quantities: Sensor (input) voltage; feedback and source resistors.

Find: Value of V_{ref} required to remove DC bias.

Schematics, Diagrams, Circuits, and Given Data: $v_{sensor}(t) = 1.8 + 0.1\cos(\omega t)$; $R_F = 220$ k Ω; $R_S = 10$ k Ω.

Page 419

Assumptions: Assume an ideal op-amp.

Analysis: The output voltage can be computed quite easily using the principle of superposition. When the reference voltage source V_{ref} is set to zero and replaced by a short-circuit, the sensor input voltage v_{sensor} sees an inverting amplifier such that:

$$\frac{v_{o_1}}{v_{sensor}} = -\frac{R_F}{R_S}$$

When the sensor input voltage source is set to zero and replaced by a short-circuit, the reference voltage source (the battery) sees a noninverting amplifier such that:

$$\frac{v_{o_2}}{V_{ref}} = 1 + \frac{R_F}{R_S}$$

Thus, the total output voltage is the sum of contributions from the two sources:

$$v_o = v_{o_1} + v_{o_2} = -\frac{R_F}{R_S} v_{sensor} + \left(1 + \frac{R_F}{R_S}\right) V_{ref}$$

Substitute the expression for v_{sensor} into the previous equation to find:

$$v_o = -\frac{R_F}{R_S}[1.8 + 0.1\cos(\omega t)] + \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}}$$

$$= -\frac{R_F}{R_S}[0.1\cos(\omega t)] - \frac{R_F}{R_S}(1.8) + \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}}$$

To remove the DC offset, require:

$$-\frac{R_F}{R_S}(1.8) + \left(1 + \frac{R_F}{R_S}\right) V_{\text{ref}} = 0$$

or

$$V_{\text{ref}} = 1.8\frac{R_F/R_S}{1 + R_F/R_S} = 1.722\,\text{V}$$

Comments: The presence of a precision voltage source in the circuit is undesirable because it may add considerable expense to the circuit design and, in the case of a battery, it is not adjustable. The circuit of Figure 6.32 illustrates how an adjustable voltage reference can be produced from the DC supplies already used by the op-amp, two fixed resistors R, and a potentiometer R_p. The fixed resistors are included to guarantee a minimum resistance R from the wiper to either power supply at all times and thus prevent possible overheating of the potentiometer. An expression for V_{ref} is obtained from voltage division:

$$\frac{V_{\text{ref}} - V_S^-}{V_S^+ - V_S^-} = \frac{R + \Delta R}{2R + R_p}$$

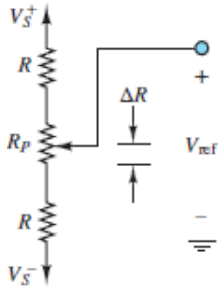


Figure 6.32

If the voltage supplies are symmetric, as is usually the case, $V_S^+ = -V_S^-$ such that:

$$\frac{V_{\text{ref}} + V_S^+}{2V_S^+} = \frac{R + \Delta R}{2R + R_p}$$

Rearrange terms to find:

$$V_{\text{ref}} = \frac{2\Delta R - R_p}{2R + R_p} V_S^+$$

The value of V_{ref} is determined by the position of the wiper ΔR. Also, when $R_p \gg R$, the range of V_{ref} is approximately $\pm V_S^+$.

EXAMPLE 6.4 Temperature Control Using Op-Amps

Problem

Op-amps often serve as building blocks in analog control systems. The objective of this example is to illustrate the use of op-amps in a temperature control circuit. Figure 6.33(a) depicts a system for which the temperature is to be maintained constant at 20°C in a variable temperature environment. The system temperature is measured via a thermocouple (see the "Temperature Measurement" section in Chapter 7). Heat is added to the system by a coil of resistance R_{coil}. The heat flux is $q_{\text{in}} = i^2 R_{\text{coil}}$, where i is the current provided by a power amplifier. The system is insulated on three sides. The fourth side is not insulated such that heat is transferred across the boundary by convection, which is represented by an equivalent thermal resistance R_t. The system has mass m, specific heat c, and thermal capacitance $C_t = mc$ (see the Make the Connection boxes "Thermal Capacitance" and "Thermal System Dynamics" in Chapter 4).

Solution

Known Quantities: Sensor (input) voltage; feedback and source resistors, thermal system component values.

Find: Select desired value of proportional gain K_P to achieve automatic temperature control.

Schematics, Diagrams, Circuits, and Given Data: $R_{coil} = 5\ \Omega$; $R_t = 2°C/W$; $C_t = 50\ J/°C$; $\alpha = 1\ V/°C$. Figure 6.33(a) to (e).

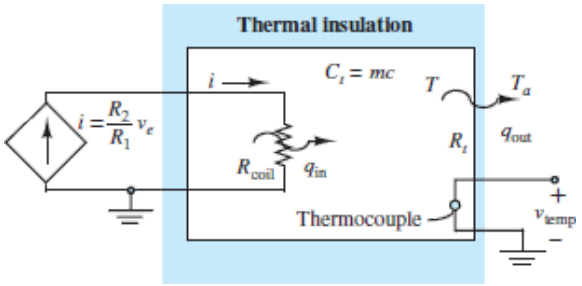


Figure 6.33(a) Thermal system

Assumptions: Assume ideal op-amps.

Analysis: Conservation of energy requires that:

$$q_{in} - q_{out} = \frac{dE_{stored}}{dt}$$

Page 421

where q_{in} represents the heat added to the system by the electrical heater, q_{out} represents the heat lost from the system through convection to the surrounding air, and E_{stored} represents the energy stored in the system due to its thermal capacitance. The system temperature T is measured by a thermocouple whose output voltage is proportional to temperature: $v_{temp} = \alpha T$. Further, assume that the power amplifier is modeled by a *voltage-controlled current source* (VCCS) such that:

$$i = K_v K_p v_e = \frac{R_2}{R_1} v_e = \frac{R_2}{R_1}(v_{ref} - v_{temp}) = \frac{R_2}{R_1}\alpha(T_{ref} - T)$$

where v_e is the error or difference between the reference voltage and the measured voltage. The negative feedback system shown in Figure 6.33(b) tends to drive v_e to zero. When v_e is positive, $v_{ref} > v_{temp}$ and the system calls for heating; on the other hand, when v_e is negative, $v_{ref} < v_{temp}$ and the system calls for cooling. The power amplifier outputs a positive current for a positive v_e. Thus, the block diagram shown in Figure 6.33(b) corresponds to an *automatic control system* that increases or decreases the heating coil current to maintain the system temperature at the desired (reference) value. The *proportional gains* K_v of the

voltage amplifier and K_p of the power amplifier determine the increase in coil current and can be used to optimize the response of the system for a specific design requirement. For example, a system specification could require that the automatic temperature control system be designed so as to maintain the temperature to within 1 degree of the reference temperature for external temperature disturbances as large as 10 degrees. The response of the system can be adjusted by varying the proportional gain.

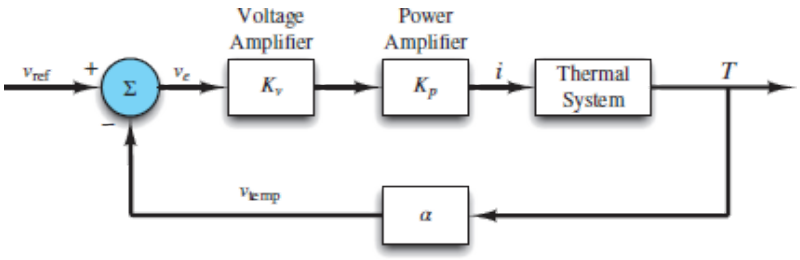


Figure 6.33(b) Block diagram of control system

The voltage amplifier can be realized by a two-stage amplifier using two op-amps as shown in Figure 6.33(c). The first stage is an inverting amplifier with closed-loop gain $G_1 = -1$ such that the voltage at node a is $-v_{\text{ref}}$. The second stage is a summing amplifier with a closed-loop gain of $G_2 = -R_2/R_1$ for each input. Thus, the output voltage at node b is

$$v_b = -\frac{R_2}{R_1}(v_a + v_{\text{temp}}) = \frac{R_2}{R_1}(v_{\text{ref}} - v_{\text{temp}}) = \frac{R_2}{R_1}(v_e)$$

The coefficient R_2/R_1 is the voltage gain K_v. In other words, *selecting the feedback resistor R_2 is equivalent to choosing K_v.*

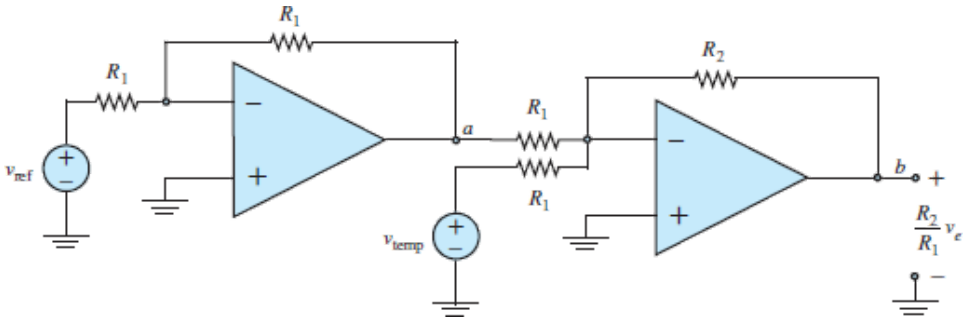


Figure 6.33(c) Circuit for generating proportional gain of error voltage

The thermal system itself is described by the conservation of energy equation given above. The rate of energy added to the system by the heating coil is simply i^2R_{coil}. The rate of energy subtracted from the system by convective heat transfer is defined as $(T - T_a)/R_t$, where R_t is a lumped parameter called the *thermal resistance*. Small values of R_t correspond Page 422to large values of the convective heat transfer coefficient, and vice versa. Finally, the net rate at which energy is stored in the system is proportional to the rate at which the system temperature T is changing, where the constant of proportionality C_t is known as the *thermal capacitance*. With these definitions in place, the conservation of energy equation can be rewritten as:

$$i^2R_{\text{coil}} - \frac{T - T_a}{R_t} = C_t\frac{dT}{dt}$$

or

$$R_tC_t\frac{dT}{dt} + T = R_tR_{\text{coil}}i^2 + T_a$$

where $i = K_pK_v\upsilon_e = K_pK_v\alpha(T_{\text{ref}} - T)$. This first-order ordinary differential equation is nonlinear. The time constant is $\tau = R_tC_t = 2°\text{C/W} \times 50\ \text{J/°C} = 100\ \text{s}$.

When $K_p = 0$, no current is supplied to the heating coil and the thermal system response is simply its own natural response; that is, no automatic control is active when $K_p = 0$ and the system response is the *open-loop* response. In that case, the governing differential equation is

$$\tau\frac{dT}{dt} + T = T_a$$

The solution is (see Chapter 4)

$$T = (T_0 - T_a)e^{-t/\tau} + T_a$$

where T_0 is the initial value of the system temperature. For the case when $T_0 = 20°\text{C}$ and $T_a = 10°\text{C}$, the solution is

$$T = (10°\text{C})e^{-t/100} + 10°\text{C} \qquad \text{Open-loop response}$$

When the gain K_p is increased to 1, υ_e increases as soon as the temperature drops below the reference value. The transduction constant of the thermocouple was given as $\alpha = 1$ such that the voltage υ_{temp} is numerically equal to the system

temperature. Figure 6.33(d) shows the temperature response for values of K_p ranging from 1 to 10. As the gain increases, the error between the desired and actual temperatures decreases very quickly. Observe that the error is less than 1 degree (recall the design specification) for K_p = 5. To better understand the workings of the complete control system, it is helpful to observe the heater current, which is an amplified version of the error voltage. Figure 6.33(e) shows that when K_p = 1 the current increases to a final value of roughly 2.7 A; when K_p = 5 and 10, the current increases more rapidly, and eventually settles to values of 3 and 3.1 A, respectively. The steady-state value of the current is reached in about 17 s for K_p = 5, and in about 8 s for K_p = 10.

Page 423

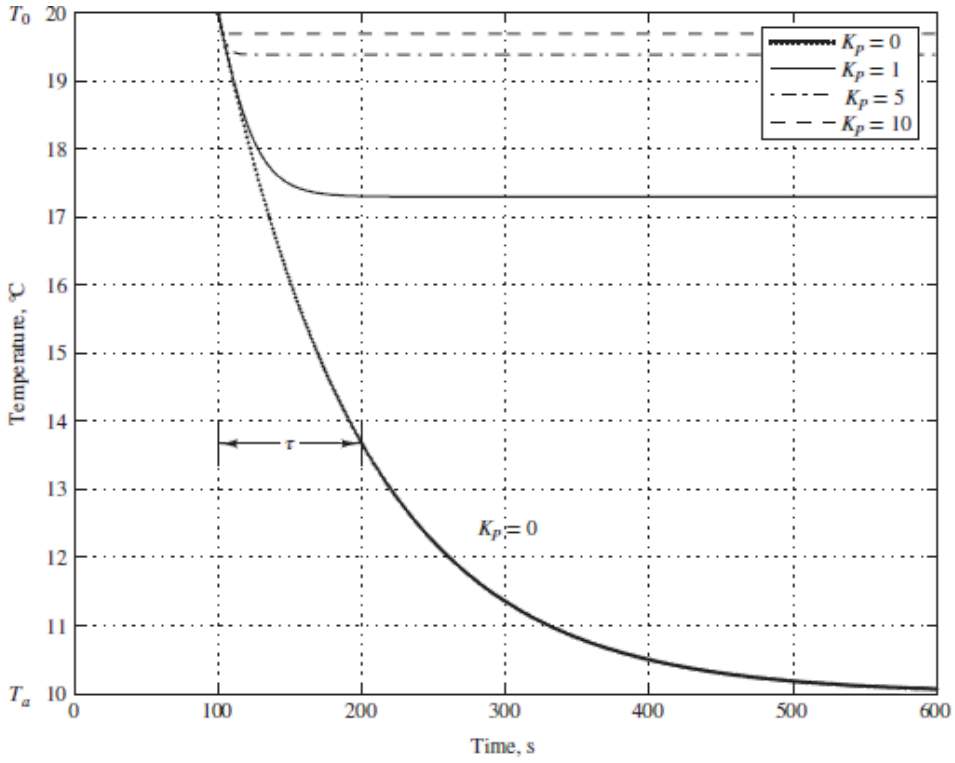


Figure 6.33(d) Response of thermal system for various values of proportional gain K_p

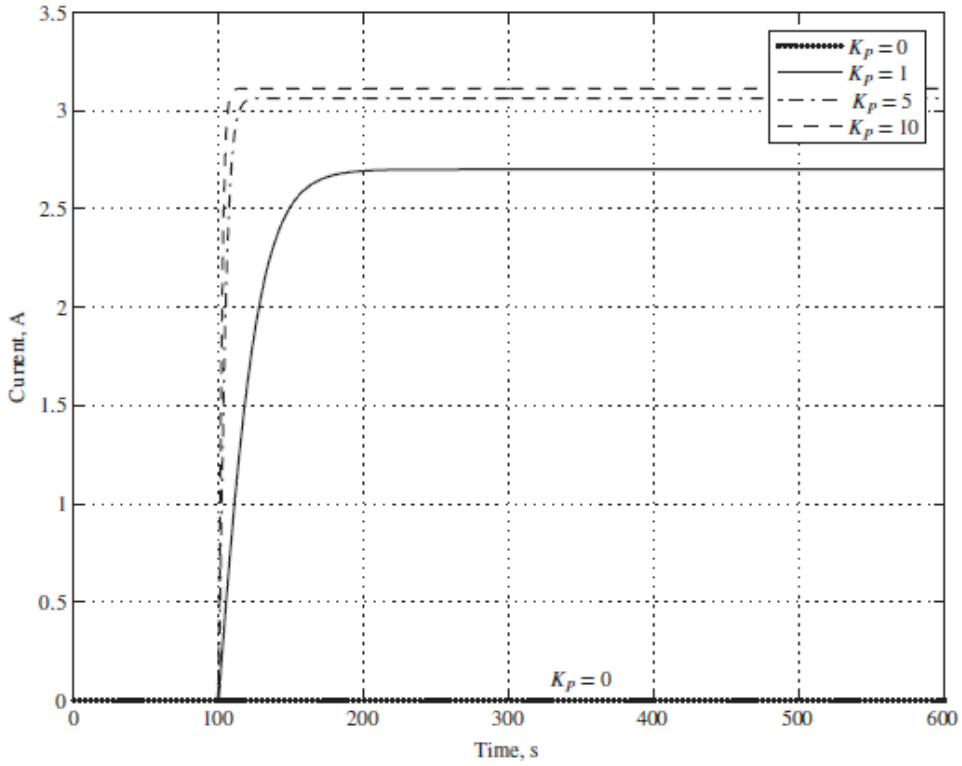


Figure 6.33(e) Power amplifier output current for various proportional gain K_p

Page 424

Comments: As K_p increases, the system's speed of response increases; however, the system's *steady-state error* also increases. The design specifications anticipate this effect by setting a tolerance of 1°C.

CHECK YOUR UNDERSTANDING

Consider an ideal inverting amplifier (see Figure 6.8) with a nominal closed-loop gain of −1,000. The impact of a nonideal op-amp with a finite, but large, open-loop gain A on the closed-loop gain can be derived by assuming that the voltage v^- at the inverting terminal is only approximately equal to the voltage $v^+ = 0$ at the noninverting terminal. Under this assumption, $v_{out} = -Av^-$. The first golden rule still applies such that $i_{in} = 0$ and R_S is virtually in series with R_F. Use this information to find an expression for the closed-loop gain as a function of the open-loop gain A. Compute the closed-loop gain when A equals 10^7, 10^6, 10^5,

10^4. How large is the open-loop gain when the closed-loop gain is less than 0.1 percent away from its nominal value?

Answer: 999.1; 999.0; 990.1; 909.1. For 0.1 percent accuracy, $A = 10^6$.

CHECK YOUR UNDERSTANDING

For Example 6.1, calculate the uncertainty in the gain when 1 percent "precision" resistors are used.

Answer: +1.98 to −2.02 percent

CHECK YOUR UNDERSTANDING

Derive an expression for the closed-loop gain of an isolation buffer when the open-loop gain A is finite. How large is the open-loop gain when the closed-loop gain is only 0.1 percent away from unity?

Answer: The expression for the closed-loop gain is $v_{out}/v_{in} = 1 + 1/A$; thus A should equal 10^4 for 0.1 percent accuracy.

Page 425

CHECK YOUR UNDERSTANDING

For Example 6.3, find the value ΔR that removes the DC bias from the sensor signal. Assume the supply voltages are symmetric at ±15 V and a 10-kΩ

potentiometer is tied to two 10-kΩ fixed resistors as in Figure 6.32. What is the range of V_{ref} when a 1-kΩ potentiometer is tied to two 10-kΩ fixed resistors?

Answer: $\Delta R = 6.722$ kΩ; V_{ref} is between ± 0.714 V

CHECK YOUR UNDERSTANDING

How much steady-state power, in watts, will be input to the thermal system of Example 6.4 to maintain its temperature in the face of a 10°C ambient temperature drop for values of K_P of 1, 5, and 10?

Answer: $K_P = 1$: 36.5 W; $K_P = 5$: 45 W; $K_P = 10$: 48 W

CHECK YOUR UNDERSTANDING

With reference to the Focus on Measurements box, "Sensor Calibration Circuit," find numerical values of R_F/R_S and V_{ref} if the temperature sensor has $\beta = 0.235$ and $V_0 = 0.7$ V and the desired relationship is $v_o = 10\ T$.

Answer: $R_F/R_S = 42.55$; $V_{\text{ref}} = 0.684$ V

6.3 ACTIVE FILTERS

The range of useful applications of an operational amplifier is greatly expanded if energy storage elements are introduced into the design; the frequency-dependent properties of these elements, studied in Chapters 3 and 5, will prove useful in the design of various types of op-amp circuits. In particular, it will be shown that it is

possible to shape the frequency response of an amplifier by appropriate use of complex impedances in the input and feedback paths. The class of filters one can obtain by means of op-amp designs is called **active filters** because op-amps can provide amplification (gain) in addition to the filtering effects already studied in Chapter 5 for passive circuits (i.e., circuits comprised of only resistors, capacitors, and inductors).

The easiest way to see how the frequency response of an op-amp can be shaped (almost) arbitrarily is to replace the resistors R_F and R_S in Figures 6.8 and 6.9 with impedances $\mathbf{Z}_F$ and $\mathbf{Z}_S$, as shown in Figure 6.34. It is a straightforward Page 426matter to show that in the case of the inverting amplifier, the expression for the closed-loop gain is given by

$$\frac{\mathbf{V}_o}{\mathbf{V}_S}(j\omega) = -\frac{\mathbf{Z}_F}{\mathbf{Z}_S} \tag{6.71}$$

whereas for the noninverting case, the gain is

$$\frac{\mathbf{V}_o}{\mathbf{V}_S}(j\omega) = 1 + \frac{\mathbf{Z}_F}{\mathbf{Z}_S} \tag{6.72}$$

where $\mathbf{Z}_F$ and $\mathbf{Z}_S$ can be arbitrarily complex impedance functions and where $\mathbf{V}_S$, $\mathbf{V}_o$, $\mathbf{I}_F$, and $\mathbf{I}_S$ are all phasors. Thus, it is possible to shape the frequency response of an active filter simply by selecting suitable ratios of feedback impedance to input impedance. By connecting a circuit similar to the low-pass filters studied in Chapter 5 in the feedback loop of an op-amp, the same filtering effect can be achieved and, in addition, the signal can be amplified.

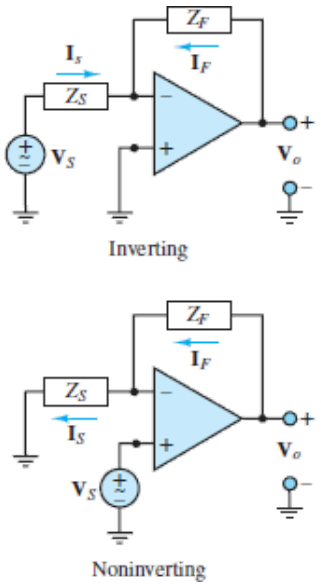


Figure 6.34 Op-amp circuits employing complex impedances

The simplest active low-pass filter is shown in Figure 6.35. The closed-loop gain, as a function of frequency, is given by

$$\mathbf{G}_{\mathrm{LP}}(j\omega) = -\frac{\mathbf{Z}_F}{\mathbf{Z}_S} \tag{6.73}$$

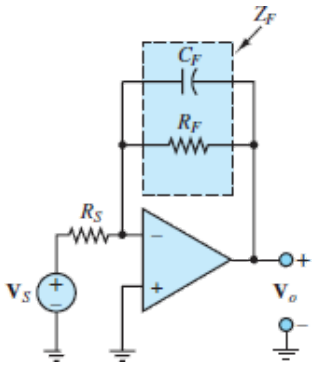


Figure 6.35 Active low-pass filter

where

$$\mathbf{Z}_F = R_F \| \frac{1}{j\omega C_F} = \frac{R_F}{1 + j\omega C_F R_F} \tag{6.74}$$

and

$$\mathbf{Z}_S = R_S \tag{6.75}$$

Note the similarity between $\mathbf{Z}_F$ and the low-pass characteristic of a passive *RC* low-pass filter. The closed-loop gain $\mathbf{G}_{\text{LP}}(j\omega)$ is then computed to be

$$\mathbf{G}_{\text{LP}}(j\omega) = -\frac{\mathbf{Z}_F}{\mathbf{Z}_S} = -\frac{R_F/R_S}{1 + j\omega C_F R_F} \qquad \text{Low-pass filter} \tag{6.76}$$

This expression can be factored into two terms. The first is an amplification factor equivalent to the amplification that would be obtained with a simple inverting amplifier (i.e., the same circuit as that of Figure 6.35 with the capacitor removed); the second is a low-pass filter, with a cutoff frequency dictated by the parallel combination of R_F and C_F in the feedback loop. The filtering effect is equivalent to what would be attained by the passive circuit shown in Figure 6.36.

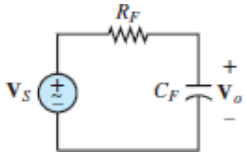


Figure 6.36 Passive low-pass filter

The response of this op-amp filter is an amplified version of that of a passive low-pass *RC* filter. Figure 6.37 depicts the amplitude response of the active low-pass filter (in the figure, $R_F/R_S = 10$ and $R_F C_F = 1$) in two different graphs: The first plots the amplitude ratio while the second plots the Bode magnitude (in dB), both versus ω on a logarithmic scale. The cutoff frequency ω_0 is

$$\omega_0 = \frac{1}{R_F C_F} \tag{6.77}$$

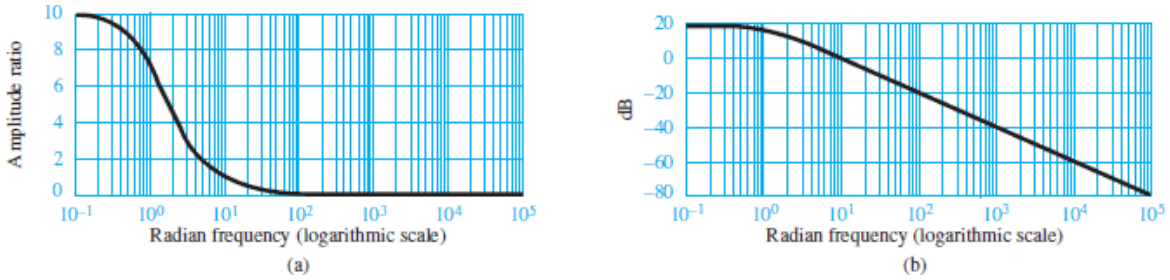


Figure 6.37 Normalized response of active low-pass filter (a) amplitude ratio response; (b) Bode magnitude (dB) response

and the slope of the Bode magnitude response is −20 dB/decade when $\omega \gg \omega_0$. The magnitude (in dB) at the cutoff frequency is

$$|\mathbf{G}_{\text{LP}}(j\omega_0)|_{\text{dB}} = 20\log_{10}\frac{R_F}{R_S} - 20\log\sqrt{2} \tag{6.78}$$

where

$$-20\log_{10}\sqrt{2} = -3 \text{ dB} \tag{6.79}$$

Thus, ω_0 is also called the *3-dB frequency.* It is also known as the *break frequency* or *cutoff frequency*.

Among the advantages of such active low-pass filters is the ease with which the gain and the bandwidth are dictated by choosing R_F, R_S and C_F.

It is also possible to arrange resistors and capacitors to produce other types of filters. For example, the circuit shown in Figure 6.38 is an active high-pass filter. The input impedance is

$$\mathbf{Z}_S = R_S + \frac{1}{j\omega C_S} \tag{6.80}$$

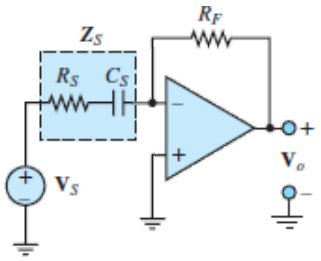


Figure 6.38 Active high-pass filter

The impedance of the feedback path is

$$\mathbf{Z}_F = R_F \tag{6.81}$$

The closed-loop gain for this inverting amplifier is

$$\mathbf{G}_{\text{HP}}(j\omega) = -\frac{\mathbf{Z}_F}{\mathbf{Z}_S} = -\frac{j\omega C_S R_F}{1 + j\omega R_S C_S} \qquad \text{High-pass filter} \tag{6.82}$$

Note that $G \to 0$ as $\omega \to 0$. Also note that as $\omega \to \infty$, the closed-loop gain G approaches a constant.

$$\lim_{\omega \to \infty} \mathbf{G}_{\text{HP}}(j\omega) = -\frac{R_F}{R_S} \tag{6.83}$$

That is, above a certain frequency range, the circuit acts as a linear amplifier. This is exactly the behavior one would expect of a high-pass filter. The high-pass response is depicted in Figure 6.39, in both amplitude and Bode magnitude plots (in the figure, $R_F/R_S = 10$ and $R_S C_S = 1$). The slope of the Bode magnitude plot is +20 dB/decade when $\omega \ll \omega_0$, where $\omega_0 = 1/R_S C_S$ is the 3-dB break frequency.

Page 428

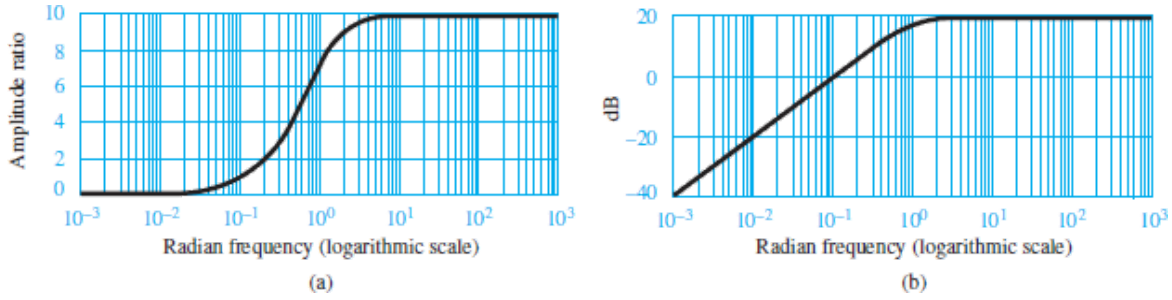


Figure 6.39 Normalized response of active high-pass filter (a) amplitude ratio response; (b) dB response

As a final example of active filters, a basic active bandpass filter configuration can be realized by combining the elements of active high- and low-pass filters. The circuit is shown in Figure 6.40.

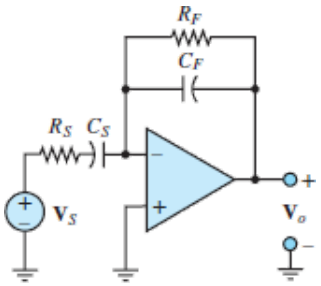


Figure 6.40 Active bandpass filter

The analysis of the bandpass circuit follows the same approach used in previous examples. The feedback and input impedances are

$$\mathbf{Z}_F = R_F \| \frac{1}{j\omega C_F} = \frac{R_F}{1 + j\omega C_F R_F} \tag{6.84}$$

$$\mathbf{Z}_S = R_S + \frac{1}{j\omega C_S} = \frac{1 + j\omega C_S R_S}{j\omega C_S} \tag{6.85}$$

The closed-loop frequency response is

LO

$$\mathbf{G}_{BP}(j\omega) = -\frac{\mathbf{Z}_F}{\mathbf{Z}_S} = -\frac{j\omega C_S R_F}{(1 + j\omega C_F R_F)(1 + j\omega C_S R_S)} \quad \text{Bandpass filter} \tag{6.86}$$

The form of the response is the same as that of the series *LC* bandpass filter shown in Figure 5.32 when $\zeta > 1$. See equation 5.55. This response is similar (although not identical) to the product of the low-pass and high-pass responses of equations 6.76 and 6.82. In particular, the denominator of $\mathbf{G}_{BP}(j\omega)$ is exactly the product of the denominators of $\mathbf{G}_{LP}(j\omega)$ and $\mathbf{G}_{HP}(j\omega)$. It is particularly enlightening to rewrite $\mathbf{G}_{LP}(j\omega)$ in a slightly different form, after making the observation that each *RC* product corresponds to one of the following critical frequencies.

$$\omega_0 = \frac{1}{R_F C_S} \qquad \omega_{LP} = \frac{1}{R_F C_F} \qquad \omega_{HP} = \frac{1}{R_S C_S} \tag{6.87}$$

It is easy to verify that for the case where

$$\omega_{HP} \gg \omega_{LP} \tag{6.88}$$

the response of the filter will be similarly shaped to that shown in Figure 6.41 in both amplitude and Bode magnitude plots. (In the figures, $\omega_0 = 1$, $\omega_{HP} = 1{,}000$, and $\omega_{LP} = 10$ radians/second.) Compare Figure 6.41 (a) to Figure 5.33 for the overdamped series LC bandpass filter. The Bode magnitude plot shows that, in effect, the bandpass response is the superposition of active low- and high-pass responses. The two 3-dB (or cutoff) frequencies are the same as in $\mathbf{G}_{LP}(j\omega)$, $1/R_FC_F$; and in $\mathbf{G}_{HP}(j\omega)$, $1/R_SC_S$. The third frequency, $\omega_0 = 1/R_FC_S$, represents the point where the response of the filter crosses the 0-dB axis (rising slope). Since 0 dB corresponds to a gain of 1, this frequency is called the **unity-gain frequency**.

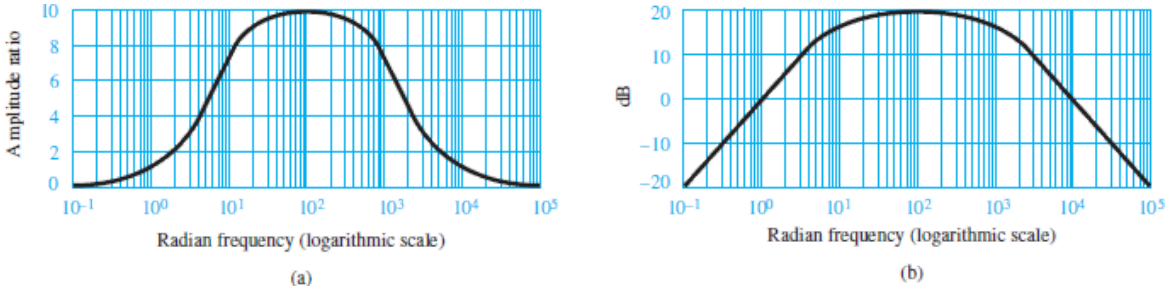


Figure 6.41 Normalized amplitude response of active bandpass filter (a) amplitude ratio response; (b) dB response

Page 429

The ideas developed thus far can be employed to construct more complex functions of frequency. In fact, most practical active filters are based on circuits of at least two capacitors. By constructing suitable functions for $\mathbf{Z}_F$ and $\mathbf{Z}_S$, it is possible to realize filters with greater frequency selectivity (i.e., sharpness of cutoff), as well as flatter bandpass or band-rejection functions (i.e., filters that either allow or reject signals in a limited band of frequencies). A few simple applications are investigated in the homework problems; advanced applications are explored in Section 6.4.

One of the advantages of these active filters is that it is possible to produce any frequency response using only capacitors. No inductors are needed. This seemingly minor fact is of great importance in practice because inductors are expensive to mass-produce to small tolerances and exact specifications and are often bulkier than capacitors with equivalent energy storage capabilities. On the other hand, capacitors are easy to manufacture in a wide variety of tolerances and

values, and in relatively compact packages, including in integrated-circuit form. Inductors are, in general, also much more prone to noise.

CHECK YOUR UNDERSTANDING

(a) Design a low-pass filter with a closed-loop gain of 100 and cutoff (3-dB) frequency equal to 800 Hz. Assume that only 0.01-μF capacitors are available. Find R_F and R_S.

(b) Repeat the design of the exercise in part a for a high-pass filter with a cutoff frequency of 2,000 Hz. This time, however, assume that only standard values of resistors are available (see Table 1.3 in Chapter 1). Select the nearest component values, and calculate the percent error in cutoff frequency.

(c) Find the frequencies corresponding to 1-dB attenuation from the low-frequency gains of the filters of parts a and b.

Answer: Part a: $R_F = 19.9$ kΩ, $R_S = 199$ Ω; part b: $R_F = 8.2$ kΩ, $R_S =$ 82 Ω, error: gain = 0 percent, $\omega_{3dB} = 2.95$ percent; part c: 407 Hz and 3.8 kHz

6.4 DESIGN OF ACTIVE FILTERS

The need to filter sensor signals that may be corrupted by noise or other interfering or undesired inputs has already been approached in an earlier chapter. In Chapter 5, passive filters made of resistors, capacitors, and inductors were Page 430analyzed. It was shown that three types of filter frequency response characteristics can be achieved with these simple circuits: low pass, high pass, and bandpass. The properties of operational amplifiers can be exploited to simplify filter design, to more easily match source and load impedances, and to eliminate the need for inductors.

Figure 6.42 depicts the general characteristics of a low-pass active filter, indicating that within the passband of the filter, a certain deviation from the nominal filter gain A is accepted, as indicated by the minimum and maximum passband gains $A + \varepsilon$ and $A - \varepsilon$. The width of the passband is indicated by the cutoff frequency ω_C. On the other hand, the stopband, starting at the frequency ω_S, does not allow a gain greater than $A_{\min}$. Different types of filter designs

achieve different types of frequency responses, which are typically characterized by having a particularly flat passband frequency response **(Butterworth filters)** or by a very rapid transition between passband and stopband (**Chebyshev filters**, and **Cauer**, or **elliptical**, **filters**), or by some other characteristic, such as a linear phase response **(Bessel filters)**. Achieving each of these properties usually involves tradeoffs; for example, a very flat passband response will usually result in a relatively slow transition from passband to stopband.

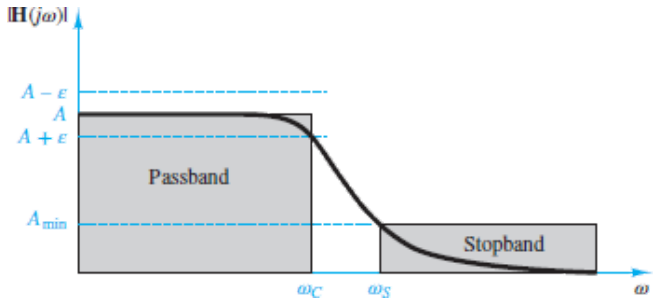


Figure 6.42 Prototype low-pass filter response

In addition to selecting a filter from a certain family, it is possible to select the *order* of the filter. In general, the higher the order, the faster the transition from passband to stopband (at the cost of greater phase shifts and amplitude distortion, however). Although the frequency response of Figure 6.42 pertains to a low-pass filter, similar definitions also apply to the other types of filters.

Butterworth filters are characterized by a *maximally flat* passband frequency response characteristic; their response is defined by a magnitude-squared function of frequency

$$|\mathbf{H}(j\omega)|^2 = \frac{H_0^2}{1 + \varepsilon^2 \omega^{2n}} \tag{6.89}$$

where $\varepsilon = 1$ for maximally flat response and n is the order of the filter. Figure 6.43 depicts the frequency response (normalized to $\omega_C = 1$) of first-, second-, third-, and fourth-order Butterworth low-pass filters. The **Butterworth polynomials**, given in Table 6.2 in factored form, permit the design of the filter by specifying the denominator as a polynomial in s. For $s = j^{\omega}$, one obtains the frequency response of the filter. Examples 6.5 and 6.6 illustrate filter design procedures that make use of these tables.

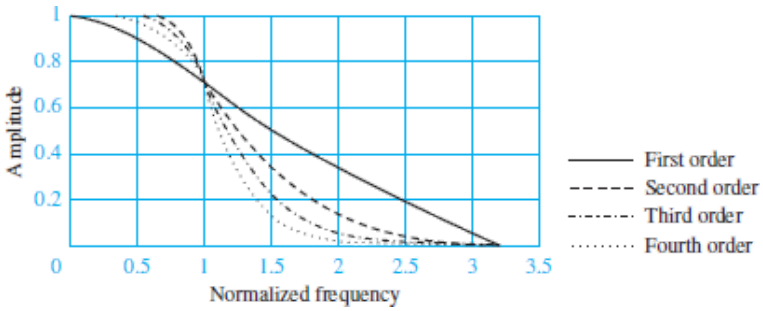


Figure 6.43 Butterworth low-pass filter frequency response

Table 6.2 Butterworth polynomials in quadratic form

Order n	Quadratic factors
1	$s + 1$
2	$s^2 + \sqrt{2}s + 1$
3	$(s + 1)(s^2 + s + 1)$
4	$(s^2 + 0.7654s + 1)(s^2 + 1.8478s + 1)$
5	$(s + 1)(s^2 + 0.6180s + 1)(s^2 + 1.6180s + 1)$

Figure 6.44 depicts the normalized frequency response of first- to fourth-order low-pass Chebyshev filters (n = 1 to 4), for ε = 1.06. Note that a certain amount of ripple is allowed in the passband; the amplitude of the ripple is defined by the parameter ε and is constant throughout the passband. Thus, these filters are also called **equiripple filters**. Cauer, or elliptical, filters are similar to Chebyshev filters, except for being characterized by equiripple both in the passband and in the stopband. Design tables are available to select the appropriate order of Butterworth, Chebyshev, or Cauer filter for various applications.

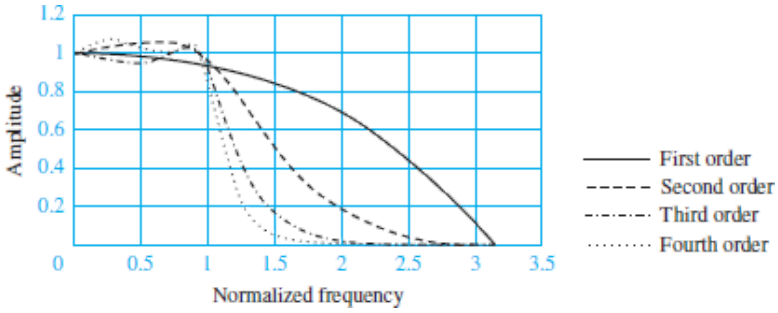


Figure 6.44 Chebyshev low-pass filter frequency response

Three common configurations of second-order active filters, which can be used to implement **second-order** (or **quadratic**) **filter sections** using a single op-amp, are shown in Figure 6.45. These filters are called **constant-*K***, or **Sallen and**

Key, **filters** (after the names of the inventors). The analysis of these active filters is based on the properties of the ideal operational amplifier.

Page 432

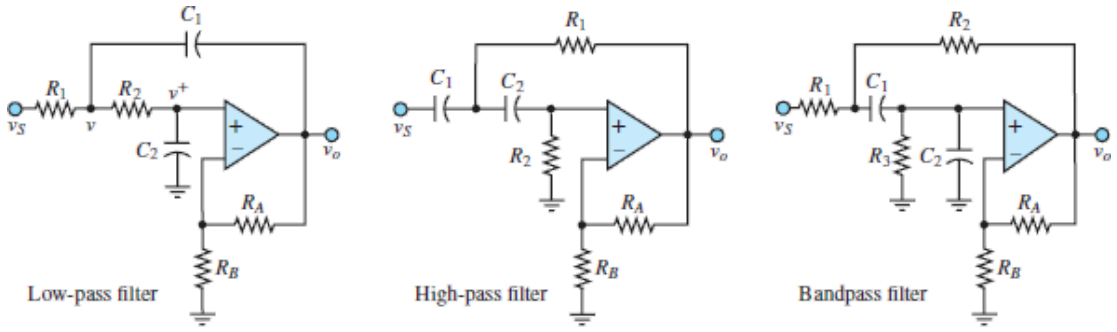


Figure 6.45 Sallen and Key active filters

Consider, for example, the low-pass filter of Figure 6.45, which is a non-inverting amplifier with **positive feedback**; that is, feedback paths are provided from the output to both the inverting and the noninverting terminals of the op-amp. The input-output relationship for the filter can be found from expressions for the input terminal voltages of the op-amp v^+ and v^-. The frequency response of this filter is

$$\mathbf{H}(j\omega) = \frac{K(1/R_1R_2C_1C_2)}{(j\omega)^2 + \left[\frac{1}{R_1C_1} + \frac{1}{R_2C_1} + \frac{1}{R_2C_2}(K-1)\right]j\omega + \frac{1}{R_1R_2C_1C_2}} \tag{6.90}$$

This frequency response can be expressed in either of two more general low-pass filter forms:

$$\mathbf{H}(j\omega) = \frac{K\omega_C^2}{(j\omega)^2 + (\omega_C/Q)(j\omega) + \omega_C^2}$$

or

$$\mathbf{H}(j\omega) = \frac{K}{(j\omega)^2/\omega_C^2 + (2\zeta/\omega_C)(j\omega) + 1} \tag{6.91}$$

These two forms are related by the identity $2\zeta\ Q = 1$, where ζ is the **dimensionless damping coefficient**, ω_C is the cutoff frequency, and Q is the **quality factor**, which represents the sharpness of the filter's **resonance**. A high-Q filter is **underdamped**, while a low-Q filter will be **overdamped**. A critically

damped circuit has $Q = 0.5$. Compare equation 6.91 to equations 5.19 and 5.24 for the voltage and current gains of series and parallel LC circuits, respectively. Apparently, it is possible to produce the same response found for those circuits but without using an inductor!

The relationships between the three parameters of the second-order filter (ω_C, ζ, and K) and the resistors and capacitors are defined below for the low-pass Sallen and Key filter. A very desirable property of this filter is the fact that its low-frequency gain K is independent of the cutoff frequency and is determined simply by the ratio of resistors R_A and R_B. The other four components define the cutoff frequency and damping coefficient as shown in equations 6.92.

$$K = 1 + \frac{R_A}{R_B} \tag{6.92}$$

$$\omega_C = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}}$$

$$\frac{1}{Q} = 2\zeta = \sqrt{\frac{R_2 C_2}{R_1 C_1}} + \sqrt{\frac{R_1 C_2}{R_2 C_1}} + (K-1)\sqrt{\frac{R_1 C_1}{R_2 C_2}}$$

The quadratic filters shown in Figure 6.45 can be staged to implement filters of higher order and of different characteristics. For example, a fourth-order Butterworth filter can be realized by cascading two second-order Sallen and Key quadratic filters and by observing that the component values of each filter can be specified given the desired gain, cutoff frequency, and damping coefficient.

EXAMPLE 6.5 Determining the Order of a Butterworth Filter

Problem

Determine the required order of a filter, given the filter specifications.

Solution

Known Quantities: Filter gain at cutoff frequencies (passband and stopband).

Find: Order n of filter.

Schematics, Diagrams, Circuits, and Given Data: Passband gain: −3 dB at ω_C = 1 rad/s; stopband gain: −40 dB at $\omega_S = 4\omega_C$.

Assumptions: Use a Butterworth filter response. Assume a low-frequency gain $H_0 = 1$.

Analysis: Using the magnitude-squared response for the Butterworth filter (equation 6.89),

$$|\mathbf{H}(j\omega)|^2 = \frac{H_0^2}{1 + \varepsilon^2 \omega^{2n}}$$

With $\varepsilon = 1$, the magnitude of the transfer function at the passband cutoff frequency ω_C is

$$|\mathbf{H}(j\omega = j\omega_C)| = \frac{H_0}{\sqrt{1 + \omega_C^{2n}}} = \frac{H_0}{\sqrt{1 + 1^{2n}}} = \frac{H_0}{\sqrt{2}}$$

This result satisfies the requirement for the passband gain (3 dB below the low-frequency gain) since

$$20 \log_{10} \frac{1}{\sqrt{2}} = -3 \text{ dB}$$

This result is a feature of Butterworth filters.

The stopband gain requirement is that the gain at frequencies above ω_S be less than −40 dB:

$$20 \log_{10} |\mathbf{H}(j\omega = j\omega_S)| = 20 \log_{10} \frac{H_0}{\sqrt{1 + \omega_S^{2n}}} = 20 \log_{10} \frac{H_0}{\sqrt{1 + 4^{2n}}} \leq -40$$

Thus,

$$20 \log_{10} H_0 - 20 \log_{10} \sqrt{1 + 4^{2n}} \leq -40$$

or

$$\begin{aligned} \log_{10}(1 + 4^{2n}) &\geq 4 \\ 1 + 4^{2n} &\geq 10^4 \\ 2n \log_{10} 4 &\geq \log_{10}(10^4 - 1) \end{aligned}$$

Solving the above inequality yields $n \leq 3.32$. Since n must be an integer, choose $n = 4$ such that the gain at the stopband frequency is

$$|\mathbf{H}(j\omega = j\omega_S)| = \frac{H_0}{\sqrt{1+\omega_S^{2n}}} = \frac{1}{\sqrt{1+4^8}} = -48.16 \text{ dB}$$

which is lower than the minimum desired gain of −40 dB.

Comment: Note that the −3-dB gain at the passband cutoff frequency is always satisfied in a Butterworth filter when $\varepsilon = 1$.

EXAMPLE 6.6 Design of Sallen and Key Filter

Problem

Determine the cutoff frequency, DC gain, and quality factor for the low-pass Sallen and Key filter of Figure 6.45.

Solution

Known Quantities: Filter resistor and capacitor values.

Find: K; ω_C; Q.

Schematics, Diagrams, Circuits, and Given Data: All resistors are 500 Ω; all capacitors are 2 μF.

Assumptions: None.

Analysis: Using the definitions given in equation 6.92, compute

$$K = 1 + \frac{R_A}{R_B} = 1 + \frac{500}{500} = 2$$

$$\omega_C = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}} = \frac{1}{\sqrt{(500)^2 (2 \times 10^{-6})^2}} = 1{,}000 \text{ rad/s}$$

$$\frac{1}{Q} = 2\zeta = \sqrt{\frac{R_2 C_2}{R_1 C_1}} + \sqrt{\frac{R_1 C_2}{R_2 C_1}} + (K-1)\sqrt{\frac{R_1 C_1}{R_2 C_2}} = 1$$

Comments: The filter response can be compared to that of a quadratic Butterworth filter (or other filter family) by determining the Q of the filter. Once the gain and cutoff frequency have been defined, Q is the parameter that distinguishes, say, a Butterworth from a Chebyshev filter. The Butterworth polynomial of order 2 is given in Table 6.2 as $s^2 + \sqrt{2}s + 1$. Compare this expression to the denominator of equation 6.91, to obtain

$$\mathbf{H}(s) = \frac{K\omega_C^2}{s^2 + (\omega_C/Q)s + \omega_C^2} = \frac{1}{s^2 + \sqrt{2}s + 1}$$

Since the expressions for the quadratic polynomials of Table 6.2 are normalized to unity gain and cutoff frequency, $K = 1$ and $\omega_C = 1$, and therefore the value of Q in a Butterworth filter can be found by setting

$$\frac{1}{Q} = \sqrt{2} \quad \text{or} \quad Q = \frac{1}{\sqrt{2}} = 0.707$$

Thus, every second-order Butterworth filter is characterized by a Q of 0.707, which corresponds to a damping coefficient $\zeta = 0.5Q^{-1} = 0.707$, that is, to a lightly underdamped response.

Page 435

6.5 INTEGRATORS AND DIFFERENTIATORS

The time domain responses of certain op-amp circuits containing energy storage elements reveal useful and familiar properties. Among these circuits are the integrator and differentiator.

The Ideal Integrator

Consider the circuit of Figure 6.46, where $v_S(t)$ is an arbitrary function of time (e.g., a pulse train, a triangular wave, or a square wave). The op-amp circuit shown provides an output that is proportional to the integral of $v_S(t)$. The analysis of the integrator circuit is, as always, based on the observation that

$$i_S(t) = -i_F(t) \tag{6.93}$$

where

$$i_S(t) = \frac{v_S(t) - 0}{R_S} \tag{6.94}$$

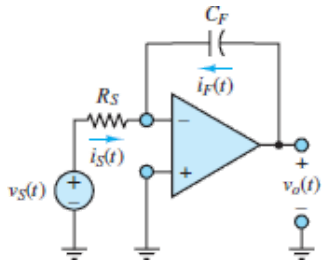


Figure 6.46 Op-amp integrator

It is also known that

$$i_F(t) = C_F \frac{d(v_o(t) - 0)}{dt} \tag{6.95}$$

from the *i-v* relationship for a capacitor. The source voltage can then be expressed as a function of the derivative of the output voltage:

$$\frac{1}{R_S C_F} v_S(t) = -\frac{dv_o(t)}{dt} \tag{6.96}$$

Integrate both sides of equation 6.96 to obtain:

LO

$$\Delta v_o(t) = -\frac{1}{R_S C_F} \int^t v_S(t')\,dt' \qquad \text{Integrator} \tag{6.97}$$

There are numerous applications for integrators. Example 6.7 illustrates the operation of the op-amp integrator.

The Ideal Differentiator

Using an argument similar to that employed for the integrator, we can derive a result for the ideal differentiator circuit of Figure 6.47. The relationship between input and output is obtained by observing that

$$i_S(t) = C_S \frac{d(v_S(t) - 0)}{dt} \tag{6.98}$$

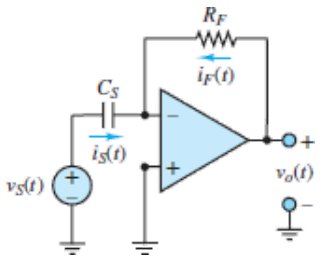


Figure 6.47 Op-amp differentiator

and

$$i_F(t) = \frac{v_o(t) - 0}{R_F} \tag{6.99}$$

so that the output of the differentiator circuit is proportional to the derivative of the input:

LO

$$v_o(t) = -R_F C_S \frac{dv_S(t)}{dt} \qquad \text{Op-amp differentiator} \tag{6.100}$$

Page 436

Although mathematically attractive, the differentiation property of this op-amp circuit is seldom used in practice because differentiation tends to amplify any noise that may be present in a signal.

FOCUS ON MEASUREMENTS

Charge Amplifiers

One of the most common families of transducers for the measurement of force, pressure, and acceleration is that of **piezoelectric transducers.** These transducers contain a piezoelectric crystal that produces a net electric charge separation in response to deformation. If an external force generates a displacement x_i, then the transducer will generate a charge separation q according to the expression:

$$q = K_P x_i$$

Figure 6.48 depicts the basic structure of the piezoelectric transducer, and a simple circuit model. The model consists of a current source in parallel with a capacitor, where the current source represents the rate of change of the charge in response to an external force; and the capacitance is a consequence of the structure of the transducer, which consists of a piezoelectric crystal (e.g., quartz or Rochelle salt) sandwiched between conducting electrodes (in effect, this is a parallel-plate capacitor).

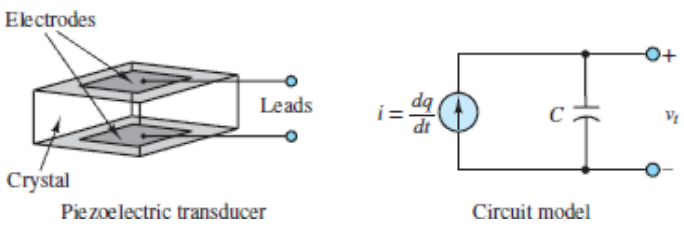


Figure 6.48 Piezoelectric transducer

Although it is possible, in principle, to employ a conventional voltage amplifier to amplify the transducer output voltage v_t, given by

$$v_t = \frac{1}{C}\int i\,dt = \frac{1}{C}\int \frac{dq}{dt}dt = \frac{q}{C} = \frac{K_p x_i}{C}$$

it is often advantageous to use a **charge amplifier**. The charge amplifier is essentially an integrator circuit, as shown in Figure 6.49, characterized by an extremely high input impedance.[3] The high impedance is essential; otherwise, the charge generated by the transducer would leak to ground through the input impedance of the amplifier.

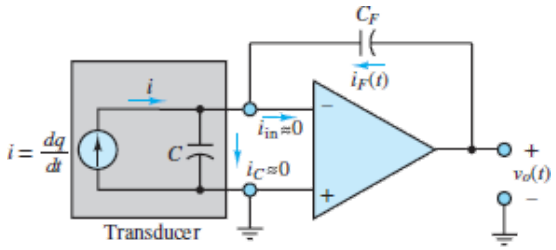


Figure 6.49 Charge amplifier

Page 437

Because of the high input impedance, the input current to the op-amp is negligible; further, because of the high open-loop gain of the op-amp, the inverting-terminal voltage is essentially at ground potential. Thus, *the voltage across the transducer is effectively zero.* As a consequence, to satisfy KCL, the feedback current $i_F(t)$ must be equal and opposite to the transducer current i:

$$i_F(t) = -i$$

It follows that the output voltage is proportional to the charge generated by the transducer, and therefore to the displacement as shown here:

$$v_o(t) = \frac{1}{C_F}\int -i\,dt = \frac{1}{C_F}\int -\frac{dq}{dt}dt = -\frac{q}{C_F} = -\frac{K_p x_i}{C_F}$$

Since the displacement is caused by an external force or pressure, this sensing principle is widely adopted in the measurement of force and pressure.

EXAMPLE 6.7 Integrating a Square Wave

Problem

Determine the output voltage for the integrator circuit of Figure 6.46 if the input is a square wave of amplitude ± A and period T, as shown in Figure 6.50.

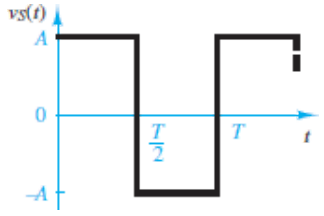


Figure 6.50

Solution

Known Quantities: Feedback and source impedances; input waveform characteristics.

Find: $v_o(t)$.

Schematics, Diagrams, Circuits, and Given Data: $T = 10$ ms; $C_F = 1$ μF; $R_S = 10$ k Ω.

Assumptions: The op-amp is ideal and $v_o = 0$ at $t = 0$.

Analysis: Equation 6.97 expresses the output of an integrator as:

$$\Delta v_o(t) = -\frac{1}{R_F C_S}\int^t v_S(t')\,dt' = -\frac{1}{R_F C_S}\int_0^t v_S(t')\,dt'$$

The square wave can be integrated in a piecewise fashion by observing that $v_S(t) = A$ for $0 \le t < T/2$ and $v_S(t) = -A$ for $T/2 \le t < T$. Thus, for the two half periods of

the waveform:

$$v_o(t) = v_0 - \frac{1}{R_F C_S}\int_0^t v_S(t')\,dt' = -100\int_0^t A\,dt'$$

$$= -100A(t-0) \qquad 0 \le t < \frac{T}{2}$$

$$v_o(t) = v_o\left(\frac{T}{2}\right) - \frac{1}{R_F C_s}\int_{T/2}^t v_S(t')\,dt' = -100A\frac{T}{2} - 100\int_{T/2}^t (-A)\,dt'$$

$$= -100A\frac{T}{2} + 100A\left(t - \frac{T}{2}\right) = 100A(t-T) \qquad \frac{T}{2} \le t < T$$

Since the waveform is periodic, the above result will repeat with period T, as shown in Figure 6.51. Note also that the average value of the output voltage is not zero.

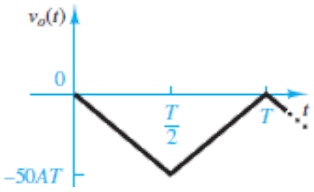


Figure 6.51

Comments: The integral of a square wave is a triangle wave. Note that the effect of the initial condition is very important since it determines the starting point of the triangular wave. The two line segments are expressed in terms of the slope and t-intercept, which is often a very friendly but neglected form.

EXAMPLE 6.8 Proportional-Integral Control with Op-Amps

Problem

The aim of this example is to illustrate the very common practice of *proportional-integral*, or *PI*, *control*. Consider the temperature control circuit of Example 6.4, shown again in Figure 6.52(a), where it was discovered that the proportional control implemented with the proportional gains K_V and K_P could still give rise to a steady-state error in the final temperature of the system. This error can be

eliminated by using an automatic control system that feeds back a component that is proportional to the *integral of the error voltage*, in addition to the proportional term. Figure 6.52(b) depicts the block diagram of such a *PI* controller. Now, the design of the control system requires selecting three gains, the *proportional gains* K_V and K_P, and the *integral gain* K_I.

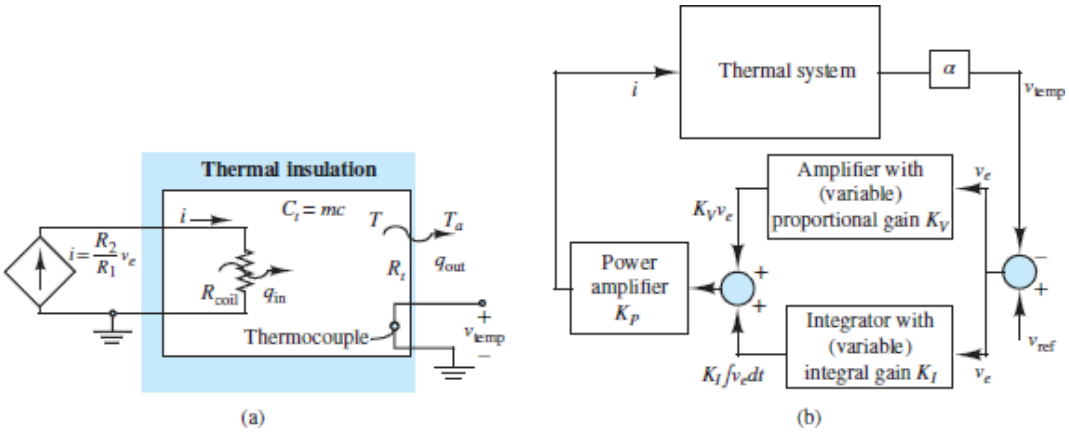


Figure 6.52 (a) Thermal system and (b) block diagram of control system

Page 439

Solution

Known Quantities: Sensor (input) voltage; feedback and source resistors, thermal system component values.

Find: Select desired value of proportional gain K_V and integral gain K_I to achieve automatic temperature control with zero steady-state error. Assume $K_P = 1$, for simplicity.

Schematics, Diagrams, Circuits and Given Data: $R_{\text{coil}} = 5\ \Omega$; $R_t = 2°\text{C/W}$; $C_t = 50\ \text{J/°C}$; $\alpha = 1\ \text{V/°C}$.

Assumptions: Assume ideal op-amps.

Analysis: The circuit of Figure 6.52(c) shows two op-amp circuits—the top circuit generates the error voltage $v_e = v_{\text{ref}} - v_{\text{temp}}$ but does not provide any gain. The bottom circuit amplifies v_e by the proportional gain $-K_V = -R_2/R_1$ and also computes the integral of v_e times the integral gain $-K_I = -1/R_3C$. These two quantities are then summed through another inverting summer circuit, which takes care of the sign change as well.

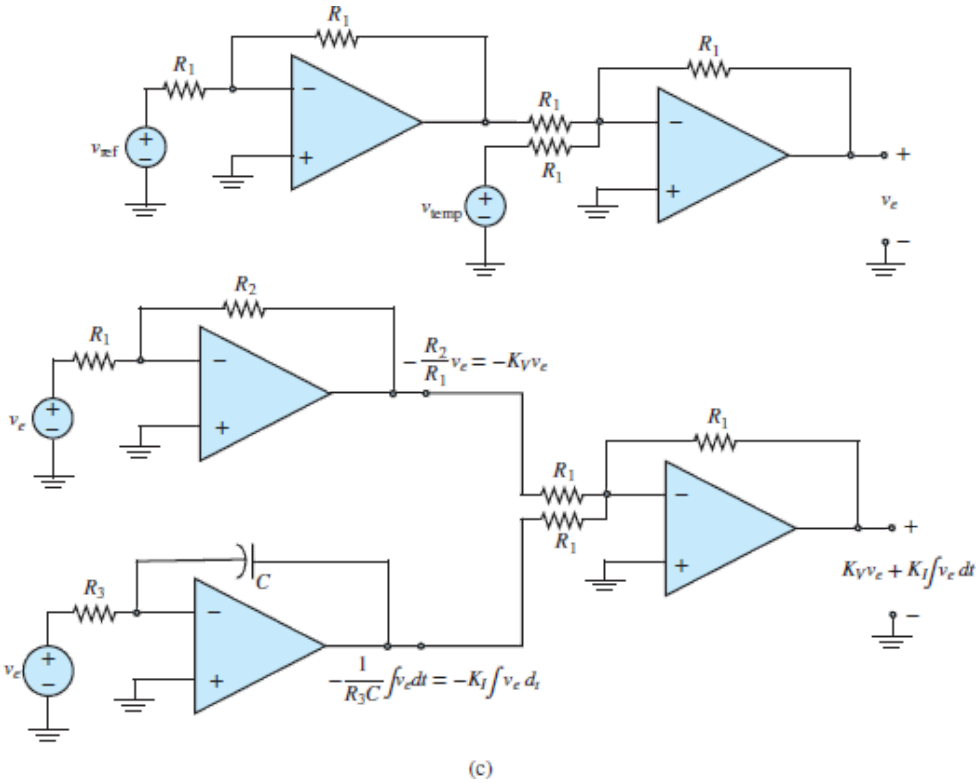


Figure 6.52(c) Circuit for generating error voltage and proportional gain

Figure 6.52(d) depicts the temperature response of the system for $K_V = 5$ and different values of K_I. Note that the steady-state error is now zero! This result is a property of integral controllers. Figure 6.52(e) shows the current supplied to the heater coil. Note the speed and shape of each response and the negligible long-term steady-state temperature error.

Comments: At sufficiently high values of K_I the system temperature oscillates in response to the −10°C temperature disturbance described in Example 6.4. This oscillation is a characteristic of an underdamped second-order system (see Chapter 4)—but the original Page 440thermal system is first order! The addition of the integral term has increased the order of the system such that it is possible for the system to oscillate. To those familiar with thermal systems, this behavior should cause a raised eyebrow! It is well known that thermal systems cannot display underdamped behavior (that is, there is no thermal system property analogous to inductance). The introduction of the integral gain can, in fact, cause temperature oscillations as if an artificial "thermal inductor" were introduced in the system.

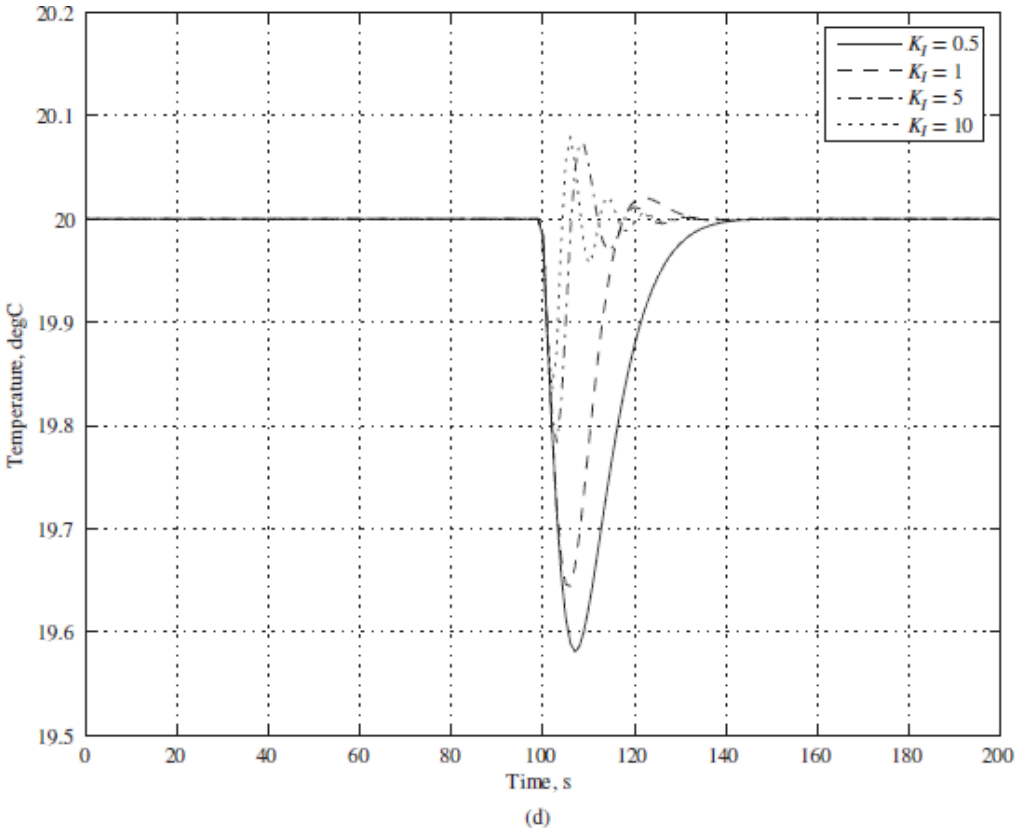


Figure 6.52(d) Response of thermal system for various values of integral gain, $K_I (K_V = 5)$

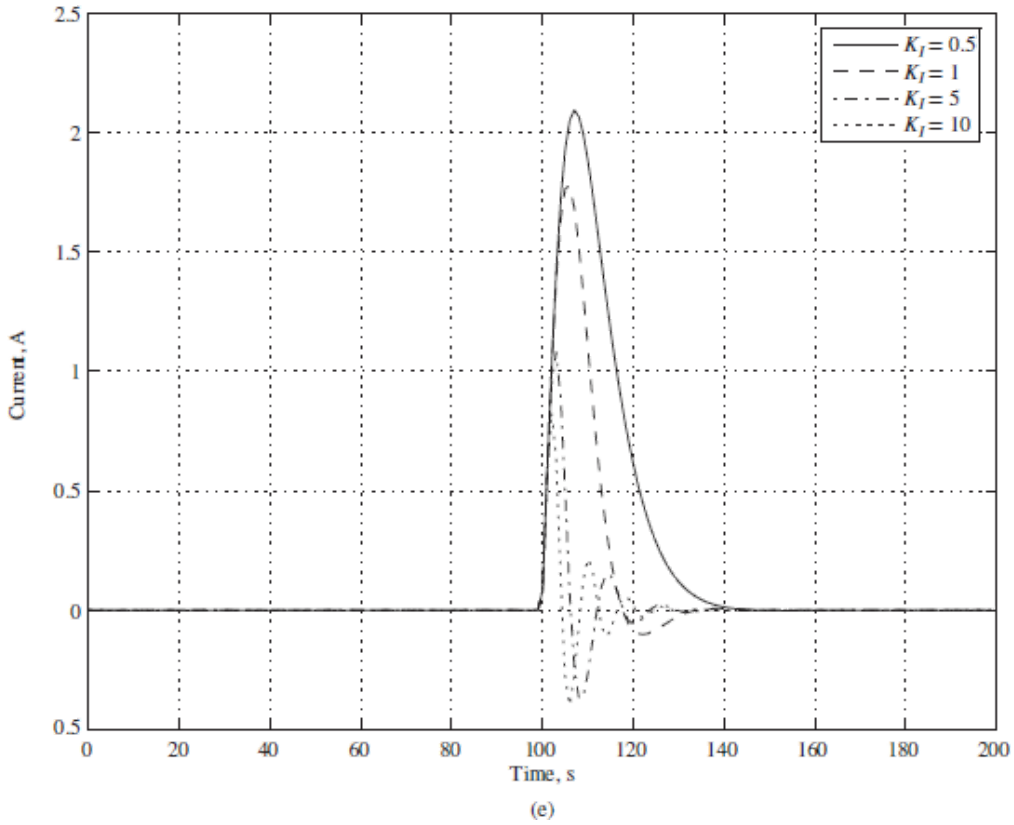


Figure 6.52(e) Power amplifier current for various values of integral gain $K_I (K_V = 5)$

Page 441

EXAMPLE 6.9 Using Cascaded Amplifiers to Simulate a Differential Equation

Problem

Derive the differential equation corresponding to the circuit shown in Figure 6.53.

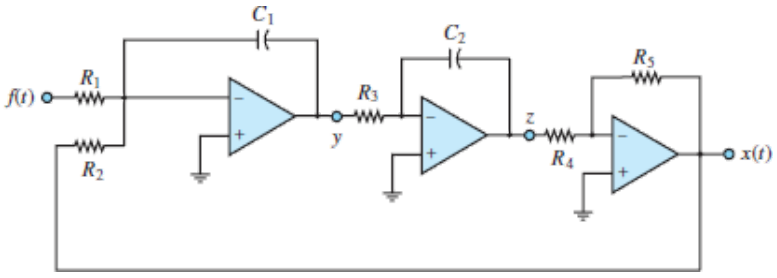


Figure 6.53 Analog computer simulation of unknown system

Solution

Known Quantities: Resistor and capacitor values.

Find: Differential equation in $x(t)$.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 0.4$ MΩ; $R_2 = R_3 = R_5 = 1$ MΩ; $R_4 = 2.5$ kΩ; $C_1 = C_2 = 1$ μF.

Assumptions: Assume ideal op-amps.

Analysis: Begin the analysis from the right-hand side of the circuit to determine the intermediate variable z as a function of x:

$$x = -\frac{R_5}{R_4} z = -400z$$

Moving to the left, next determine the relationship between y and z:

$$z = -\frac{1}{R_3 C_2} \int y(t')\,dt' \qquad \text{or} \qquad y = -R_3 C_2 \frac{dz}{dt} = -\frac{dz}{dt}$$

Finally, determine y as a function of x and f:

$$y = -\frac{1}{R_2 C_1} \int x(t')\,dt' - \frac{1}{R_1 C_1} \int f(t')\,dt' = -\int [x(t') + 2.5f(t')]\,dt'$$

or

$$\frac{dy}{dt} = -x - 2.5f$$

Page 442

Substitute the expressions into one another and eliminate y and z to obtain:

$$x = -400z$$

$$\frac{dx}{dt} = -400\frac{dz}{dt} = 400y$$

$$\frac{d^2x}{dt^2} = 400\frac{dy}{dt} = 400(-x - 2.5f)$$

Thus:

$$\frac{d^2x}{dt^2} + 400x = -1{,}000f$$

Comments: Note that the summing and integrating functions have been combined into a single block in the first amplifier.

CHECK YOUR UNDERSTANDING

Plot the frequency response of an ideal integrator in the form of a Bode plot. Determine its slope in dB/decade. Assume $R_S C_F = 10$ s.

Answer: –20 dB/decade

CHECK YOUR UNDERSTANDING

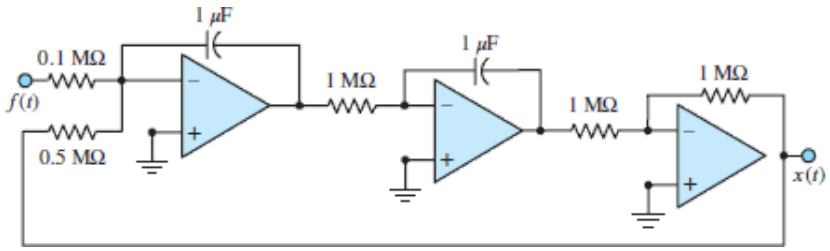
Plot the frequency response of an ideal differentiator in the form of a Bode plot. Determine its slope in dB/decade. Assume $R_F C_S = 100$ s.

Verify that, if the triangular wave of Example 6.7 is the input to the ideal differentiator of Figure 6.47, the resulting output is a square wave.

Answer: +20 dB/decade

CHECK YOUR UNDERSTANDING

Derive the differential equation corresponding to the circuit shown in the figure.



Answer: $d^2x/dt^2 + 2x = -10f(t)$

Page 443

6.6 PHYSICAL LIMITATIONS OF OPERATIONAL AMPLIFIERS

In nearly all the discussion and examples so far, the op-amp has been treated as an ideal device, characterized by infinite input impedance, zero output resistance, and infinite open-loop voltage gain. Although this model is adequate to represent its behavior in a large number of applications, practical op-amps are not ideal but exhibit limitations that should be considered in the design of instrumentation. In particular, in dealing with relatively large voltages and currents, and in the presence of high-frequency signals, it is important to be aware of the nonideal properties of the op-amp.

Voltage Supply Limits

As indicated in Figure 6.6, operational amplifiers (and all amplifiers, in general) are powered by external DC voltage supplies V_S^+ and V_S^-, which are usually symmetric and on the order of ±10 to ±20 V. Some op-amps are especially designed to operate from a single voltage supply. Amplifiers are capable of amplifying signals *only within the range of their supply voltages;* it is not possible for an amplifier to generate a voltage greater than V_S^+ or less than V_S^-.

$$V_S^- < v_{out} < V_S^+ \qquad \text{Voltage supply limitation} \tag{6.101}$$

For most op-amps, the limit is approximately 1.4 V less than the supply voltages.

Example 6.10 shows how the voltage supply limit can cause the peaks of the sine wave to be clipped in an abrupt fashion. This type of hard nonlinearity changes the characteristics of the signal quite radically. For example, a rock guitar has a characteristic sound that is very different from the sound of a classical or jazz guitar. The reason is that the "rock sound" is obtained by overamplifying the signal, attempting to exceed the voltage supply limits, and causing clipping similar in quality to the distortion introduced by voltage supply limits in an op-amp. This clipping broadens the spectral content of each tone and causes the sound to be distorted.

One of the circuits most directly affected by supply voltage limitations is the op-amp integrator.

Frequency Response Limits

Another property of all amplifiers that may pose severe limitations to the op-amp is their finite bandwidth. In the ideal op-amp model the open-loop gain is a very large constant. In reality, A is a function of frequency and is characterized by a low-pass response. For a typical op-amp,

$$A(j\omega) = \frac{A_0}{1 + j\omega/\omega_0} \qquad \text{Finite bandwidth limitation} \tag{6.102}$$

The cutoff frequency of the op-amp open-loop gain ω_0 represents approximately the point where the amplifier response starts to drop off as a function of frequency and is analogous to the cutoff frequencies of the RC and RL circuits of Chapter 5. Figure 6.54 depicts $A(j\omega)$ in both linear and decibel plots for the fairly typical values $A_0 = 10^6$ and $\omega_0 = 10\pi$. It should be apparent from Figure 6.54 that the assumption of a very large open-loop gain becomes less and less accurate for increasing frequency. Recall the initial derivation of the closed-loop gain for the

inverting amplifier: In obtaining the final result $\mathbf{V}_o/\mathbf{V}_S = -R_F/R_S$, it was assumed that $A \to \infty$. This assumption is clearly inadequate at the higher frequencies.

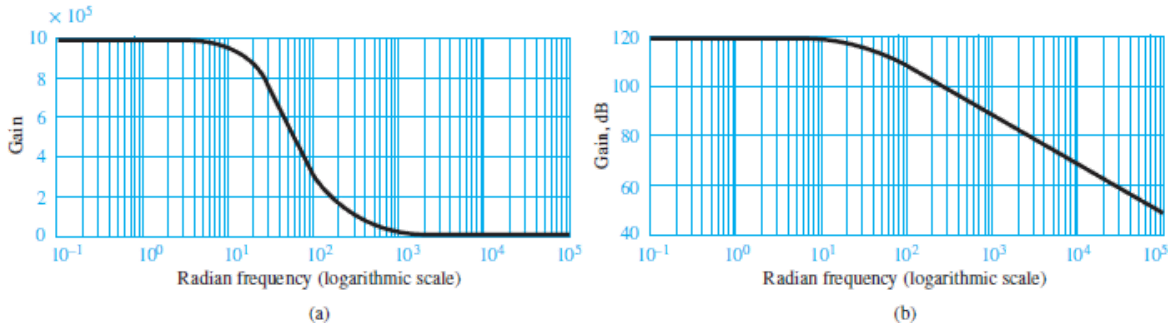


Figure 6.54 Open-loop gain of practical op-amp (a) amplitude ratio response; (b) dB response

The finite bandwidth of the practical op-amp results in a constant **gain-bandwidth product** for any given amplifier. The effect is that as the closed-loop gain of the amplifier is increased, its 3-dB bandwidth is proportionally reduced until, in the limit, if the amplifier were used in the open-loop mode, its gain would be equal to A_0 and its 3-dB bandwidth would be equal to ω_0. The constant gain-bandwidth product is therefore equal to the product of the open-loop gain and the open-loop bandwidth of the amplifier: $A_0\omega_0 = K$. When the amplifier is connected in a closed-loop configuration (e.g., as an inverting amplifier), its gain is typically much less than the open-loop gain and the 3-dB bandwidth of the amplifier is proportionally increased. To explain this further, Figure 6.55 depicts the case in which two different linear amplifiers (achieved through any two different negative feedback configurations) have been designed for the same op-amp. The first has closed-loop gain $G_1 = A_1$, and the second has closed-loop gain $G_2 = A_2$. The bold line in the figure indicates the open-loop frequency response, with gain A_0 and cutoff frequency ω_0. As the gain decreases from A_0 to A_1, the cutoff frequency increases from ω_0 to ω_1. As the gain decreases to A_2, the bandwidth increases to ω_2. Thus:

The gain-bandwidth product of an op-amp is constant.

$$A_0 \times \omega_0 = A_1 \times \omega_1 = A_2 \times \omega_2 = K \tag{6.103}$$

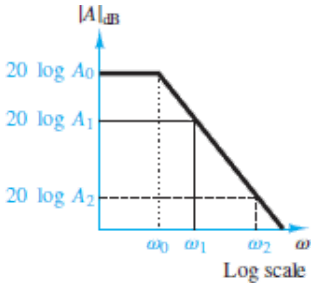


Figure 6.55

Input Offset Voltage

Another limitation of practical op-amps results because even in the absence of any external inputs, it is possible that an **offset voltage** will be present at the input of an op-amp. This voltage is usually denoted by $\pm V_{os}$, and it is caused by mismatches Page 445in the internal circuitry of the op-amp. The offset voltage appears as a differential input voltage between the inverting and noninverting input terminals. The presence of an additional input voltage will cause a DC bias error in the amplifier output. Typical and maximum values of V_{os} are quoted in manufacturers' data sheets. The worst-case effects due to the presence of offset voltages can therefore be predicted for any given application.

Input Bias Currents

Another nonideal characteristic of op-amps results from the presence of small input bias currents at the inverting and noninverting terminals. Once again, these are due to the internal construction of the input stage of an operational amplifier. Figure 6.56 illustrates the presence of nonzero input bias currents I_B going into an op-amp.

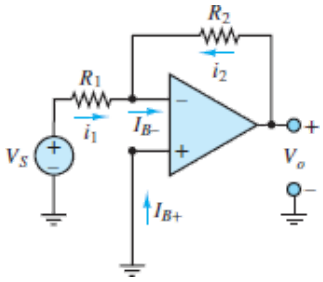


Figure 6.56

Typical values of I_{B+} and I_{B-} depend on the semiconductor technology employed in the construction of the op-amp. Op-amps with bipolar transistor input stages may see input bias currents as large as 1 μA, while for FET input devices, the input bias currents are less than 1 nA.

One often designates the **input offset current** I_{os} as

$$I_{os} = I_{B+} - I_{B-} \tag{6.104}$$

Output Offset Adjustment

Both the offset voltage and the input offset current contribute to an output offset voltage $V_{o,os}$. Some op-amps provide a means for minimizing $V_{o,os}$. For example, the μA741 op-amp provides a connection for this procedure. Figure 6.57 shows a typical pin configuration for an op-amp in an eight-pin dual-in-line package (DIP) and the circuit used for nulling the output offset voltage. The variable resistor is adjusted until v_o reaches a minimum (ideally, 0 V). Nulling the output voltage in this manner removes the effect of both input offset voltage and current on the output.

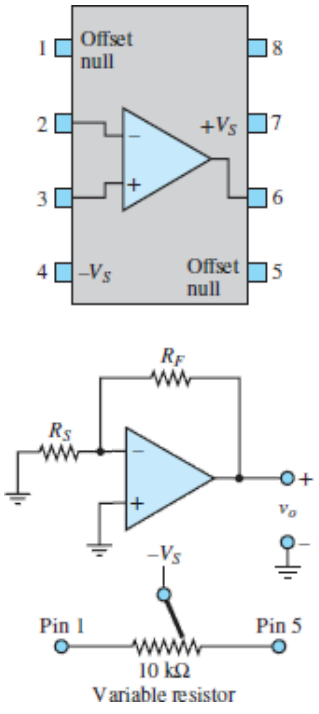


Figure 6.57 Output offset voltage adjustment

Slew Rate Limit

Another important restriction in the performance of a practical op-amp is associated with rapid changes in voltage. The op-amp can produce only a finite rate of change at its output. This limit rate is called the **slew rate**. Consider an ideal step input, where at $t = 0$ the input voltage is switched from 0 to V volts. The output may be expected to switch from 0 to AV volts, where A is the amplifier gain. However, v_o can change at only a finite rate; thus,

$$\left|\frac{dv_o}{dt}\right|_{\max} = S_0 \qquad \text{Slew rate limitation} \tag{6.105}$$

Figure 6.58 shows the response of an op-amp to an ideal step change in input voltage. Here, S_0, the slope of v_o, represents the slew rate.

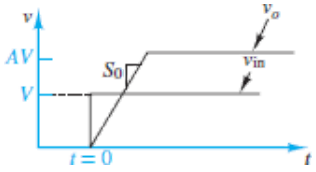


Figure 6.58 Slew rate limit in op-amps

Page 446

The slew rate limitation can affect sinusoidal signals, as well as signals that display abrupt changes, as does the step voltage of Figure 6.58. This may not be obvious until we examine the sinusoidal response more closely. It should be apparent that the maximum rate of change for a sinusoid occurs at the zero crossing, as shown by Figure 6.59. To evaluate the slope of the waveform at the zero crossing, let

$$v_{\text{in}}(t) = V \sin \omega t \qquad \text{such that} \qquad v_o(t) = AV \sin \omega t \tag{6.106}$$

Then:

$$\frac{dv_o}{dt} = \omega AV \cos \omega t \tag{6.107}$$

The maximum slope of the sinusoidal signal will therefore occur at $\omega t = 0, \pi, 2\pi, \ldots$, so that

$$\left|\frac{dv_o}{dt}\right|_{\max} = \omega AV = S_0 \tag{6.108}$$

Thus, the maximum slope of a sinusoid is proportional to both the signal frequency and the amplitude. The curve shown by a dashed line in Figure 6.59 indicates that as ω increases, so does the slope of $v(t)$ at the zero crossings. What is the direct consequence of this result, then?

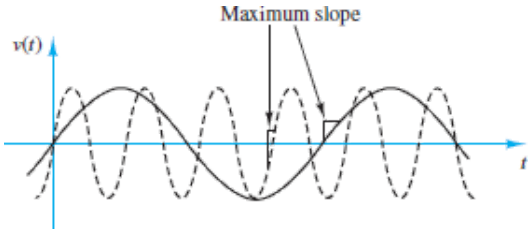


Figure 6.59 The maximum slope of a sinusoidal signal varies with the signal frequency

Short-Circuit Output Current

Recall the model for the op-amp shown in Figure 6.3, which depicted the internal circuit of the op-amp as an equivalent input impedance R_{in} and a controlled voltage source Av_{in}. In practice, the internal source is not ideal because it cannot provide an infinite amount of current (to the load, to the feedback connection, or to both). The immediate consequence of this nonideal op-amp characteristic is that the maximum output current of the amplifier is limited by the so-called short-circuit output current I_{SC}:

$$|i_o| < I_{SC} \qquad \text{Short-circuit output current limitation} \tag{6.109}$$

To further explain this point, consider that the op-amp needs to provide current to the feedback path (to "zero" the voltage differential at the input) and to whatever load resistance, R_o, may be connected to the output. Figure 6.60 illustrates this idea for the case of an inverting amplifier, where I_{SC} is the load current that would be provided to a short-circuit load ($R_o \rightarrow 0$). Clearly, as $R_o \rightarrow 0$ the output voltage $v_o \rightarrow 0$ unless $i_o \rightarrow \infty$. However, in practice i_o is limited to a finite value. Thus, Page 447eventually, as $R_o \rightarrow 0$ the magnitude of the output voltage v_o will decrease from $(R_F/R_S)\, v_S$ and approach zero. At this point, the circuit is no longer acting as one would expect for an ideal op-amp with negative feedback.

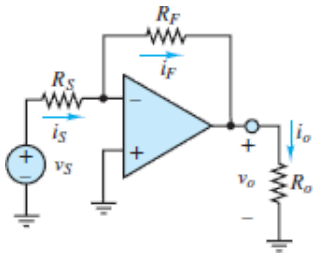


Figure 6.60

Common-Mode Rejection Ratio

The concepts of common-mode and differential-mode voltages as well as the common-mode rejection ratio (CMRR) were introduced in Section 6.2 and expressed mathematically by equations 6.66 to 6.70. The CMRR is an amplifier characteristic that can be found in the data sheet for any particular amplifier, such as a 741 operational amplifier.

$$\text{CMRR} = 20 \log \left| \frac{A_{\text{DM}}}{A_{\text{CM}}} \right| \quad \text{(in dB)}$$

Practical Op-Amp Considerations

The results presented in the preceding pages suggest that operational amplifiers permit the design of rather sophisticated circuits in a few simple steps, by selecting appropriate resistor values. This is certainly true, provided that the circuit component selection satisfies certain criteria. A few important practical criteria for selecting op-amp circuit component values are summarized here.

1. Use standard resistor values. While any arbitrary value of gain can, in principle, be achieved by selecting the appropriate combination of resistors, the designer is often constrained to the use of standard 5 percent resistor values. For example, if a design requires a gain of 25, it might be tempting to select, say, 100- and 4-kΩ resistors to achieve $R_F/R_S = 25$ for the inverting amplifier shown in Figure 6.56. However, 4 kΩ is not a standard value; the closest 5 percent tolerance resistor value is 3.9 kΩ, leading to a gain of 25.64. Can you find a combination of standard 5 percent resistors whose ratio is closer to 25?
2. Ensure that the load current is reasonable. Assume the maximum output voltage in the step 1 example is 10 V. The feedback current required by your design with $R_F = 100$ k Ω and $R_S = 4$ k Ω would be $I_F = 10/100{,}000 = 0.1$ mA. This is a very reasonable value for an op-amp. If you tried to achieve the same gain by using, say, a 10-Ω feedback resistor and a 0.39-Ω source resistor, the feedback current would become as large as 1 A. This value is generally beyond the capabilities of a general-purpose op-amp, so very low resistor values are generally not acceptable. On the other hand, 10-kΩ and 390-Ω resistors would still lead to acceptable currents. As a general rule of thumb, avoid resistor values lower than 100 Ω in practical designs.
3. Avoid stray capacitance by avoiding excessively large resistances, which can cause unwanted signals to couple into the circuit through a mechanism known as *capacitive coupling*. This phenomenon is discussed in Chapter 7. Large resistances can also cause other problems. As a general rule of thumb, avoid resistor values higher than 1 MΩ in practical designs.
4. Precision designs may be warranted. If a certain design requires that the amplifier gain be set to a very accurate value, it may be appropriate to use the (more expensive) option of precision resistors: for example, 1 percent tolerance resistors are commonly available, at a premium cost. Some of

the examples and homework problems explore the variability in gain due to the use of higher- and lower-tolerance resistors.

EXAMPLE 6.10 Voltage Supply Limits in an Inverting Amplifier

Problem

Compute and sketch the output voltage of the inverting amplifier of Figure 6.61.

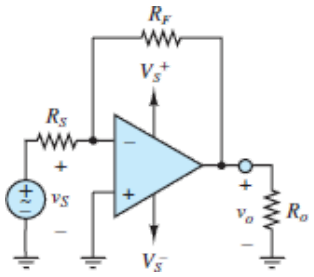


Figure 6.61

Solution

Known Quantities: Resistor and supply voltage values; input voltage.

Find: $v_o(t)$.

Schematics, Diagrams, Circuits, and Given Data: $R_S = 1$ k Ω; $R_F = 10$ k Ω; $R_o = 1$ k Ω; $V_S^+ = 15$ V; $V_S^- = -15$ V; $v_S(t) = 2\sin(1{,}000t)$.

Assumptions: Assume a supply voltage–limited op-amp.

Analysis: For an ideal op-amp the output would be

$$v_o(t) = -\frac{R_F}{R_S}v_S(t) = -10 \times 2\sin(1{,}000t) = -20\sin(1{,}000t)$$

However, the supply voltage is limited to ±15 V, and the op-amp output voltage will therefore saturate before reaching the theoretical peak output value of ±20 V.

Figure 6.62 depicts the output voltage waveform.

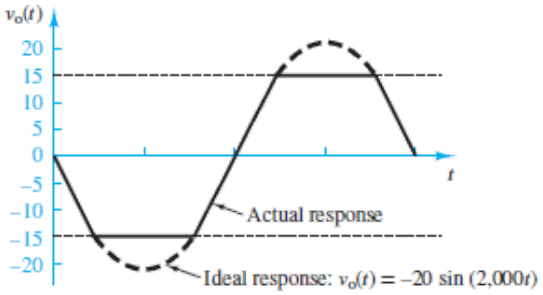


Figure 6.62 Op-amp output with voltage supply limit

Comments: In a practical op-amp, saturation would be reached at 1.4 V from the supply voltages, or at approximately ±13.6 V.

EXAMPLE 6.11 Voltage Supply Limits in an Op-Amp Integrator

Problem

Compute and sketch the output voltage of the integrator of Figure 6.46.

Solution

Known Quantities: Resistor, capacitor, and supply voltage values; input voltage.

Find: $v_o(t)$.

Page 449

Schematics, Diagrams, Circuits, and Given Data: $R_S = 10\ \text{k}\Omega$; $C_F = 20\ \mu\text{F}$; $V_S^+ = 15\ \text{V}$; $V_S^- = -15\ \text{V}$; $v_S(t) = -0.5 + 0.3\cos(10t)$.

Assumptions: Assume a supply voltage–limited op-amp. The initial condition is $v_o(0) = 0$.

Analysis: For an ideal op-amp integrator the output would be

$$v_o(t) = v_o(0) - \frac{1}{R_S C_F}\int_0^t v_S(t')dt' = -\frac{1}{0.2}\int_0^t [-0.5 + 0.3\cos(10t')]\,dt'$$

$$= 2.5t + 0.15\sin(10t)$$

However, the supply voltage is limited to ±15 V, and the integrator output voltage will therefore saturate at the upper supply voltage value of 15 V as the term 2.5*t* increases with time. Figure 6.63 depicts the output voltage waveform.

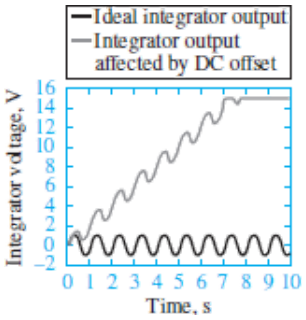


Figure 6.63 Effect of DC offset on integrator

Comments: Note that the DC offset in the waveform causes the integrator output voltage to increase linearly with time. The presence of even a very small DC offset will always cause integrator saturation.

EXAMPLE 6.12 Gain-Bandwidth Product Limit in an Op-Amp

Problem

Determine the maximum allowable closed-loop voltage gain of an op-amp if the amplifier is required to have an audio-range bandwidth of 20 kHz.

Solution

Known Quantities: Gain-bandwidth product.

Find: G_{max}.

Schematics, Diagrams, Circuits, and Given Data: $A_0 = 10^6$; $\omega_0 = 2\pi \times 5$ rad/s.

Assumptions: Assume a gain-bandwidth product limited op-amp.

Analysis: The gain-bandwidth product of the op-amp is

$$A_0 \times \omega_0 = K = 10^6 \times 2\pi \times 5 = \pi \times 10^7 \text{ rad/s}$$

The desired bandwidth is $\omega_{max} = 2\pi \times 20{,}000$ rad/s, and the maximum allowable gain will therefore be

$$G_{max} = \frac{K}{\omega_{max}} = \frac{\pi \times 10^7}{\pi \times 4 \times 10^4} = 250 \frac{V}{V}$$

For any closed-loop voltage gain greater than 250, the amplifier would have reduced bandwidth.

Comments: To achieve gains greater than 250 and maintain the same bandwidth, two options are available: (1) Use a different op-amp with greater gain-bandwidth product, or (2) connect two amplifiers in cascade, each with lower gain and greater bandwidth, such that the product of the gains would be greater than 250.

Page 450

EXAMPLE 6.13 Increasing the Gain-Bandwidth Product by Means of Amplifiers in Cascade

Problem

Determine the overall 3-dB bandwidth of the cascade amplifier of Figure 6.64.

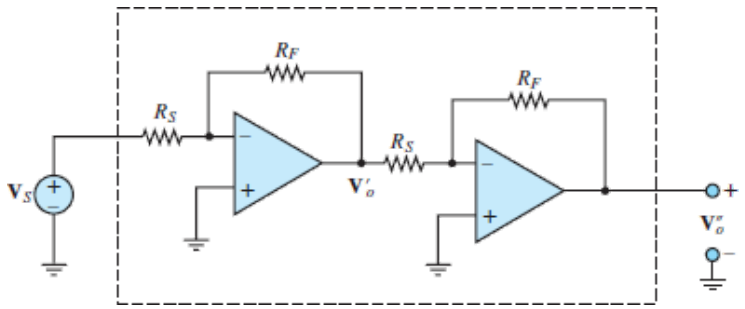


Figure 6.64 Cascade amplifier

Solution

Known Quantities: Gain-bandwidth product and gain of each amplifier.

Find: $\omega_{3\text{ dB}}$ of cascade amplifier.

Schematics, Diagrams, Circuits, and Given Data: $A_0\omega_0 = K = 4\pi \times 10^6$ for each amplifier. $R_F/R_S = 100$ for each amplifier.

Assumptions: Assume gain-bandwidth product limited (otherwise ideal) op-amps.

Analysis: Let G_1 and ω_1 denote the gain and the 3-dB bandwidth of the first amplifier, respectively, and G_2 and ω_2 those of the second amplifier.

The 3-dB bandwidth of the first amplifier is

$$\omega_1 = \frac{K}{G_1} = \frac{4\pi \times 10^6}{10^2} = 4\pi \times 10^4 \frac{\text{rad}}{\text{s}}$$

The second amplifier will also have

$$\omega_2 = \frac{K}{G_2} = \frac{4\pi \times 10^6}{10^2} = 4\pi \times 10^4 \frac{\text{rad}}{\text{s}}$$

Thus, the approximate bandwidth of the cascade amplifier is $4\pi \times 10^4$, and the gain of the cascade amplifier is $G_1G_2 = 100 \times 100 = 10^4$ or 80 dB.

For a single-stage amplifier having the same K the bandwidth is 100 times smaller.

$$\omega_3 = \frac{K}{G_3} = \frac{4\pi \times 10^6}{10^4} = 4\pi \times 10^2 \frac{\text{rad}}{\text{s}}$$

Comments: In practice, the actual 3-dB bandwidth of the cascade amplifier is not quite as large as that of each of the two stages because the gain of each amplifier starts decreasing at frequencies somewhat lower than the nominal cutoff frequency. A more detailed analysis shows that the actual 3-dB bandwidth is roughly $\omega_1\sqrt{\sqrt{2} - 1}$.

EXAMPLE 6.14 Effect of Input Offset Voltage on an Amplifier

Problem

Determine the effect of the input offset voltage V_{os} on the output of the amplifier shown in Figure 6.65.

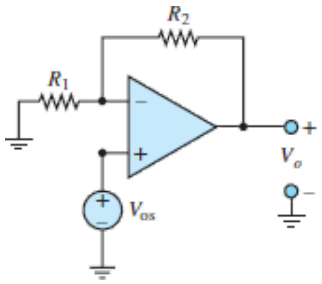


Figure 6.65 Op-amp input offset voltage

Solution

Known Quantities: Nominal closed-loop voltage gain; input offset voltage.

Find: The offset voltage component in the output voltage V_o, os.

Schematics, Diagrams, Circuits, and Given Data: $A_{nom} = 100$; $V_{os} = 1.5$ mV.

Assumptions: Assume an input offset voltage–limited (otherwise ideal) op-amp.

Analysis: The amplifier is connected in a noninverting configuration; thus its nominal closed-loop gain is

$$G_{nom} = 100 = 1 + \frac{R_F}{R_S}$$

The DC offset voltage, represented by an ideal voltage source, is directly applied to the noninverting input. Thus

$$V_{o,os} = G_{nom} V_{os} = 100\, V_{os} = 150 \text{ mV}$$

Thus, we should expect the output of the amplifier to be shifted upward by 150 mV.

Comments: The input offset voltage is not, of course, an external source, but is a voltage offset between the inputs of the op-amp. Figure 6.57 depicts how such an offset can be zeroed. The worst-case offset voltage is usually listed in the device data sheets. Typical values are 2 mV for the 741c general-purpose op-amp and 5 mV for the FET-input TLO81.

EXAMPLE 6.15 Effect of Input Offset Current on an Amplifier

Problem

Determine the effect of the input offset current I_{os} on the output of the amplifier of Figure 6.66.

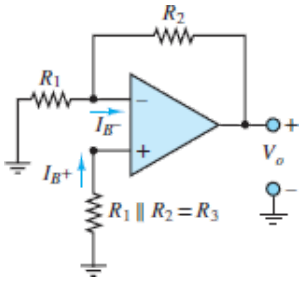


Figure 6.66

Solution

Known Quantities: Resistor values; input offset current.

Find: The offset voltage component in the output voltage $v_{out,\ os}$.

Schematics, Diagrams, Circuits, and Given Data: $I_{os} = 1\ \mu$ A; $R_2 = 10$ k Ω.

Assumptions: Assume an input offset current–limited (otherwise ideal) op-amp.

Analysis: Calculate the inverting and noninverting terminal voltages caused by the offset current in the absence of an external input:

$$v^+ = -R_3 I_{B^+} \qquad v^- = v^+ = -R_3 I_{B^+}$$

Page 452

With these values apply KCL at the inverting node and write

$$\frac{V_o - v^-}{R_2} - \frac{v^- - 0}{R_1} = I_{B^-}$$

$$\frac{V_o}{R_2} - \frac{-R_3 I_{B^+}}{R_2} - \frac{-R_3 I_{B^+}}{R_1} = I_{B^-}$$

$$V_o = R_2\left[-I_{B^+}R_3\left(\frac{1}{R_2} + \frac{1}{R_1}\right) + I_{B^-}\right] = R_2[-I_{B^+} + I_{B^-}] = -R_2 I_{os}$$

Thus, we should expect the output of the amplifier to be shifted downward by $R_2 I_{os}$, or $10^4 \times 10^{-6} = 10$ mV for the data given in this example.

Comments: Usually, the worst-case input offset currents (or input bias currents) are listed in the device data sheets. Values can range from 100 pA (for CMOS op-amps, for example, LMC6061) to around 200 nA for a low-cost general-purpose amplifier (for example, μA741c).

EXAMPLE 6.16 Effect of Slew Rate Limit on an Amplifier

Problem

Determine the effect of the slew rate limit S_0 on the output of an inverting amplifier for a sinusoidal input voltage of known amplitude and frequency.

Solution

Known Quantities: Slew rate limit S_0; amplitude and frequency of sinusoidal input voltage; amplifier closed-loop gain.

Find: Sketch the theoretically correct output and the actual output of the amplifier in the same graph.

Schematics, Diagrams, Circuits, and Given Data: $S_0 = 1$ V/μ *s*; $v_S = \sin(2\pi \times 10^5 t)$; $G = 10$.

Assumptions: Assume the op-amp is slew rate limited, but otherwise ideal.

Analysis: Given the closed-loop voltage gain of 10, compute the theoretical output voltage to be:

$$v_o = -10\sin(2\pi \times 10^5 t)$$

The maximum slope of the output voltage is then computed as follows:

$$\left|\frac{dv_o}{dt}\right|_{\max} = G\omega = 10 \times 2\pi \times 10^5 = 6.28\,\frac{\text{V}}{\mu\text{s}}$$

Clearly, the value calculated above far exceeds the slew rate limit. Figure 6.67 depicts the approximate appearance of the waveforms that one would measure in an experiment.

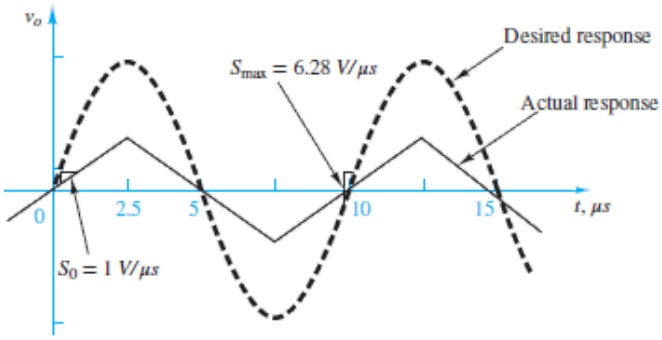


Figure 6.67 Distortion introduced by slew rate limit

Comments: Note that in this example the slew rate limit has been exceeded severely, and the output waveform is visibly distorted, to the point that it has effectively become a triangular wave. The effect of the slew rate limit is not always necessarily so dramatic and visible; thus one needs to pay attention to the specifications of a given op-amp. The slew Page 453rate limit is listed in the device data sheets. Typical values can range from 13 V/μ *s*, for the TLO81, to around 0.5 V/μ *s* for a low-cost general-purpose amplifier (for example, μ A741c).

EXAMPLE 6.17 Effect of Short-Circuit Current Limit on an Amplifier

Problem

Determine the effect of the short-circuit limit I_{SC} on the output of an inverting amplifier for a sinusoidal input voltage of known amplitude.

Solution

Known Quantities: Short-circuit current limit I_{SC}; amplitude of sinusoidal input voltage; amplifier closed-loop gain.

Find: Compute the minimum allowable load resistance value $R_{o_{min}}$, and sketch the theoretical and actual output voltage waveforms for resistances smaller than $R_{o_{min}}$.

Schematics, Diagrams, Circuits, and Given Data: I_{SC} = 50 mA; v_S = 0.05 sin(ωt); G = 100.

Assumptions: Assume the op-amp is short-circuit current limited, but otherwise ideal.

Analysis: Given the closed-loop voltage gain of 100, compute the theoretical output voltage to be:

$$v_o(t) = -G\,v_S(t) = -5\sin(\omega t)$$

To assess the effect of the short-circuit current limit, calculate the peak value of the output voltage since this is the condition that will require the maximum output current from the op-amp:

$$V_{o_{peak}} = 5\text{V}$$

$$I_{SC} = 50\,\text{mA}$$

$$R_{o_{min}} = \frac{V_{o_{peak}}}{I_{SC}} = \frac{5\,\text{V}}{50\,\text{mA}} = 100\,\Omega$$

For any load resistance less than 100 Ω, the required load current will be greater than I_{SC}. For a 75-Ω load resistor:

$$V_{o_{peak}} = I_{SC} \times R_o = 3.75\,\text{V}$$

That is, the output voltage cannot reach the theoretically correct 5-V peak and would be "compressed" to reach a peak voltage of only 3.75 V. This effect is depicted in Figure 6.68.

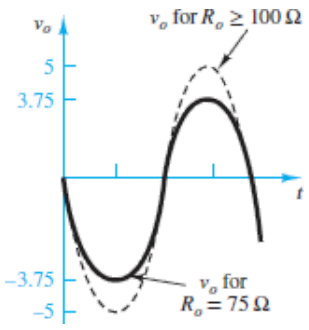


Figure 6.68 Distortion introduced by short-circuit current limit

Page 454

Comments: The short-circuit current limit is listed in the device data sheets. Typical values for a low-cost general-purpose amplifier (say, the 741c) are in the tens of milliamperes.

CHECK YOUR UNDERSTANDING

How long will it take (approximately) for the integrator of Example 6.11 to saturate if the input signal has a 0.1-VDC bias [that is, $v_S(t) = 0.1 + 0.3\cos(10t)$]?

Answer: Approximately 30 s

CHECK YOUR UNDERSTANDING

What is the maximum gain that could be achieved by the op-amp of Example 6.12 if the desired bandwidth is 100 kHz?

Answer: $A_{max} = 50$

CHECK YOUR UNDERSTANDING

In Example 6.13, the closed-loop gain of each amplifier was assumed constant at frequencies below the cutoff frequency. In practice, the open-loop gain A of each op-amp decreases slowly with frequency at frequencies lower than the closed-loop gain cutoff frequency. The frequency response for the open-loop gain of an op-amp is well approximated by:

$$A(j\omega) = \frac{A_0}{1 + j\omega/\omega_0}$$

Use this expression to find an expression for the closed-loop gain of the cascade amplifier. (*Hint:* The combined gain is equal to the product of the individual closed-loop gains.) What is the actual gain in decibels at the cutoff frequency ω_0 for the cascade amplifier?

What is the 3-dB bandwidth of the cascade amplifier of Example 6.13? *[Hint:* The gain of the cascade amplifier is the product of the individual op-amp frequency responses. Compute the magnitude of this product, set it equal to $(1/\sqrt{2}) \times 10^4$, and solve for ω.]

Answer: 74 dB; $\omega_{3\text{ dB}} = 2\pi \times 12{,}800$ rad/s

Page 455

CHECK YOUR UNDERSTANDING

What is the maximum gain that can be accepted in the op-amp circuit of Example 6.14 if the offset is not to exceed 50 mV?

Answer: $A_{V_{max}} = 33.3$

CHECK YOUR UNDERSTANDING

Given the desired peak output amplitude (10 V), what is the maximum frequency that will not result in violating the slew rate limit for the op-amp of Example 6.16?

Answer: $f_{max} = 159$ kHz

Conclusion

Operational amplifiers are the single most important integrated circuit in analog electronics. Upon completing this chapter, the following learning objectives should have been achieved:

1. *Understand the properties of ideal amplifiers and the concepts of gain, input impedance, output impedance, and feedback.* Ideal amplifiers represent fundamental building blocks of electronic instrumentation. With the concept of an ideal amplifier clearly established, one can design practical amplifiers, filters, integrators, and many other signal processing circuits. An ideal op-amp closely approximates a practical op-amp in many respects.
2. *Understand the difference between open-loop and closed-loop op-amp configuration; and compute the gain (or complete the design of) simple inverting, noninverting, summing, and difference amplifiers using ideal op-amp analysis. Analyze more advanced op-amp circuits, using ideal op-amp analysis, and identify important performance parameters in op-amp data sheets.* Analysis of op-amp circuits is made easy by a few simplifying assumptions, which are based on the op-amp having a very large input impedance, a very small output impedance, and a large open-loop gain. The inverting and noninverting amplifier configurations permit the design of useful circuits by appropriately selecting and placing a few resistors.

3. *Analyze and design simple active filters. Analyze and design ideal integrator and differentiator circuits.* The use of capacitors in op-amp circuits extends their applications to include filtering, integration, and differentiation.
4. *Understand the structure and behavior of analog computers, and design analog computer circuits to solve differential equations.* The properties of op-amp summing amplifiers and integrators make it possible to construct analog computers that solve differential equations and simulate dynamic systems. While digital computer-based numerical simulations are readily available, there is still a role for analog computers in specialized applications.
5. *Understand the principal physical limitations of an op-amp.* It is important to understand that there are limitations in the performance of practical op-amp circuits that are not included in many simple op-amp models. It is important to consider issues related to voltage supply limits, bandwidth limits, offsets, slew rate limits, and output current limits in the design of an op-amp circuit.

Page 456

HOMEWORK PROBLEMS

Section 6.1: Ideal Amplifiers

6.1 The circuit shown in Figure P6.1 has a DC signal source, two stages of amplification, and a load. Determine, in decibels, the power gain $G = P_0/P_S = V_o\, I_o/V_S\, I_S$, where:

$R_s = 0.5\ \text{k}\Omega$ $\quad R_{o3} = 0.7\ \text{k}\Omega$
$R_{i1} = 3.2\ \text{k}\Omega$ $\quad R_{i2} = 2.8\ \text{k}\Omega$
$R_{o1} = 2.2\ \text{k}\Omega$ $\quad R_{o2} = 2.2\ \text{k}\Omega$
$A_1 = 90\ \text{V/V}$ $\quad H_2 = 300\ \text{mS}$

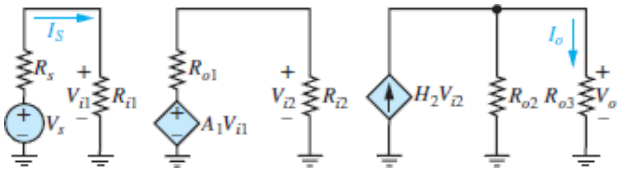


Figure P6.1

6.2 A temperature sensor in a production line under normal operating conditions produces a no-load (i.e., sensor current = 0) voltage:

$$v_s = V_{pk}\cos(\omega t) \qquad R_s = 400\ \Omega$$
$$V_{pk} = 500\ \text{mV} \qquad \omega = 6.28\ \text{krad/s}$$

The temperature is monitored on a display (the load) with a vertical line of light-emitting diodes. Normal conditions are indicated when a string of the bottommost diodes 2 cm in length is on. This requires that a voltage be supplied to the display input terminals where

$$R_o = 12\ \text{k}\Omega \qquad v_o = V_m\cos(\omega t) \qquad V_m = 6\ \text{V}$$

The signal from the sensor must be amplified.

Therefore, a voltage amplifier, shown in Figure P6.2, is connected between the sensor and CRT with

$$R_i = 2\ \text{k}\Omega \qquad R_o = 3\ \text{k}\Omega$$

Determine the required no-load gain of the amplifier.

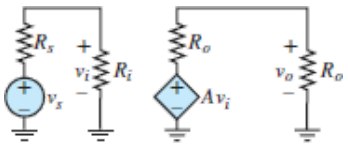


Figure P6.2

6.3 What are the golden rules of an ideal operational amplifier? What conditions do these rules rely upon?

6.4 What approximations are usually made about the circuit components and parameters of the amplifier model shown in Figure P6.4?

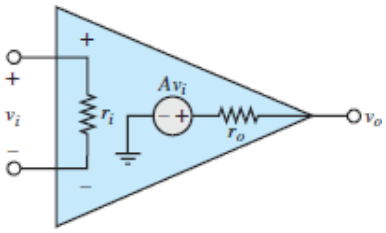


Figure P6.4

Section 6.2: The Operational Amplifier

6.5 Find v_1 in the circuits of Figure P6.5(a) and (b). In Figure P6.5(a) the 3-k Ω resistor "loads" the output; that is, v_1 is changed by attaching the 3-k Ω resistor in parallel with the lower 6-k Ω resistor. However, in Figure P6.5(b) the isolation buffer holds v_1 to $v_g/2$, regardless of the presence of the 3-k Ω resistor and its value!

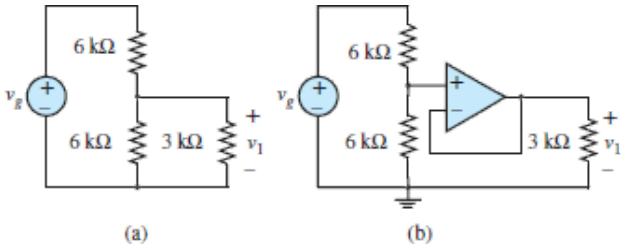


Figure P6.5

6.6 Find the current i in the circuit of Figure P6.6.

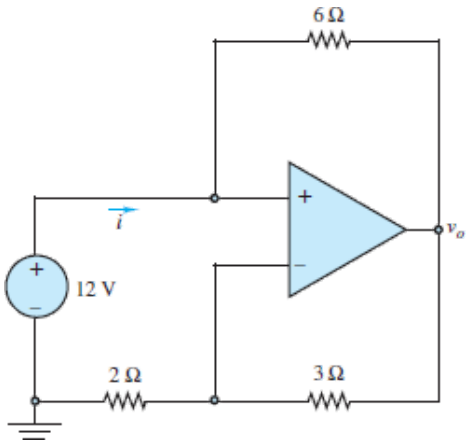


Figure P6.6

Page 457

6.7 Find the voltage v_o in Figure P6.7 by finding the Thévenin equivalent network seen to the left of nodes a and b to form an archetypical inverting amplifier.

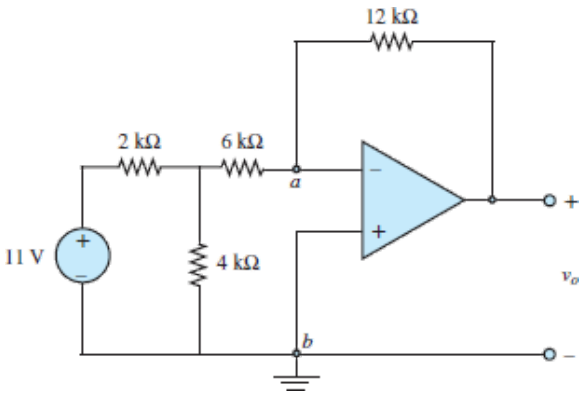


Figure P6.7

6.8 Find the Thévenin equivalent network seen between the noninverting terminal node and the reference node in Figure P6.8. Use it to find v_3 in terms of v_1 and v_2.

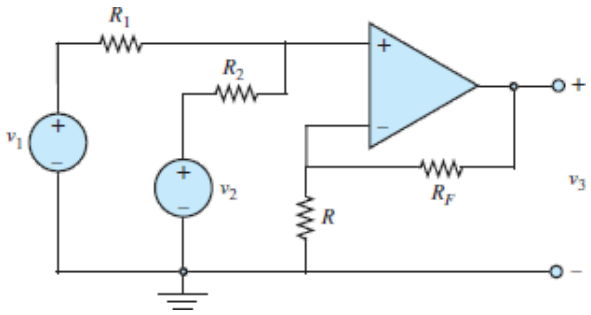


Figure P6.8

6.9 Determine an expression for the closed-loop voltage gain $G = v_o/v_1$ for the circuit of Figure P6.9. Find the input resistance v_1/i_1 seen by the voltage source.

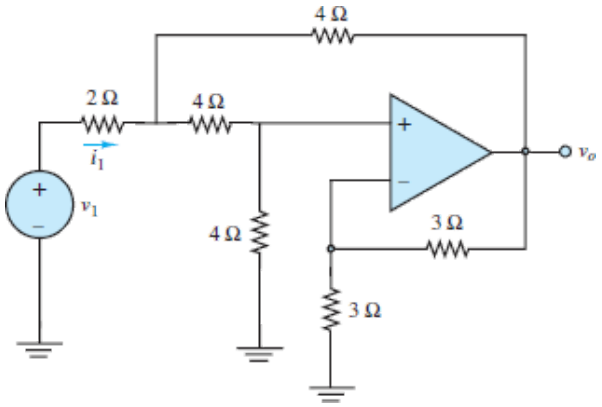


Figure P6.9

6.10 Difference amplifiers are often used in conjunction with a Wheatstone bridge, such as that shown in Figure P6.10, where each resistor is a temperature sensing element, and their change in resistance ΔR is directly proportional to their change in temperature ΔT. The constant of proportionality is the temperature coefficient $\pm\alpha$, which can be positive (PTC) or negative (NTC). Find the Thévenin equivalent network seen by the amplifier to the left of nodes *a* and *b*. Assume that $\Delta R = \pm\Delta T$ and $|\Delta R| \ll R_0$.

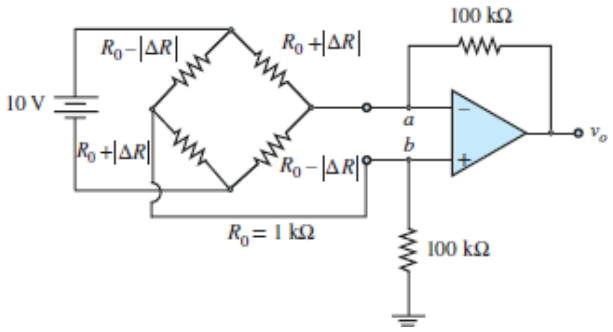


Figure P6.10

6.11 The circuit shown in Figure P6.11 is a *negative impedance converter.* Find the input impedance $\mathbf{Z}_{in}$:

$$\mathbf{Z}_{in} = \frac{\mathbf{V}_1}{\mathbf{I}_1}$$

when:

a. $\mathbf{Z}_o = R$

b. $\mathbf{Z}_o = \frac{1}{j\omega C}$

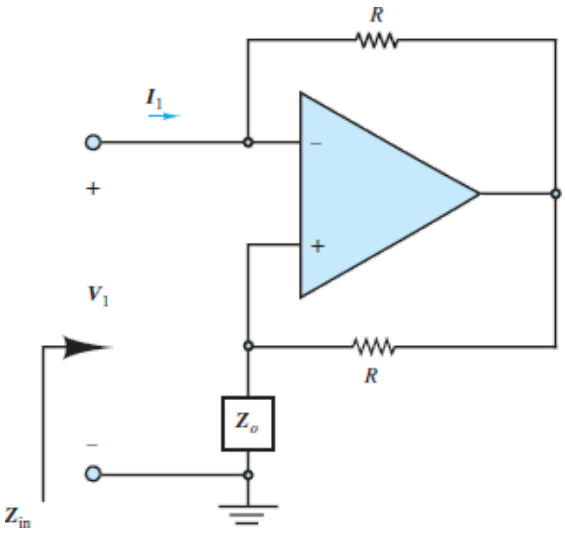


Figure P6.11

6.12 The circuit of Figure P6.12 demonstrates that op-amp feedback can create a resonant circuit without Page 458the use of an inductor. Assume $R_1 = R_2 = 1$ Ω, $C_1 = 2Q$ F, and $C_2 = 1/2Q$ F, where Q is the quality factor introduced in Chapter 5. Notice that $v_2 = v_o$ and use it and KCL to find the voltage gain v_o/v_{in}.

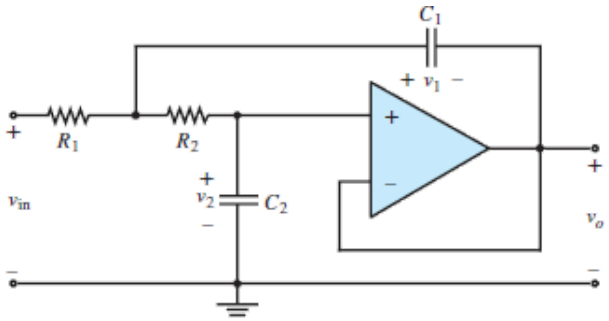


Figure P6.12

6.13 Inductors are difficult to use as components of integrated circuits due to the need for large coils of wire, which require significant space and tend to act as excellent antennas for ambient noise. As an alternative, a "solid-state inductor" can be constructed as shown in Figure P6.13.

a. Determine the input impedance $\mathbf{Z}_{in} = \boldsymbol{V}_1/\boldsymbol{I}_1$.

b. What is $\mathbf{Z}_{in}$ when $R = 1.0$ k Ω and $C = 0.02$ μF?

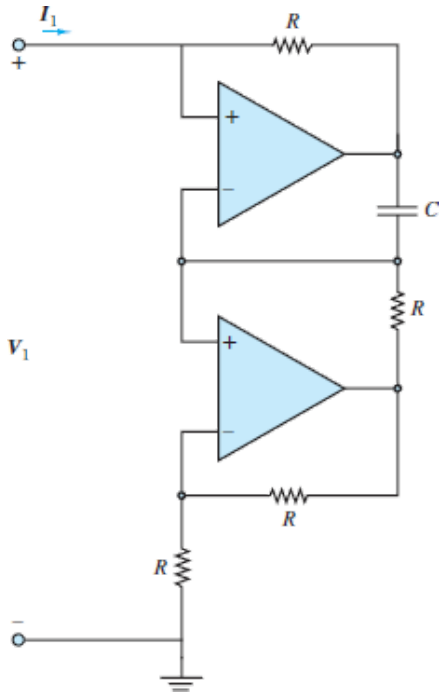


Figure P6.13

6.14 In the circuit of Figure P6.14, determine the input impedance $\mathbf{Z}_{in} = \boldsymbol{V}_1/\boldsymbol{I}_1$. Note the difference between this circuit and that shown in Figure P6.13.

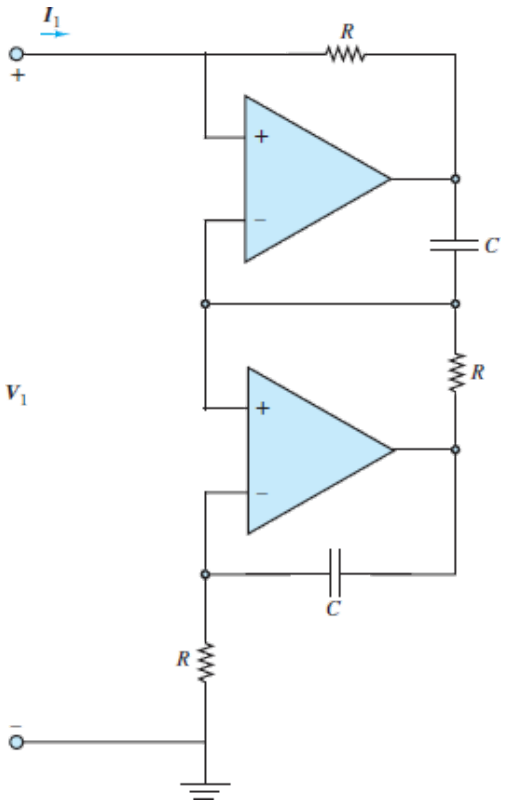


Figure P6.14

6.15 It is easy to construct a current source using an inverting amplifier configuration as shown in Figure P6.15. Verify that the current I through R_o is independent of the value of R_o , assuming that the op-amp stays in its linear operating region, and find the value of I.

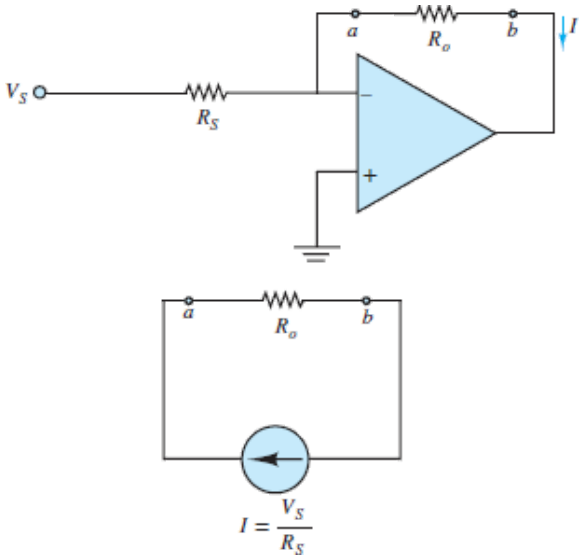


Figure P6.15

Page 459

6.16 A "super diode" or "precision diode" circuit is shown in Figure P6.16. The diode permits current from anode to cathode when the anode voltage is V_D higher than the cathode voltage, where V_D is the diode offset voltage. Current is not permitted from cathode to anode. Determine the output voltage $v_o(t)$ for the given input voltage $v_{\text{in}}(t)$. Show that the entire circuit behaves like a diode but without the offset voltage.

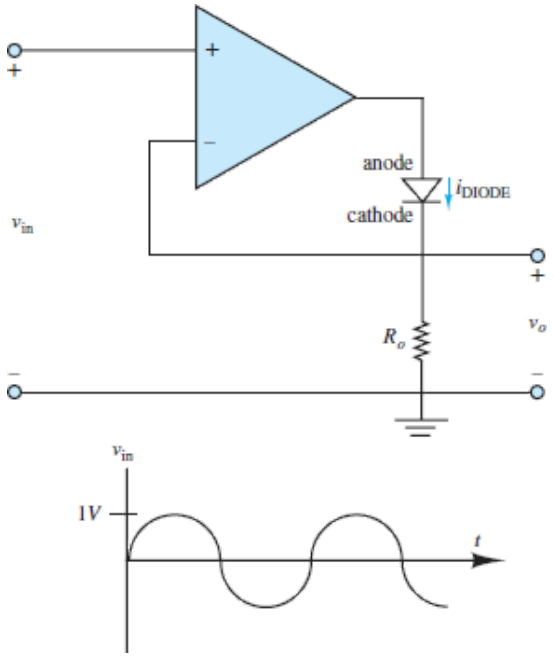


Figure P6.16

6.17 Determine the response function $\mathbf{V}_2/\mathbf{V}_1$ for the circuit of Figure P6.17.

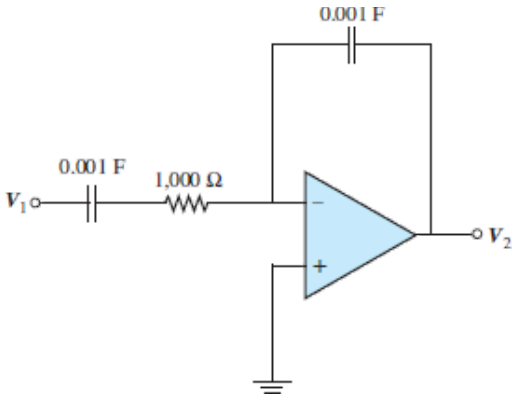


Figure P6.17

6.18 Time delays are often encountered in engineering systems. They can be approximated using Euler's definition as

$$e^{-sT} = \lim_{N\to\infty}\left[\frac{1}{sT/N+1}\right]^N$$

With the addition of a unity gain inverting amplifier show that the transfer function of the circuit shown in Figure P6.17 equals the argument of the limit given above when $s = j\omega$.

6.19 Use the circuit shown in Figure P6.17 and the result of Problem 6.18 to design a circuit whose transfer function equals the argument of the limit when $T = 1$ and $N = 4$ and thus is an approximation of a time delay.

6.20 For the circuit of Figure P6.20, apply the principle of superposition to find v_o .

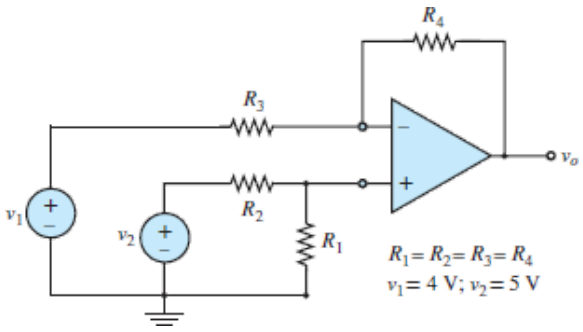


Figure P6.20

6.21 Difference amplifiers are often used in conjunction with a Wheatstone bridge, such as that shown in Figure P6.10, where each resistor is a

temperature sensing element, and their change in resistance ΔR is directly proportional to their change in temperature ΔT. The constant of proportionality is the temperature coefficient $\pm\alpha$, which can be positive (PTC) or negative (NTC). Assume $|\Delta R| = K\Delta T$, where K = constant. Find an expression for $v_o(\Delta T)$.

6.22 Consider the circuit of Figure P6.22. Assume ω = 1000 Rad/s:

a. If $\mathbf{V}_1 - \mathbf{V}_2 = 1\angle 0$V, use phasor analysis to find $|\boldsymbol{V}_o|$.

b. Use phasor analysis to find $\angle \boldsymbol{V}_o$.

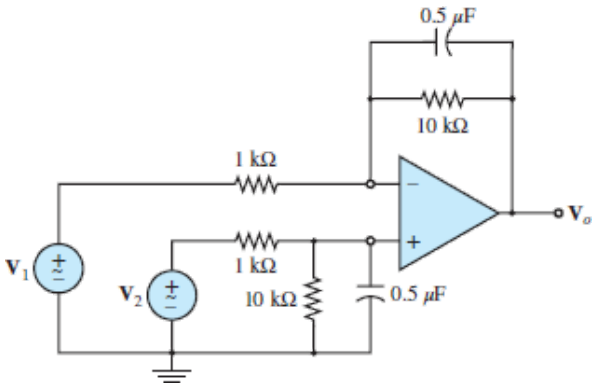


Figure P6.22

Page 460

6.23 Find an expression for the voltage gain $\mathbf{V}_o/\mathbf{V}_{in}$ of the circuit of Figure P6.12. Assume R_1 = 3 Ω, R_2 = 2 Ω, and $C_1 = C_2 = \frac{1}{6}$ F.

6.24 In the circuit of Figure P6.24, assume R_F = 12 k Ω and that it is critical that the voltage gain v_o/v_S remain within ±2 percent of the nominal gain of 20. What value of R_S is needed for the nominal gain? What are the allowed maximum and minimum values of R_S? Will a standard 5 percent tolerance resistor be adequate to satisfy this requirement? (See Table 1.3 of standard resistor values in Chapter 1.)

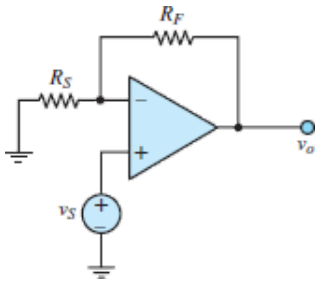


Figure P6.24

6.25 The two 5 percent tolerance resistors of an inverting amplifier (see Figure 6.8) have nominal values $R_F = 33$ k Ω and $R_S = 1.5$ k Ω.

a. What is the nominal voltage gain $G = v_o/v_S$ of the amplifier?

b. What is the maximum value of G if the resistor values can swing ±5 percent?

c. What is the minimum value of G if the resistor values can swing ±5 percent?

6.26 The circuit of Figure P6.26 is another form of a *level shifter,* which adjusts the DC portion of the input voltage $v_1(t)$ while also amplifying the AC portion. Let:

$v_1(t) = 10 + 10^{-3} \sin \omega t$ V, $R_F = 10$ k Ω, and $V_{batt} = 20$ V.

a. Find R_S such that no DC voltage appears at the output.

b. What is $v_o(t)$, using R_S from part a?

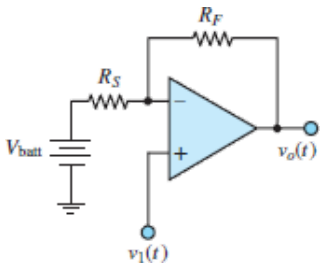


Figure P6.26

6.27 Figure P6.27 shows a simple practical amplifier that uses a 741 op-amp chip. Pin numbers are as indicated. Assume the op amp has a 2-M Ω input

resistance, an open-loop gain $A = 200{,}000$, and an output impedance $R_o = 50\ \Omega$. Find the closed-loop gain $G = v_o/v_i$.

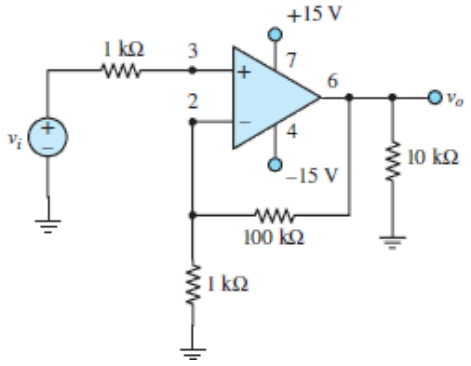


Figure P6.27

6.28 Design an inverting summing amplifier to obtain the following weighted sum of four different signal sources:

$$v_o = -(2\sin\omega_1 t + 4\sin\omega_2 t + 8\sin\omega_3 t + 16\sin\omega_4 t)$$

Assume that $R_F = 5\ \text{k}\Omega$, and determine the required source resistors.

6.29 The amplifier shown in Figure P6.29 has a signal source (v_s in series with R_s) and load R_o separated by an amplification stage built upon the Motorola MC1741C op-amp. Assume:

$$R_s = 2.2\,\text{k}\Omega \qquad R_1 = 1\,\text{k}\Omega$$
$$R_F = 8.7\,\text{k}\Omega \qquad R_o = 20\,\Omega$$

The op-amp itself has a 2-M Ω input resistance, a 75-Ω output resistance, and a 200K open-loop gain. To a first approximation, the op-amp would be modeled as ideal. A better model would include the effects of the parameters listed above. See Figure 6.6 and equation 6.23.

a. Assume the op-amp is not ideal, and derive an expression for the input resistance $r_i = v_i/i_i$ of the overall amplifier, where $v_i = v_s - i_i R_s$.

b. Determine the value of that input resistance, and compare it to the input resistance derived for an ideal op-amp.

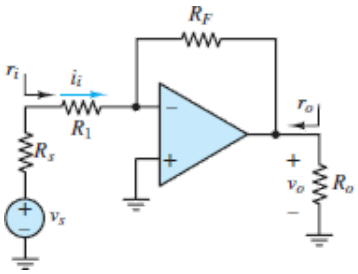


Figure P6.29

Page 461

6.30 In the circuit shown in Figure P6.30, assume $R_1 = 40$ k Ω, $R_2 = 2$ k Ω, $R_F = 150$ k Ω, $R_o = 75$ Ω and $v_s = 0.01 + 0.005 \cos(\omega t)$ V. Determine an expression for the output voltage v_o and its value.

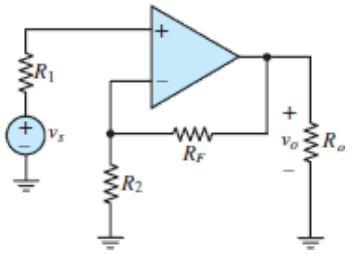


Figure P6.30

6.31 For the circuit shown in Figure P6.31, assume $v_S = 0.3 + 0.2 \cos(\omega t)$, $R_s = 4$ Ω, and $R_o = 15$ Ω. Determine the output voltage v_o for an ideal op-amp and also for a Motorola MC1741C op-amp with characteristics as given in Problem 6.29.

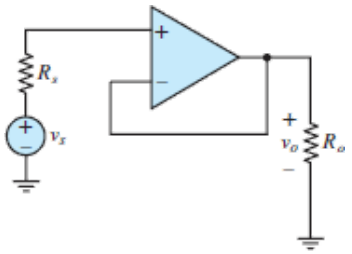


Figure P6.31

6.32 For the circuit shown in Figure P6.32, assume:

$v_{S1} = 2.9 \times 10^{-3} \cos(\omega t)$ V
$v_{S2} = 3.1 \times 10^{-3} \cos(\omega t)$ V
$R_1 = 1$ kΩ $\quad R_2 = 3.3$ kΩ
$R_3 = 10$ kΩ $\quad R_4 = 18$ kΩ $\quad R_o = 75$ Ω

Determine an expression and value for the output voltage v_o .

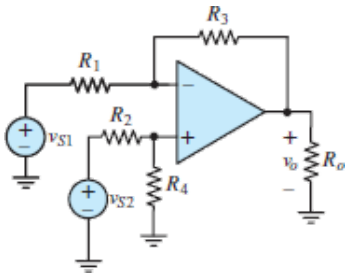


Figure P6.32

6.33 For the circuit shown in Figure P6.33, assume $v_{S1} = -2$ V, $v_{S2} = 2 \sin (2\pi \cdot 2{,}000\ t)$ V, $R_1 = 100$ k Ω, $R_2 = 50$ k Ω, $R_o = 50$Ω and $R_F = 150$ k Ω. Determine the output voltage v_o .

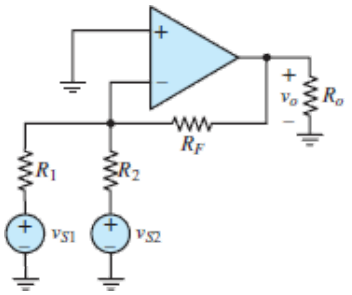


Figure P6.33

6.34 For the circuit shown in Figure P6.33, assume: $v_{S1} = v_{S2} = 5$ mV, $R_o = 75$ Ω, $R_1 = 50$ Ω, $R_2 = 2$ k Ω, and $R_F = 2$ k Ω. The nonideal MC1741C op-amp has a 2-m Ω input resistance, a 75-Ω output resistance, and an open-loop gain of 200K. (See Figure 6.6 and equation 6.23.) Determine:

a. An expression for the output voltage v_o .

b. The voltage gain for each of the two input signals.

6.35 In the circuit shown in Figure P6.35, determine the output voltage V_o. All resistances are equal and $V_{in} = 4 \angle 0$ V.

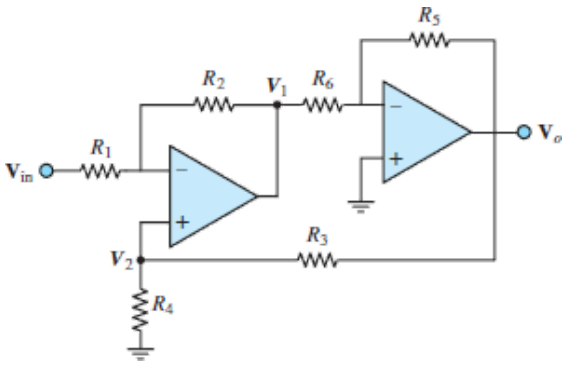


Figure P6.35

6.36 In the circuit shown in Figure P6.36, assume $V_2 = 8 \angle 0$ V and find the input voltage V_{in} such that $V_o = 0$.

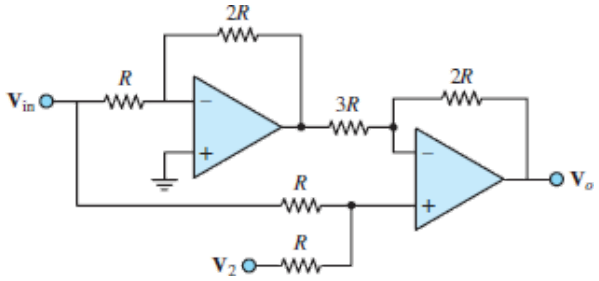


Figure P6.36

Page 462

6.37 In the circuit shown in Figure P6.32, assume:

$$v_{S1} = 1.3 \text{ V} \qquad v_{S2} = 1.9 \text{ V}$$
$$R_1 = R_2 = 4.7 \text{ k}\Omega$$
$$R_3 = R_4 = 10 \text{ k}\Omega \qquad R_o = 1.8 \text{ k}\Omega$$

Determine:

a. The output voltage v_o .

b. The common-mode component of v_o .

c. The difference-mode component of v_o .

6.38 In the circuit shown in Figure P6.38, determine the output voltage V_o . Let $R_1 = R_2 = 10\text{k } \Omega$, $R_3 = 15 \text{ k } \Omega$, $R_4 = 10 \text{ k } \Omega$, $R_F = 50 \text{ k } \Omega$, and $V_{in} = 6$ V.

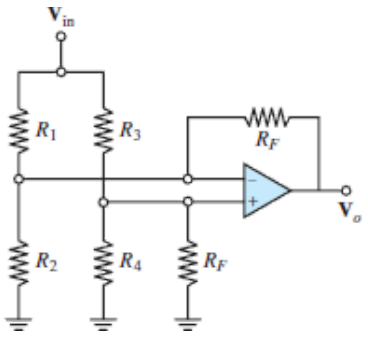


Figure P6.38

6.39 A linear potentiometer R_P is used to sense and generate a voltage v_y proportional to the y-coordinate of an xy inkjet printer head. A reference signal v_R is supplied by the software controlling the printer. The difference between these voltages is amplified to drive a motor. The motor changes the position of the printer head until that difference equals zero. For proper operation, the motor voltage must be 10 times the difference between the signal and reference voltage. For rotation in the proper direction, the motor voltage must be negative with respect to v_y . In addition, i_P must be negligibly small to avoid loading the pot and causing an erroneous signal voltage.

a. Design an op-amp circuit that satisfies these specifications. Redraw Figure P6.39, replacing the dotted line box with your amplifier circuit. Be sure to indicate component values.

b. Mark the pin numbers on your redrawn figure for an eight-pin single μA741C op-amp chip.

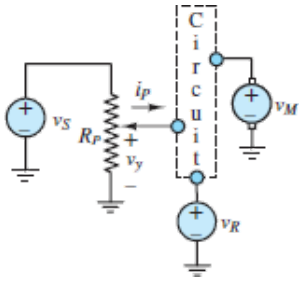


Figure P6.39

6.40 Compute the current I_{batt} delivered by the battery in Figure P6.40. Assume: $R_{S1} = R_{S2} = 30\text{k}\ \Omega$, $R_{F1} = 100\ \text{k}\ \Omega$, $R_{F2} = 60\ \text{k}\ \Omega$, $R_1 = 5\ \text{k}\ \Omega$, $R_2 = 7\ \text{k}\ \Omega$, and

$V_{batt} = 3$ V.

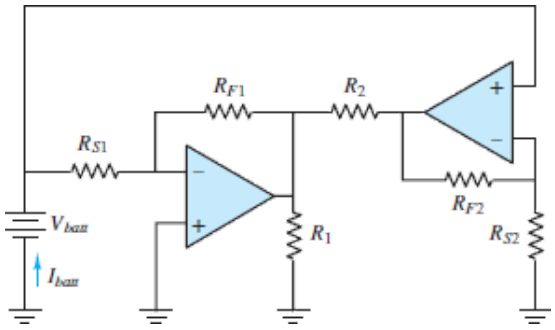


Figure P6.40

6.41 Figure P6.41 shows a simple voltage-to-current converter. Show that the current I_o through the light-emitting diode (LED), and therefore its brightness, is proportional to the source voltage V_s as long as $V_s > 0$. The LED permits current in the direction shown only.

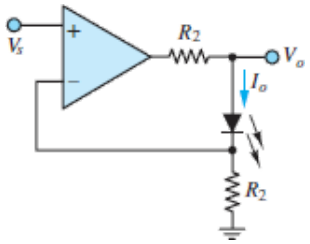


Figure P6.41

6.42 Figure P6.42 shows a simple current-to-voltage converter. Show that the voltage V_o is proportional to the current generated by the cadmium sulfide (CdS) solar cell. Also show that the transimpedance of the circuit V_o/I_s is $-R$!

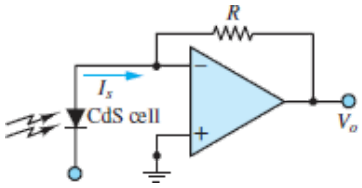


Figure P6.42

6.43 A nonideal op-amp voltmeter circuit as in Figure P6.43 is required to measure a maximum input of $V_S = 15$ mV. The op-amp input current is $I_B = 0.25\ \mu$ A. The ammeter is designed for full-scale deflection when $I_m = 80\ \mu$ A and $r_m = 8$ k Ω. Determine Page 463suitable values for R_3 and R_4 so that the full-scale deflection of the ammeter corresponds to $V_S = 15$ mV.

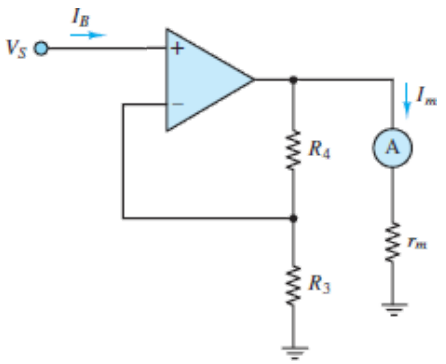


Figure P6.43

6.44 Find an expression for the voltage gain v_o/v_s in Figure P6.44. Assume $R_{S1} = R_{S2} = 2.5$ k Ω and $R_{F1} = R_{F2} = 9.0$ k Ω.

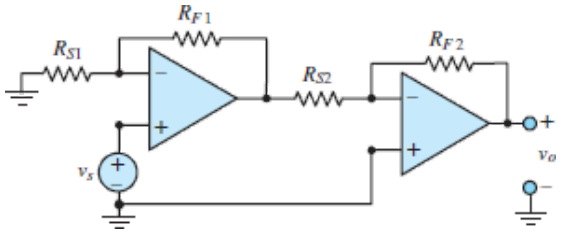


Figure P6.44

6.45 Select appropriate components using standard 5 percent resistors to obtain a voltage gain v_o/v_s as close to −80 as possible for the circuit of Figure P6.44.

6.46 For the circuit in Figure P6.44 compute the maximum and minimum possible voltage gains if the resistor values are allowed to swing ±5 percent.

6.47 The circuit shown in Figure P6.47 can function as a precision ammeter. Assume that the voltmeter has a range of 0 to 10 V and an internal resistance of 20 k Ω. The full-scale reading of the ammeter is intended to be 1 mA. Find the resistance R such that the voltmeter reading is 10 V when $i_{\text{in}} = 1$ mA.

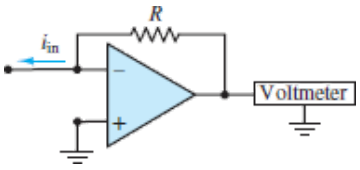


Figure P6.47

6.48 Select appropriate components using standard 5 percent resistors to obtain a voltage gain v_o/v_s as close to 20 as possible for the circuit of Figure P6.30.

6.49 For the circuit in Figure P6.30 compute the maximum and minimum possible voltage gains if the resistor values are allowed to swing ±5 percent. Use the component values listed in problem 6.30.

6.50 Select appropriate components using standard 1 percent resistors to obtain a difference gain as close to 15 as possible in the circuit of Figure P6.32. Assume that $R_3 = R_4$ and $R_1 = R_2$.

6.51 For the circuit in Figure P6.32 compute the maximum and minimum possible voltage gains if the resistor values are allowed to swing ±1 percent. Also compute the maximum common-mode output for the same allowed ±1 percent swing. Pick the nominal resistor values so that $R_3 = R_4$ and $R_1 = R_2$. Use the component values listed in problem 6.32.

Section 6.3: Active Filters

6.52 The circuit shown in Figure P6.52 with input V_s and output V_o is an active high-pass filter. Assume:

$$C = 1\ \mu\text{F} \qquad R = 10\ \text{k}\Omega \qquad R_o = 1\ \text{k}\Omega$$

Determine:

a. The voltage gain $|V_o/V_s|$ (in dB) in the passband.

b. The cutoff frequency.

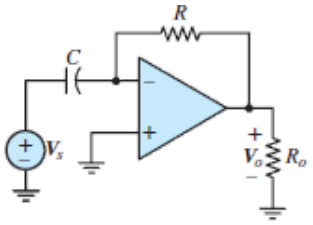


Figure P6.52

6.53 The op-amp circuit shown in Figure P6.53 is used as a high-pass filter. Assume:

$$C = 0.2\,\mu\text{F} \qquad R_o = 222\,\Omega$$
$$R_1 = 1.5\,\text{k}\Omega \qquad R_2 = 5.5\,\text{k}\Omega$$

Determine:

a. The voltage gain $|V_o/V_s|$, (in dB), in the passband.

b. The cutoff frequency.

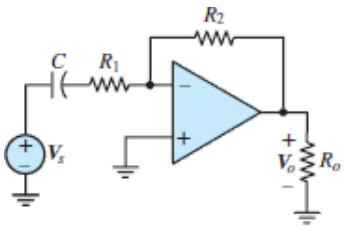


Figure P6.53

Page 464

6.54 The op-amp circuit shown in Figure P6.53 is used as a high-pass filter. Assume:

$$C = 200\ \text{pF} \qquad R_o = 1\,\text{k}\Omega$$
$$R_1 = 10\,\text{k}\Omega \qquad R_2 = 220\,\text{k}\Omega$$

Determine:

a. The voltage gain $|V_o/V_s|$, (in dB), in the passband.

b. The cutoff frequency.

6.55 The circuit shown in Figure P6.55 is an active filter. Assume:

$$C = 120\ \text{pF} \qquad R_o = 180\,\text{k}\Omega$$
$$R_1 = 3\,\text{k}\Omega \qquad R_2 = 50\,\text{k}\Omega$$

Determine the break frequencies and $|V_o/V_i|$ (in dB) at very low and at very high frequencies.

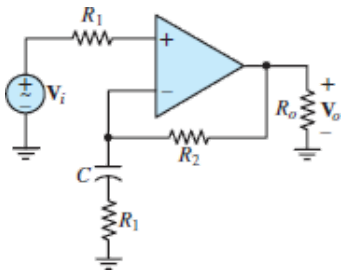


Figure P6.55

6.56 The circuit shown in Figure P6.56 is an active filter. Assume:

$$C = 15\ \text{nF} \qquad R_o = 4\ \text{k}\Omega$$
$$R_1 = 1.2\ \text{k}\Omega \qquad R_2 = 5.6\ \text{k}\Omega$$
$$R_3 = 62\ \text{k}\Omega$$

Determine:

a. An expression for the voltage gain in standard form:

$$\mathbf{G}_V(j\omega) = \frac{\mathbf{V}_o(j\omega)}{\mathbf{V}_i(j\omega)}$$

b. The break frequencies.

c. The passband gain.

d. The Bode magnitude and phase plots of $\mathbf{V}_o/\mathbf{V}_i$.

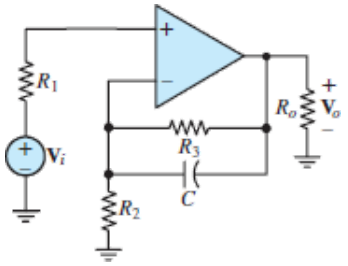


Figure P6.56

6.57 The op-amp circuit shown in Figure P6.57 is used as a low-pass filter. Assume:

$$C = 0.8\ \mu\text{F} \qquad R_o = 1\ \text{k}\Omega$$
$$R_1 = 5\ \text{k}\Omega \qquad R_2 = 15\ \text{k}\Omega$$

Determine:

a. An expression in standard form for the voltage gain $\mathbf{V}_o/\mathbf{V}_s$.

b. The gain, in dB, in the passband and at the cutoff frequency.

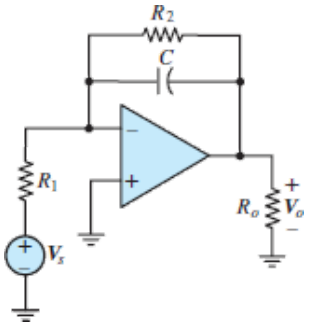


Figure P6.57

6.58 The op-amp circuit shown in Figure P6.57 is used as a low-pass filter. Assume:

$$R_1 = 2.2\ \text{k}\Omega \qquad R_2 = 68\ \text{k}\Omega$$
$$C = 0.47\ \text{nF} \qquad R_o = 1\ \text{k}\Omega$$

Determine:

a. An expression in standard form for the voltage gain $\boldsymbol{V}_o/\boldsymbol{V}_s$.

b. The gain, in dB, in the passband and at the cutoff frequency.

6.59 The circuit shown in Figure P6.59 is a bandpass filter. Assume:

$$R_1 = R_2 = 10\ \text{k}\Omega \qquad R_o = 4.7\ \text{k}\Omega$$
$$C_1 = C_2 = 0.1\ \mu\text{F}$$

Determine:

a. The voltage gain $|\ \boldsymbol{V}_o/\boldsymbol{V}_i\ |$ in the passband.

b. The resonant frequency.

c. The break frequencies.

d. The quality factor Q.

e. The Bode magnitude and phase plots of $\boldsymbol{V}_o/\boldsymbol{V}_i$.

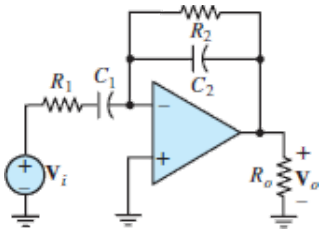


Figure P6.59

6.60 The op-amp circuit shown in Figure P6.57 is used as a low-pass filter. Assume:

$$R_1 = 12\ \text{k}\Omega \qquad R_2 = 4.7\ \text{k}\Omega \qquad R_o = 3.3\ \text{k}$$
$$C = 0.7\ \text{nF}$$

Determine:

a. An expression in standard form for the voltage gain $\mathbf{V}_o/\mathbf{V}_s$.

b. The gain, in dB, in the passband and at the cutoff frequency.

6.61 The circuit shown in Figure P6.59 is used as a bandpass filter. Assume:

$$R_1 = 2.2\ \text{k}\Omega \qquad R_2 = 100\ \text{k}\Omega$$
$$C_1 = 2.2\ \mu\text{F} \qquad C_2 = 1\ \text{nF}$$

Determine the passband voltage gain.

6.62 Derive the frequency response function $\boldsymbol{V}_o/\boldsymbol{V}_{\text{in}}$ for the circuit shown in Figure P6.62.

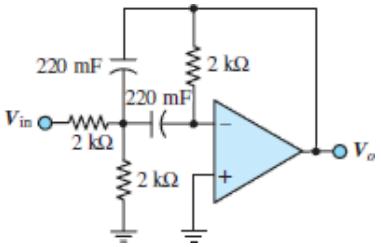


Figure P6.62

6.63 The circuit shown in Figure P6.63 can be used as a low-pass filter.

a. Derive the frequency response $\mathbf{V}_o/\mathbf{V}_{\text{in}}$ of the circuit.

b. If $R_1 = R_2 = 100$ k Ω and $C = 0.1$ μ F, compute the attenuation, in dB, of $\mathbf{V}_o/\mathbf{V}_{in}$ at $\omega = 1{,}000$ rad/s.

c. Compute the amplitude and phase of $\mathbf{V}_o/\mathbf{V}_{in}$ at $\omega = 2{,}500$ rad/s.

d. Find the range of frequencies over which the attenuation of $\mathbf{V}_o/\mathbf{V}_{in}$ is less than 1 dB.

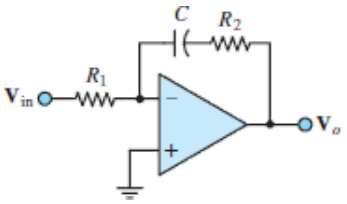


Figure P6.63

6.64 Determine a symbolic expression in standard form for the voltage gain $\mathbf{V}_o/\mathbf{V}_{in}$ in Figure P6.64. What kind of a filter does the voltage gain represent?

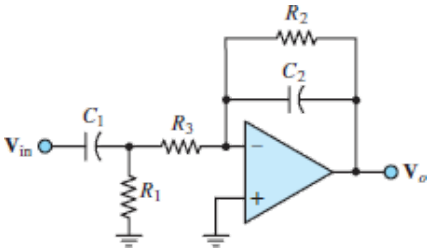


Figure P6.64

6.65 For the circuit of Figure P6.65, sketch the amplitude response of $\mathbf{V}_2/\mathbf{V}_1$, indicating the half-power frequencies.

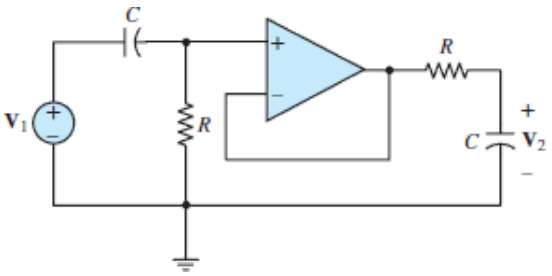


Figure P6.65

6.66 Determine a symbolic expression for the voltage gain $\mathbf{V}_o/\mathbf{V}_{S1}$ of Figure P6.66. What kind of a filter does the gain represent?

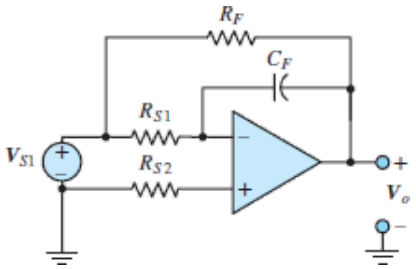


Figure P6.66

Page 466

6.67 Determine a symbolic expression for the voltage gain $\mathbf{V}_o/\mathbf{V}_S$ of Figure P6.67. What kind of a filter does the gain represent?

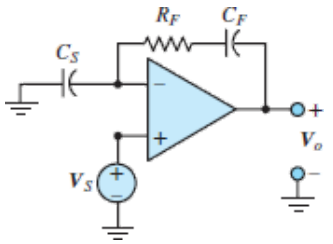


Figure P6.67

6.68 Replace the passband specification of Example 6.5 with ω_C = 10 rad/s, and determine the order of the filter required to achieve 40-dB attenuation at ω_S = 24 rad/s.

6.69 The circuit of Figure P6.69 acts as an active low-pass filter.

a. Derive the relationship between output amplitude and input amplitude.

b. Derive the relationship between output phase angle and input phase angle.

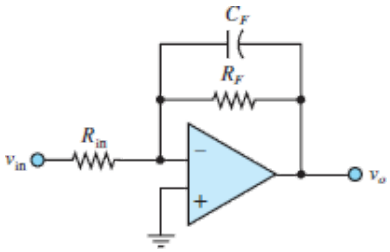


Figure P6.69

6.70 Consider the circuit of Figure P6.69. Let R_{in} = 20 k Ω, R_F = 100 k Ω, and C_F = 100 pF. Determine an expression for $v_o(t)$ if $v_{in}(t) = 2 \sin(2{,}000\ \pi\ t)$ V.

6.71 Derive the frequency response of the low-pass filter of Figure 6.45.

6.72 Derive the frequency response of the high-pass filter of Figure 6.45.

6.73 Derive the frequency response of the bandpass filter of Figure 6.45.

6.74 Consider the circuit of Figure P6.69. Let C_F = 100 pF. Determine appropriate values for R_{in} and R_F to produce a cutoff frequency of 20 kHz and a gain magnitude of 5.

Section 6.4: Design of Active Filters

6.75 Design a second-order Butterworth high-pass filter with a 10-kHz cutoff frequency, a DC gain of 10, $Q = 5$, and $V_S = \pm 15$ V.

6.76 Design a second-order Butterworth high-pass filter with a 25-kHz cutoff frequency, a DC gain of 15, $Q = 10$, and $V_s = \pm 15$ V.

6.77 The circuit shown in Figure P6.77 is claimed to exhibit a second-order Butterworth low-pass voltage gain characteristic. Derive the characteristic and verify the claim.

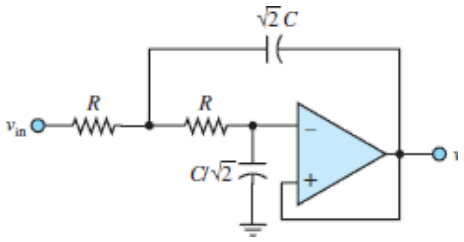


Figure P6.77

6.78 Design a second-order Butterworth low-pass filter with a 15-kHz cutoff frequency, a DC gain of 15, $Q = 5$, and $V_S = \pm 15$ V.

6.79 Design a bandpass filter with a low cutoff frequency of 200 Hz, a high cutoff frequency of 1 kHz, and a passband gain of 4. Calculate the value of Q for the filter. Also draw the approximate frequency response of this filter.

6.80 Using the circuit of Figure P6.77, design a second-order low-pass Butterworth filter with a cutoff frequency of 10 Hz.

6.81 A low-pass Sallen and Key filter is shown in Figure P6.81. Find the voltage gain $\boldsymbol{V_o}/\mathbf{V}_{\text{in}}$ as a function of frequency and generate its Bode magnitude plot. Show and observe that the cutoff frequency is $1/2\pi RC$ and that the low-frequency gain is R_4/R_3.

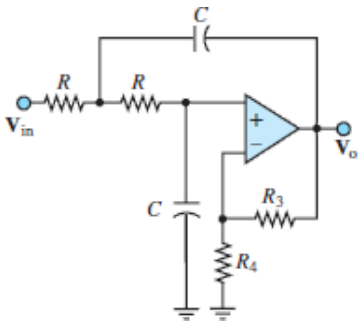


Figure P6.81

Page 467

6.82 The circuit shown in Figure P6.82 exhibits low-pass, high-pass, and bandpass voltage gain characteristics, depending on whether the output is taken at node 1, node 2, or node 3. Find the transfer functions relating each of these outputs to $\mathbf{V}_{\text{in}}$, and determine which is which.

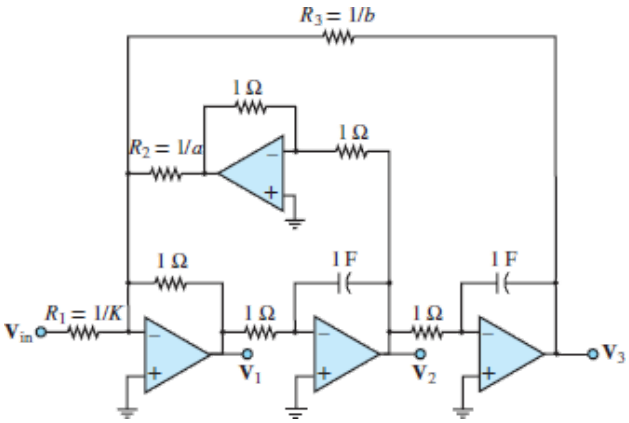


Figure P6.82

6.83 The filter shown in Figure P6.83 is called an *infinite-gain multiple-feedback filter*. Derive the following expression for the filter's frequency response $\mathbf{V}_o/\mathbf{V}_i$.

$$\frac{-(1/R_3R_2C_1C_2)R_3/R_1}{(j\omega)^2 + \left(\frac{1}{R_1C_1} + \frac{1}{R_2C_1} + \frac{1}{R_3C_1}\right)j\omega + \frac{1}{R_3R_2C_1C_2}}$$

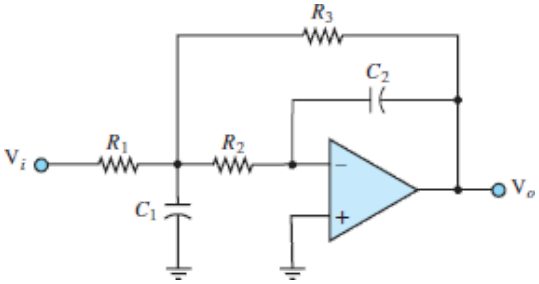


Figure P6.83

6.84 The filter shown in Figure P6.84 is a Sallen and Key bandpass filter circuit, where K is the DC gain of the filter. Derive the following expression for the filter's frequency response $\mathbf{V}_o/\mathbf{V}_i$.

$$\frac{j\omega K/R_1C_1}{(j\omega)^2 + j\omega\left(\frac{1}{R_1C_1} + \frac{1}{R_3C_2} + \frac{1}{R_3C_1} + \frac{1-K}{R_2C_1}\right) + \frac{R_1+R_2}{R_1R_2R_3C_1C_2}}$$

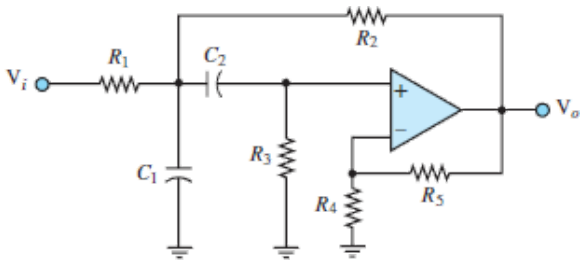


Figure P6.84

6.85 Show that the expression for Q in the filter of Problem 6.83 is given by

$$\frac{1}{Q} = \sqrt{R_2 R_3 \frac{C_2}{C_1}}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)$$

Section 6.5: Integrators and Differentiators

6.86 The circuit shown in Figure P6.86(a) produces an output voltage v_o which is the derivative of the source voltage v_s shown in Figure P6.86(b) multiplied by some gain. Assume:

$$C = 1.5\ \mu\text{F} \qquad R = 5\ \text{k}\Omega \qquad R_o = 1.5\ \text{k}\Omega$$

Determine the output voltage as a function of time and plot it.

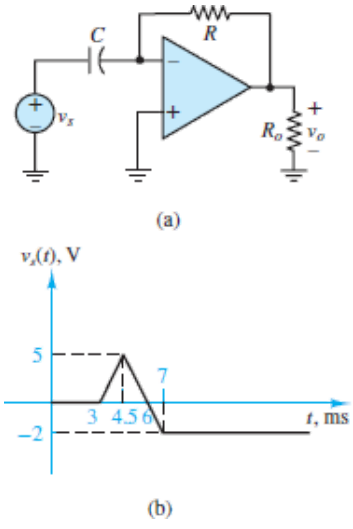


Figure P6.86

6.87 The circuit shown in Figure P6.87(a) produces an output voltage v_o which is either the integral or the derivative of the source voltage v_s shown in Figure P6.87(b) multiplied by some gain. Assume:

$$C = 0.5\,\mu\text{F} \qquad R = 8\,\text{k}\Omega \qquad R_o = 2\,\text{k}\Omega$$

Page 468

For the given source voltage, determine the output voltage as a function of time and plot it.

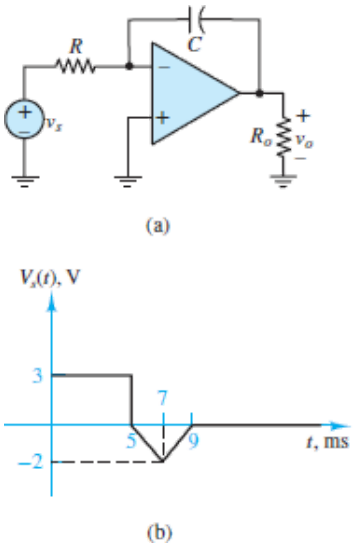


Figure P6.87

6.88 The circuit shown in Figure P6.88 is an integrator. The capacitor is initially uncharged, and the source voltage is

$$v_{\text{in}}(t) = 10^{-2}V + \sin(2{,}000\pi t)\ \text{V}$$

a. At $t = 0$, the switch S_1 is closed. How long does it take before clipping occurs at the output if $R_s = 10$ k Ω and $C_F = 0.008$ μ F?

b. At what times does the integration of the DC input cause the op-amp to saturate fully?

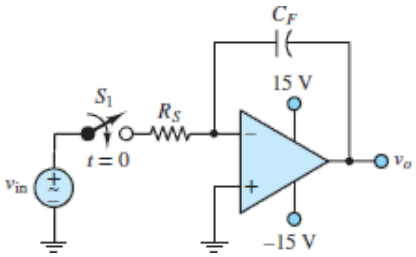


Figure P6.88

6.89 A practical integrator is shown in Figure 6.35. Note that the resistor in parallel with the feedback capacitor provides a path for the capacitor to discharge DC voltage. Usually, the time constant $R_F C_F$ is chosen to be large enough not to interfere with the integration.

a. If R_S = 10 k Ω, R_F = 2 M Ω, C_F = 0.008 μ F, and $v_S(t)$ = 10 V + sin(2,000 π t) V, find $v_o(t)$, using phasor analysis.

b. Repeat part a if R_F = 200 k Ω and if R_F = 20 k Ω.

c. Compare the time constants $R_F C_F$ with the period of the waveform for parts a and b. What can you say about the time constant and the ability of the circuit to integrate?

6.90 The circuit of Figure 6.40 is a practical differentiator. Assume an ideal op-amp, and $v_S(t) = 10^4 \sin(2{,}000\pi t)$ mV, C_S = 100 μ F, C_F = 0.008 μ F, R_F = 2 M Ω, and R_S = 10 k Ω.

a. Determine the voltage gain $\boldsymbol{V_o/V_S}$.

b. Sum the DC and AC components of $v_o(t)$ to find the total output voltage.

6.91 Derive the differential equation in $x(t)$ for the circuit of Figure P6.91.

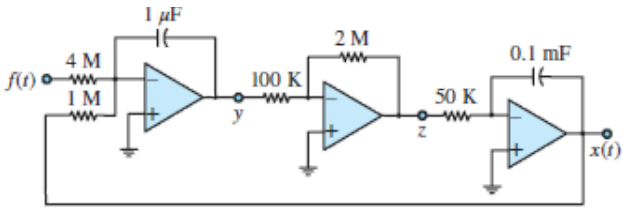


Figure P6.91

6.92 Construct a circuit corresponding to the following differential equation:

$$\frac{d^2x}{dt^2} + 100\frac{dx}{dt} + 10x = -5f(t)$$

Section 6.6: Physical Limitations of Operational Amplifiers

6.93 Consider the noninverting amplifier of Figure 6.65. Find V_o when the op-amp has an input offset voltage of 2 mV. Assume the input bias currents are zero and $R_1 = R_F = 4.7$ k Ω.

6.94 In the circuit shown in Figure P6.94, sketch the output voltage $v_o(t)$ for the two input voltages $v_1(t)$ and $v_2(t)$. Assume R_1 = 120 k Ω, R_2 = 150 k Ω, and C = 2 nF. Also assume the op-amp slew rate limit is S_0 = 1.0 V/μs and the capacitor is initially uncharged.

Page 469

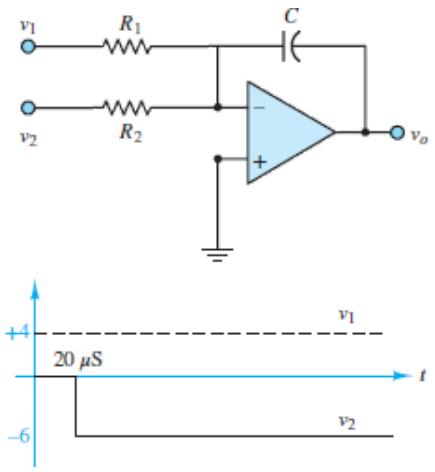


Figure P6.94

6.95 Consider a standard inverting amplifier, as shown in Figure P6.95. Assume that the offset voltage can be neglected and that the two input bias currents are equal. Find the relationship between >R_{30}, R_1 and R_{30} that eliminates the error in the output voltage due to the bias currents.

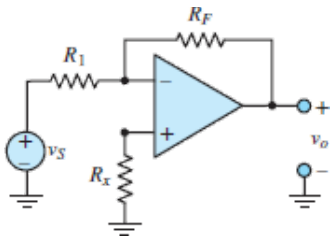


Figure P6.95

6.96 Determine the effect of the slew rate limit $S_0 = 0.5$ V/μs on the output of the unity-gain isolation buffer shown in Figure 6.10 for each of the following sinusoidal input voltages:

a. $v_S = 0.8 \sin (2\pi \cdot 6{,}000\, t)$V

b. $v_S = 0.9 \sin (2\pi \cdot 7{,}500\, t)$V

c. $v_S = 0.9 \sin (2\pi \cdot 15{,}000\, t)$V

6.97 In the circuit shown in Figure P6.97, derive the output voltage $v_o(t)$ as a function of $v_{in}(t)$. Assume the data given for S_0 and v_{in} in Problem 6.96. Also assume that $R_{30} = 15$ kΩ and $C = 0.8$ μF. For each of the three expressions of v_{in}, determine the maximum value of the low frequency gain R_{30}/R_{30} such that the slew rate limit is not exceeded.

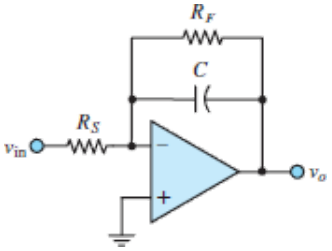


Figure P6.97

6.98 Determine the effect of the slew rate limit $S_0 = 0.5$ V/μ s on the output of a noninverting amplifier with closed-loop voltage gain $G = 10$ for a symmetric square wave v_{in} . Sketch the output waveform for each following case:

a. v_{in} switches between ±0.5 V and f = 500 Hz.

b. v_{in} switches between ±1.25 V and f = 5 kHz.

c. v_{in} switches between ±0.5 V and f = 25 kHz.

6.99 Consider a difference amplifier with a desired common-mode output of less than 1 percent of the difference-mode output. See Figure 6.18, equation 6.59 and the discussion on common and difference modes. Find the minimum decibel common-mode rejection ratio to fulfill this requirement if the differential-mode gain A_{dm} = 1,000. Let

$$v_1 = \sin(2{,}000\pi t) + 0.1\sin(120\pi t)\ \text{V}$$
$$v_2 = \sin(2{,}000\pi t + 180°) + 0.1\sin(120\pi t)\ \text{V}$$
$$v_o = A_{DM}(v_1 - v_2) + A_{CM}\frac{v_1 + v_2}{2}$$

6.100 Square wave testing can be used with operational amplifiers to estimate the *slew rate,* which is defined as the maximum rate at which the output can change (in volts per microsecond). Input and output waveforms for a noninverting op-amp circuit are shown in Figure P6.100. As indicated, the rise time t_R of the output waveform is defined as the time it takes for that waveform to increase from 10 percent to 90 percent of its final value, or

$$t_R \triangleq t_B - t_A = -\tau(\ln 0.1 - \ln 0.9) = 2.2\tau$$

where τ is the circuit time constant. Derive this expression and estimate the slew rate for the op-amp.

Page 470

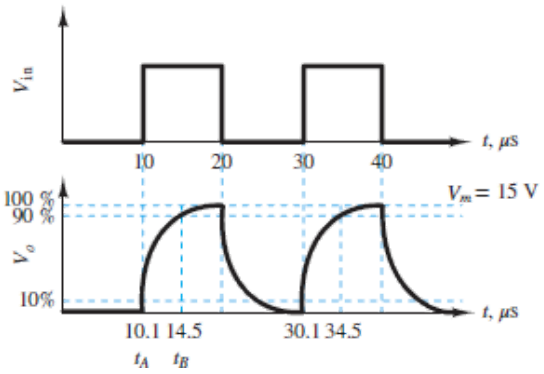


Figure P6.100

6.101 The nonideal op-amp used in the inverting amplifier of Figure 6.8 has an open-loop voltage gain $A = 250 \times 10^3$. Assume that v^- is small but

nonzero. The input terminal currents i_{in} can still be assumed zero. Apply equation 6.23 to find:

$$\frac{v_o}{v_S} = \frac{-R_F/R_S}{1 + (1/A)[(R_F + R_S)/R_S]}$$

a. If $R_S = 10$ k Ω and $R_F = 1$ M Ω, find the closed-loop voltage gain $G = v_o/v_S$.

b. Repeat part a for $R_F = 10$ M Ω.

c. Repeat part a for $R_F = 100$ M Ω.

d. Evaluate G as $A \to \infty$ for parts a to c.

6.102 A nonideal op-amp used in the noninverting amplifier of Figure P6.102 has an open-loop voltage gain $A = 250 \times 10^3$. Assume $v_{in} = v^- + \Delta v$, where Δv is small but nonzero, as suggested in equation 6.23. The input terminal currents i_{in} can still be assumed zero. Find:

a. The closed-loop gain v_o/v_{in} for $R_F = R_S = 7.5$ k Ω;

b. The closed-loop gain v_o/v_{in} for $R_F = R_S = 7.5$ k Ω.

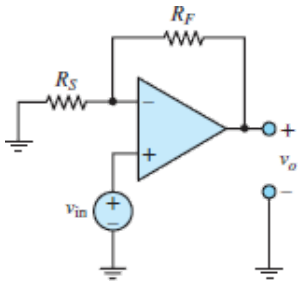


Figure P6.102

6.103 Given the unity-gain bandwidth for an ideal op-amp equal to 5.0 MHz, find the voltage gain at a frequency of $f = 500$ kHz.

6.104 The open-loop gain A of real (nonideal) op-amps is very large at low frequencies but decreases markedly as frequency increases. As a result, the closed-loop gain of op-amp circuits can be strongly dependent on frequency. Determine the frequency dependent relationship between the finite and frequency-dependent open-loop gain $A(j\omega)$ expressed in equation 6.102 and the closed-loop gain $G(\omega)$ of the inverting amplifier

shown in Figure 6.8. Plot G versus ω. Notice that $-R_F/R_S$ is the low-frequency closed-loop gain.

6.105 A sinusoidal sound (pressure) wave $p(t)$ impinges upon a condenser microphone of sensitivity S. The voltage output of the microphone v_s is amplified by two cascaded inverting amplifiers to produce an amplified signal v_0. Determine the peak amplitude of the sound wave (in dB) if $v_0 = 5\ V_{RMS}$. Estimate the maximum peak magnitude of the sound wave in order that v_0 not contain any saturation effects of the op-amps. Assume $S = 10.0$ mV/Pa, $G = 5$ for each amplifier and $V^+ = -V^- = 12$ V.

6.106 For the circuit shown in Figure P6.106, assume a nonideal op-amp and:

$$v_{S1} = 2.8 + 0.01\cos(\omega t) \quad \text{V}$$
$$v_{S2} = 3.5 - 0.007\cos(\omega t) \quad \text{V}$$
$$R_2 = 1.0\ \text{k}\Omega \quad R_F = 100.0\ \text{k}\Omega \quad \omega = 4\ \text{krad/s}$$

Refer to equations 6.23 and 6.66–6.70 to determine for each of the following op-amp open-loop gains of 10^6, 10^4 and 10^2 the:

a. Common- and difference-mode input signals.

b. Common- and difference-mode gains A_{CM} and A_{DM}, respectively.

c. Common- and difference-mode components of the output voltage.

d. Total output voltage.

e. Common-mode rejection ratio (CMRR), in dB.

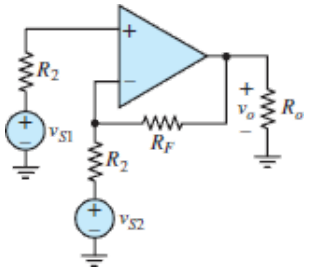


Figure P6.106

6.107 For the circuit shown in Figure P6.106, assume a nonideal op-amp and:

$$v_{S1} = 3.5 + 0.01\cos(\omega t) \quad \text{V}$$
$$v_{S2} = 3.5 - 0.01\cos(\omega t) \quad \text{V}$$
$$A_{CM} = 10\ \text{dB} \quad A_{DM} = 20\ \text{dB} \quad \omega = 4\ \text{krad/s}$$

where A_{CM} and A_{DM} are the common- and difference-mode open-loop voltage gains, respectively. Refer to Page 471equations 6.23 and 6.66–6.70 to determine for each of the following op-amp open-loop gains of 10^6, 10^4 and 10^2 the:

a. Common- and difference-mode input voltages.

b. The individual voltage gains for v_{S1} and v_{S2}.

c. The common- and difference-mode components of the output voltage.

d. The common-mode rejection ratio (CMRR), in dB.

e. The individual closed-loop voltage gains for v_{S1} and v_{S2}.

6.108 In the circuit shown in Figure P6.108, the two voltage sources are temperature sensors with T = temperature (Kelvin) and

$$v_{S1} = kT_1 \qquad v_{S2} = kT_2$$

where

$$k = 120\ \mu\text{V/K}$$
$$R_1 = R_3 = R_4 = 5\ \text{k}\Omega$$
$$R_2 = 3\ \text{k}\Omega \qquad R_o = 600\ \Omega$$

If

$$T_1 = 310\ \text{K} \qquad T_2 = 335\ \text{K}$$

determine

a. The voltage gains for the two input voltages.

b. The common-mode and difference-mode input voltages.

c. The common-mode and difference-mode gains.

d. The common-mode component and the difference-mode component of the output voltage.

e. The common-mode rejection ratio (CMRR), in dB.

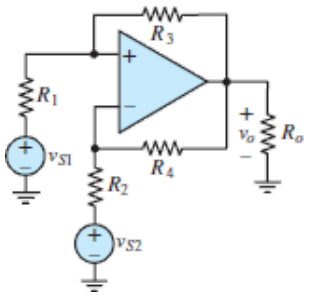


Figure P6.108

6.109 In the difference amplifier shown in Figure P6.108,

$v_{S1} = 13$ mV $\quad v_{S2} = 9$ mV

$v_o = v_{oc} + v_{od}$

$v_{oc} = 33$ mV (common-mode output voltage)

$v_{od} = 18$ V (difference-mode output voltage)

Determine

a. The common-mode gain.

b. The difference-mode gain.

c. The common-mode rejection ratio, in dB.

6.110 The ideal charge amplifier discussed in the Focus on Measurements box, "Charge Amplifiers," will saturate in the presence of any DC offsets. Figure P6.110 presents a practical charge amplifier in which the user is provided with a choice of three time constants—RC_F, $10RC_F$, and $100RC_F$—which can be selected by means of a switch. Assume that $R = 0.1$ MΩ, and $C_F = 0.1$ μ F. Analyze the frequency response of the practical charge amplifier for each time constant, and determine the lowest input signal frequency that can be amplified without excessive distortion for each case. Can this circuit amplify a DC signal?

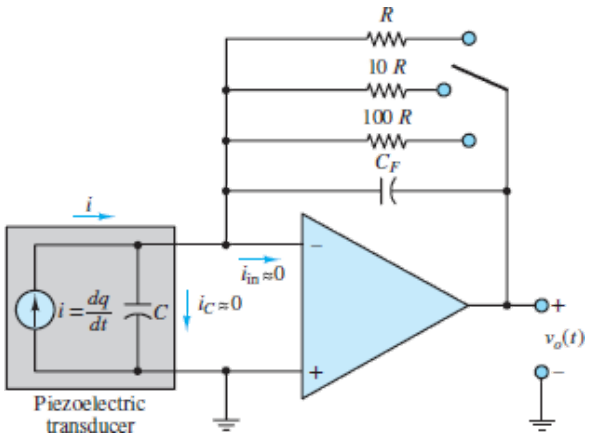


Figure P6.10

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]The voltage gain G and the open-loop gain A may also be designated as A_V and AVOL, respectively. Electrical conductance is also designated as G; as always, it is important to correctly interpret a symbol from the context in which it is used. Happily, conductance G is rarely used in engineering work. Its inverse, resistance R, is preferred instead.

[2]The operational amplifier of Figure 6.6 is a *voltage amplifier*; another type of operational amplifier, called a *current* or *transconductance amplifier*, is described in the homework problems.

[3]Special op-amps are employed to achieve extremely high input impedance, through FET input circuits. See Chapter 10.

C H A P T E R
7

ELECTRONIC INSTRUMENTATION AND MEASUREMENTS

Measurement and instrumentation systems are indispensable to engineers and scientists. While these systems are often packaged and sold as plug-and-play devices, many times it is necessary to understand their detailed specifications to properly interpret the generated data and to detect and correct errors in that data. This chapter follows a logical thread, starting with physical sensors, proceeding (in order) through wiring, grounding, signal conditioning, analog-to-digital conversion, and digital data transmission.

Section 7.1 presents an overview of sensors commonly used in engineering measurements. Some sensing devices have already been covered in earlier chapters, and others will be discussed in later chapters; the main emphasis in this chapter is on classifying physical sensors and on providing additional details not presented elsewhere in this book—most notably, temperature transducers. Section 7.2 describes the common signal connections and proper wiring and grounding techniques, with emphasis on noise sources and techniques for reducing undesired interference. Section 7.3 provides an essential introduction to analog signal conditioning, namely, a discussion of instrumentation amplifiers and active

filters. Sections 7.4 through 7.6 introduce analog-to-digital conversion, other integrated circuits used in instrumentation systems, and digital data transmission, respectively.

Page 474

Learning Objectives

Students will learn to...

1. Review the major classes of sensors. *Section 7.1*.
2. Learn how to properly ground circuits, and learn methods for noise shielding and reduction. *Section 7.2*.
3. Design signal conditioning amplifiers. *Section 7.3*.
4. Understand A/D and D/A conversion and select the specifications of the appropriate conversion system for a given application. *Section 7.4*.
5. Analyze and design simple comparator and timing circuits using integrated circuits. Review other common instrumentation integrated circuits. *Sections 7.5 and 7.6.* (Section 7.6> may be found on the book website.)

7.1 MEASUREMENT SYSTEMS AND TRANSDUCERS

Measurement Systems

In virtually every engineering application there is a need to measure physical quantities, such as forces, stresses, temperatures, pressures, flows, or displacements. These measurements are performed by **sensors** or **transducers**, which are capable of converting one type of quantity into another. Most sensors convert the quantity to be measured (e.g., humidity, temperature) to a corresponding electrical quantity (e.g., voltage or current). Often the electrical output of the sensor requires additional manipulation before it is in a useful form. For example, the change in resistance resulting from a change in the surface stresses of a material—the quantity measured by the resistance strain gauges described in Chapter 2[1]—must be first converted to a change in voltage through a suitable circuit (the Wheatstone bridge) and then amplified from the millivolt to the volt level. The manipulations needed to produce the desired end result are referred to as *signal conditioning*. The wiring of the sensor to the signal

conditioning circuitry requires significant attention to *grounding* and *shielding* procedures, to ensure that the resulting signal is as free from noise and interference as possible. Very often, the conditioned sensor signal is then converted to *digital* form and recorded in a computer for additional manipulation or is displayed in some form. The apparatus used in manipulating a sensor output to produce a result that can be suitably displayed or stored is called a **measurement system**. Figure 7.1 depicts a typical measurement system in block diagram form.

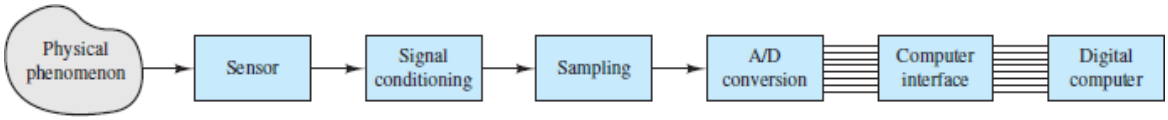


Figure 7.1 Measurement system

Page 475

Sensor Classification

Sensors may be grouped according to their physical characteristics (e.g., active electronic sensors, passive resistive sensors) or by the quantity measured by the sensor (e.g., temperature, flow rate). Other classifications are also possible. Table 7.1 presents a partial classification of sensors grouped according to the quantity sensed. Most measurements of interest to an engineer are likely to fall in the categories listed in Table 7.1. The table includes references to sensors described in the Focus on Measurement sections.

Table 7.1 Sensor classification

Sensed variables	Sensors	Chapter reference
Motion and dimensional variables	Resistive potentiometers	Resistive Throttle Position Sensor (Chapter 1)
	Strain gauges	Resistance Strain Gauges, The Wheatstone Bridge; and Force Measurements (Chapter 1)
	Differential transformers (LVDTs)	Linear Variable Differential Transformer (LVDT) (Chapter 18, online)
	Variable-reluctance sensors	Magnetic Reluctance Position Sensor (Chapter 18, online)
	Capacitive sensors	Capacitive Displacement Transducer and Microphone (Chapter 3); Peak Detector Circuit for Capacitive Displacement Transducer (Chapter 8)
	Piezoelectric sensors	Piezoelectric Sensor and Charge Amplifiers (Chapter 6)
	Electro-optical sensors	Digital Position Encoders; Digital Measurement of Angular Position and Velocity (Chapter 11)
	Moving-coil transducers	Seismic Transducer (Chapter 5)
	Seismic sensors	Seismic Transducer (Chapter 5)
Force, torque, and pressure	Strain gauges	Resistance Strain Gauges, The Wheatstone Bridge; and Force Measurements (Chapter 1)
	Piezoelectric sensors	Piezoelectric Sensor and Charge Amplifiers (Chapter 6)
	Capacitive sensors	Capacitive Displacement Transducer and Microphone (Chapter 3); Peak Detector Circuit for Capacitive Displacement Transducer (Chapter 8)
Flow	Pitot tube	
	Hot-wire anemometer	Hot-Wire Anemometer (Chapter 7)
	Differential pressure sensors	Differential Pressure Sensor (Chapter 7)
	Turbine meters	Turbine Meters (Chapter 7)
	Vortex shedding meters	
	Ultrasonic sensors	
	Electromagnetic sensors	
	Imaging systems	
Temperature	Thermocouples	Thermocouples (Chapter 7)
	Resistance thermometers (RTDs)	Resistance Thermometers (RTDs) (Chapter 7)
	Semiconductor thermometers	Diode Thermometer (Chapter 8)
	Radiation detectors	
Liquid level	Motion transducers	
	Force transducers	
	Differential pressure measurement devices	
Humidity	Semiconductor sensors	
Chemical composition	Gas analysis equipment	
	Solid-state gas sensors	

Page 476

A sensor is usually accompanied by a set of specifications that indicate its overall effectiveness. Some of these specifications are defined below:

Accuracy: conformity of a measurement to the true value, usually in percent of full-scale reading

Error: difference between measurement and true value, usually in percent of full-scale reading

Precision: the degree to which the value of a measurement can be reliably reproduced, usually expressed in bits or significant figures

Resolution: smallest measurable increment

Span: linear operating range

Range: the range of measurable values

Linearity: conformity to an ideal linear calibration curve, usually in percent of reading or of full-scale reading (whichever is greater)

Motion and Dimensional Measurements

Motion and dimension are perhaps the most commonly measured engineering quantities, including absolute position, relative position (displacement), velocity, acceleration, and jerk (the derivative of acceleration). These can be either translational or rotational measurements. These measurements are often made by sensing elementary properties, such as changes in resistance (e.g., strain gauges, potentiometers), in electric field (e.g., capacitive sensors), or in magnetic field (e.g., inductive, variable-reluctance, or eddy current sensors). Other mechanisms may be based on special materials (e.g., piezoelectric crystals) or on optical signals and imaging systems.

Force, Torque, and Pressure Measurements

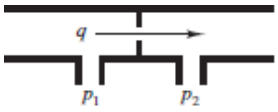
Another common class of measurements is that of pressure and force, and the related measurement of torque. Perhaps the largest family of force and pressure sensors are those based on strain gauges, such as load cells and diaphragm pressure transducers. Piezoelectric and capacitive sensors are also common.

Flow Measurements

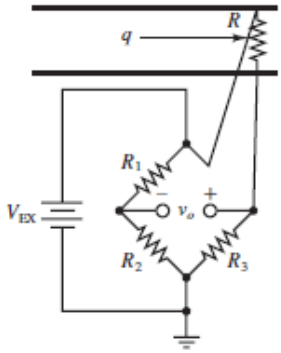
In many engineering applications it is desirable to sense the flow rate of a fluid, whether compressible (gas) or incompressible (liquid). The measurement of fluid flow rate is a complex subject. Three common flow rate measurement systems are described in Figure 7.2. The measurement described in Figure 7.2(a) is based on

the **differential-pressure** across a **calibrated orifice**, where the relationship between the differential pressure $p_1 - p_2$ and flow rate q is given by theory and calibration constants.

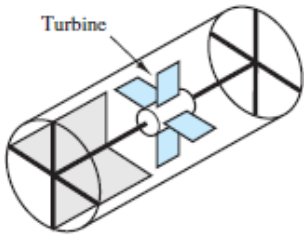
Page 477



Differential-pressure flowmeter: A calibrated orifice and a pair of pressure transducers permit the measurement of flow rate.
(a)



Hot-wire anemometer: A heated wire is cooled by the gas flow. The resistance of the wire changes with temperature.
(b)



Turbine flowmeter: Fluid flow induces rotation of the turbine; measurement of turbine velocity provides an indication of flow rate.
(c)

Figure 7.2 Devices for the measurement of flow

The system shown in Figure 7.2(b) is a **hot-wire anemometer**, which relies on a heated wire being cooled by a flowing gas. Since the resistance R of the wire changes with temperature, a Wheatstone bridge circuit can be used to convert the change in resistance to a change in voltage. **Hot-film anemometers** employ the same approach to sense the air mass-flow rate into an automotive engine and determine its air-to-fuel ratio.

Figure 7.2(c) depicts a **turbine flowmeter** in which the fluid flow causes a turbine to rotate. The angular velocity of the turbine, which is related to the fluid flow rate, can be measured by a noncontact sensor, such as a magnetic pickup.[2]

Many other techniques exist for measuring fluid flow.

Temperature Measurements

One of the most frequently measured physical quantities is temperature. The need to measure temperature arises in just about every field of engineering. This subsection is devoted to summarizing two common temperature sensors—the **thermocouple** and the **resistance temperature detector (RTD)**—and their related signal conditioning needs.

Thermocouples

A thermocouple is formed by the junction of two dissimilar metals. This junction results in an open-circuit **thermoelectric voltage** due to the **Seebeck effect**, named after Thomas Seebeck, who discovered the phenomenon in 1821. Various types of thermocouples exist; they are usually classified according to the data of Table 7.2. The Seebeck coefficient is specified at a given temperature because the output voltage v of a thermocouple has a nonlinear relationship to temperature T, which is typically expressed as a polynomial of the following form:

$$T = a_0 + a_1 v + a_2 v^2 + a_3 v^3 + \cdots + a_n v^n \tag{7.1}$$

Table 7.2 Thermocouple data

Type	Elements +/−	Seebeck coefficient (μV/°C)	Range (°C)	Range (mV)
E	Chromel/constantan	58.70 at 0°C	−270 to 1,000	−9.835 to 76.358
J	Iron/constantan	50.37 at 0°C	−210 to 1,200	−8.096 to 69.536
K	Chromel/alumel	39.48 at 0°C	−270 to 1,372	−6.548 to 54.874
R	Pt(10%)—Rh/Pt	10.19 at 600°C	−50 to 1,768	−0.236 to 18.698
T	Copper/constantan	38.74 at 0°C	−270 to 400	−6.258 to 20.869
S	Pt(13%)—Rh/Pt	11.35 at 600°C	−50 to 1,768	−0.226 to 21.108

For example, the coefficients of the J thermocouple in the range of −100 to +1,000°C are as follows:

$$a_0 = -0.048868252 \quad a_1 = 19{,}873.14503 \quad a_2 = -128{,}614.5353$$
$$a_3 = 11{,}569{,}199.78 \quad a_4 = -264{,}917{,}531.4 \quad a_5 = 2{,}018{,}441{,}314$$

Page 478

The electrical leads connected to a thermocouple must be chosen carefully because they may produce additional thermoelectric junctions that will, in effect, act as additional thermocouples. For example, each junction shown in Figure 7.3 will generate a thermoelectric voltage, which is dependent on the temperature at these junctions. However, only junction J_1 is intended to respond to temperature. To accommodate the effects of junctions J_2 and J_3 it would be necessary to know the voltages across these junctions. To address this problem the voltmeter connections are made within an isothermal block, as shown in Figure 7.4, so that their contributions to the measured voltage cancel. In addition, a reference junction at known temperature, such as the ice water bath **cold junction** shown in Figure 7.4, can be employed to reference the temperature of junction J_1 to a known temperature. In this way, when junction J_1 is held at 0°C (32°F), the measured voltage will be zero. Other J_1 temperatures will result in voltmeter readings related to $T_1 - T_{ref}$.

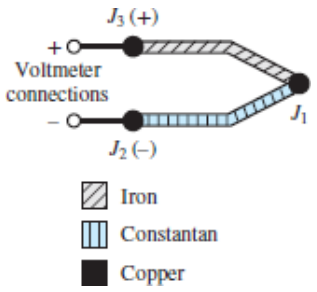


Figure 7.3 *J* thermocouple circuit

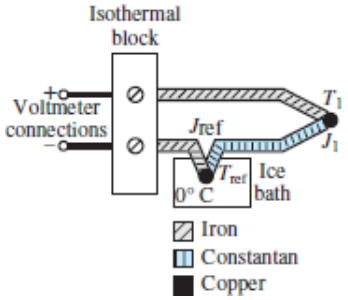


Figure 7.4 Cold junction−compensated thermocouple circuit

Resistance Temperature Detectors

A resistance temperature detector (RTD) is a device whose resistance is a function of temperature. RTDs offer greater accuracy and stability than thermocouples. **Thermistors** are part of the RTD family. All RTDs are *passive* devices. The change in resistance in an RTD is usually converted to a change in voltage by forcing a current through the device. RTDs are susceptible to a **self-heating error** caused by i^2R resistive heating. The sensitivity of an RTD to this error is usually denoted by the power required to raise its temperature by 1°C. A smaller current will reduce self-heating, but it will also reduce the output voltage.

RTDs have fairly linear dependence on temperature. The RTD **temperature coefficient** α, which can be positive or negative, is commonly defined in terms of the change in resistance $R_{100} - R_0$ from 0 to 100°C.

$$\alpha \equiv \frac{R_{100} - R_0}{100 - 0} \frac{\Omega}{^\circ\text{C}} \tag{7.2}$$

A more accurate cubic equation relating the RTD resistance to its temperature depends upon published tables of coefficients. As an example, a platinum RTD Page 479could be described either by the temperature coefficient α = 0.003911 or by the equation

$$\begin{aligned} R_T &= R_0(1 + AT - BT^2 - CT^3) \\ &= R_0(1 + 3.6962 \times 10^{-3}T - 5.8495 \times 10^{-7}T^2 \\ &\quad - 4.2325 \times 10^{-12}T^3) \end{aligned} \tag{7.3}$$

where the coefficient C is equal to zero for temperatures above 0°C.

Because RTDs have fairly low resistance, they are sensitive to error introduced by the added resistance of lead wires. Figure 7.5 depicts the effect of the lead resistances r_L on the RTD measurement. Note that the measured voltage includes the resistance of the RTD and the leads. Thus, if the lead resistance is significant, its impact on the measurement will also be significant. This impact can be mitigated by the *four-wire* RTD circuit and the *three-wire* Wheatstone bridge circuit, shown in Figure 7.6(a) and (b), respectively. In the circuit of Figure 7.6(a), the resistances of the lead wires from the excitation r_{L1} and r_{L4} may be arbitrarily large since the measurement is affected by the resistances of only the output lead wires r_{L2} and r_{L3}. The circuit of Figure 7.6(b) takes advantage of

the properties of the Wheatstone bridge to cancel out the unwanted effect of the lead wires.

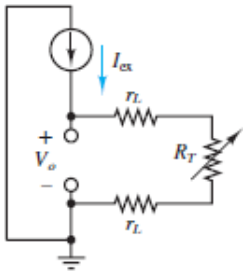
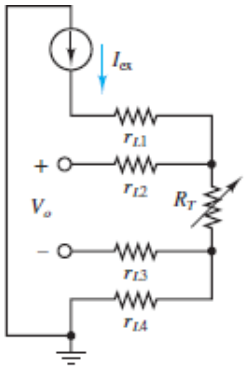
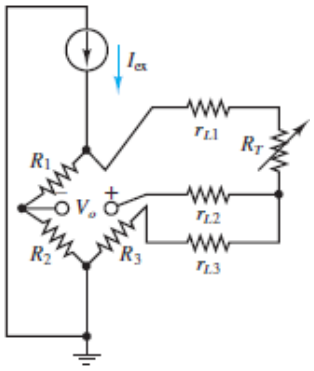


Figure 7.5 Effect of connection leads on RTD measurement



(a)



(b)

Figure 7.6 (a) Four-wire RTD circuit and (b) three-wire Wheatstone bridge RTD circuit

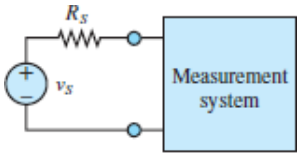
7.2 WIRING, GROUNDING AND NOISE

It is difficult to overstate the importance of proper circuit connections. This section summarizes some important considerations regarding signal source

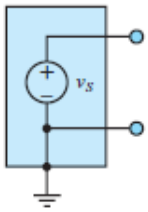
connections, various types of input configurations, noise sources and coupling mechanisms, and means of minimizing the influence of noise on a measurement.

Signal Sources and Measurement System Configurations

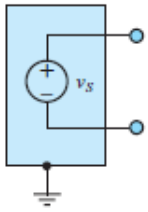
Every sensor can be thought of as a signal source. A general representation of a sensor connected to a measurement system is shown in Figure 7.7(a). The sensor is modeled as an ideal voltage source in series with a source resistance. Although this representation is not appropriate for all sensors, it permits the discussion of an important wiring issue. Figure 7.7(b) and (c) show two types of sources: **grounded** and **floating**. One terminal of a grounded source is tied to a reference ground, such as the case or housing of the source. (The case or housing of an electric device is commonly tied to earth ground through the thick, round prong of a typical three-prong AC plug.) Neither terminal of a floating source is tied to a reference ground, thus, the voltage across its terminals is unrelated to the reference ground. A thermocouple acts *intrinsically* as a floating source because its output is the difference of two voltages. A thermocouple *could* be used as a grounded source, but this is usually not a desirable arrangement for this particular sensor.



(a) Ideal signal source connected to measurement system



(b) Grounded signal source



(c) Floating signal source

Figure 7.7 Measurement system and types of signal sources

A measurement system can also be either **ground-referenced** or **differential**. In a ground-referenced system, the signal low connection is tied to the instrument case ground; in a differential system, neither of the two signal connections is tied to ground. Thus, a differential measurement system is well suited to measuring the difference between two signal levels (such as the output of an ungrounded thermocouple). Multimeters and oscilloscopes are examples of differential and ground-referenced measurement systems, respectively.

Page 480

One of the potential dangers in dealing with grounded sources is the introduction of **ground loops**. A ground loop is an undesired current path caused by the connection of two reference voltages to each other, as shown in Figure 7.8, where a grounded source is connected to a ground-referenced measurement system. The source ground and the measurement system ground are denoted by two different symbols because there may exist a voltage difference Δv between them. This voltage difference, if it exists, is the result of a current through the nonzero resistance of the nonideal wire connecting the two grounds. The net effect of the ground loop is that the measured voltage v_m would include the unknown voltage difference Δv, as shown in Figure 7.8. Ground loops can cause substantial errors in measurement systems. In addition, ground loops are the primary cause of unwanted noise.

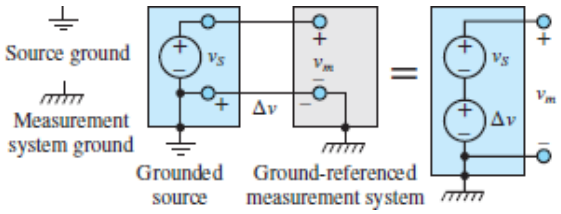


Figure 7.8 Ground loop in ground-referenced measurement system

A differential measurement system, such as that shown in Figure 7.9, can be employed to eliminate the impact of a ground loop Δv. Notice in the figure that the source and measurement system grounds are not connected to each other through their cases or otherwise.

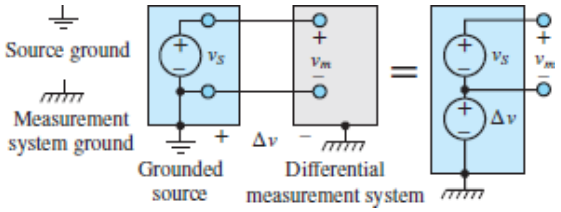


Figure 7.9 Differential (nonreferenced) measurement system

If the source connected to the differential measurement system is floating, as shown in Figure 7.10, it is often recommended to reference the source to the instrument ground by means of two identical resistors that provide a return path to ground for any currents present at the instrument. An example of such input currents would be the input bias currents always present at the input of an operational or instrumentation amplifier.

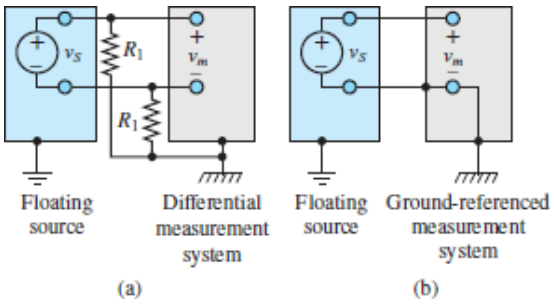


Figure 7.10 Measuring signals from a floating source: (a) differential input; (b) single-ended input

Noise Sources and Coupling Mechanisms

Noise—meaning any undesired signal in a measurement—is unavoidable in measurements. The block diagram shown in Figure 7.11 depicts the two essential requirements for a noisy measurement: a **noise source** and a **noise coupling mechanism**. Page 481Noise sources are always present and are often impossible to mitigate completely; typical sources of noise in practical measurements are the electromagnetic fields caused by fluorescent light fixtures, video monitors, power supplies, switching circuits, and high-voltage (or current) circuits. Many other sources exist, of course, but often the simple sources in our everyday environment are the most difficult to defeat.

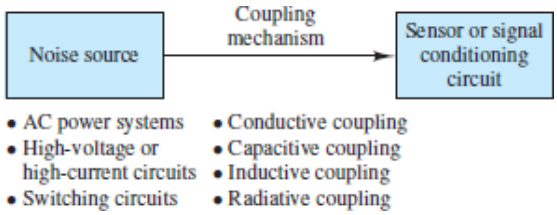


Figure 7.11 Noise sources and coupling mechanisms

Various coupling mechanisms can exist between a noise source and an instrument. Noise coupling can be conductive; that is, noisy currents may be conducted directly from the source to the instrument. Noise can also be coupled capacitively, inductively, and radiatively.

Figure 7.12(a) illustrates how interference can be **conductively coupled** by way of a ground loop. Notice that a power supply is connected to both a load and a sensor. KCL requires that the current i is the sum of the load and sensor currents. If the sensor and load currents share a substantial portion of the ground return path and the load current is substantial compared to the sensor current, the voltage at junction a can be substantially higher than ground because of the nonzero resistance of the nonideal wire. Thus, the sensor voltage will include not only v_{ba} but also v_a itself, which may be large. In other words, the measured sensor output will no longer be v_o, but $v_o + v_{ba} + v_a$. If the load is switched on and off, its current changes abruptly and these changes will be manifested in the sensor output voltage as noise.

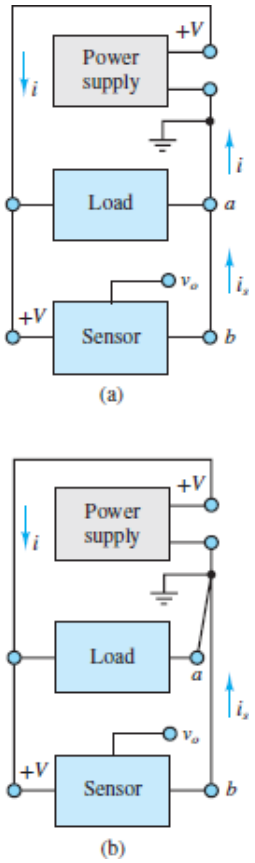


Figure 7.12 Conductive coupling: ground loop and separate ground returns

This problem can be effectively mitigated by providing separate *ground returns* for the load and sensor to eliminate the shared ground loop. Figure 7.12(b) depicts this simple modification. The sensor output voltage is now $v_o + v_{ba}$, which is unaffected by the load current.

The mechanism of **capacitive coupling** noise is rooted in electric fields due to external sources. The electromagnetic principle is depicted in Figure 7.13(a), where a noise source is shown to generate an electric field. If a conductor in Page 482the noise source is sufficiently close to a conductor in the measurement system, the electric field separating the two conductors will be impacted by changes in the distance between the two conductors, which effectively form a capacitor. Figure 7.13(b) depicts an equivalent circuit in which the noise voltage v_N couples to the measurement circuit through a capacitor that represents the capacitance of the noise path. Most notebook computer touchpads rely on capacitive coupling between the pad itself and a conductive human finger to

operate. If this coupling is not well designed, it is common to find the cursor jumping all over the screen due to the motion of a human hand nearby.

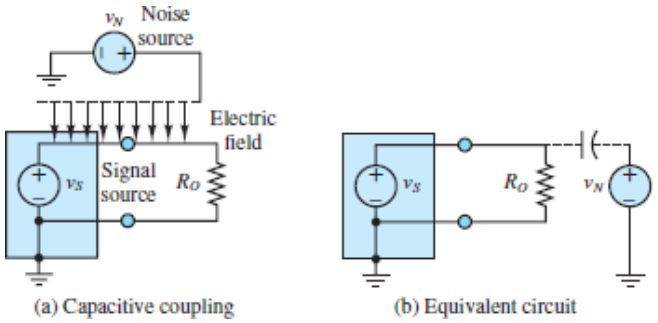


Figure 7.13 Capacitive coupling and equivalent-circuit representation

Inductive coupling noise is due to the presence of conductive loops in a measurement system interacting with spurious magnetic fields. The **mutual inductance** between the noise source and the measurement system enables the coupling, as depicted in Figure 7.14.

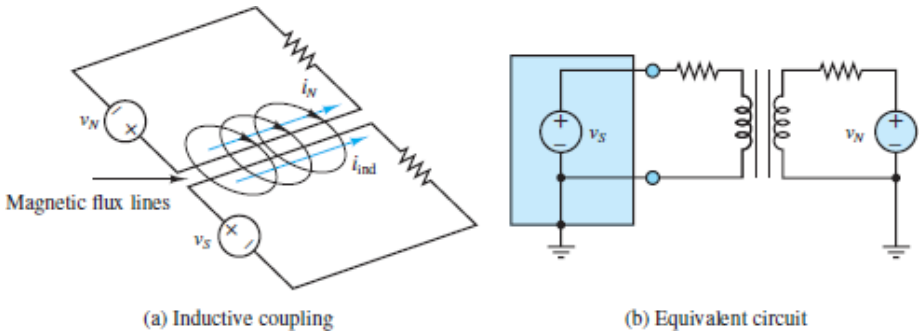


Figure 7.14 Inductive coupling and equivalent-circuit representation

Noise Reduction

Various techniques exist for minimizing the effect of undesired interference, in addition to proper wiring and grounding procedures. The two most common methods are **shielding** and the use of **twisted-pair wire**. A shielded cable is

shown in Figure 7.15. The shield is made of a copper braid or of foil and is usually grounded at the source end *but not at the instrument end* because this would result in a ground loop. The shield can protect the signal from a significant amount of electromagnetic interference, especially at lower frequencies. Shielded cables with various numbers of conductors are available commercially. However, shielding cannot prevent inductive coupling. The simplest method for minimizing inductive coupling is the use of twisted-pair wire; the reason for using twisted-pair wire is that Page 483untwisted wire can offer large loops that can couple a substantial amount of electromagnetic radiation. Twisting drastically reduces the enclosed loop area and with it the interference. Twisted pair is available commercially.

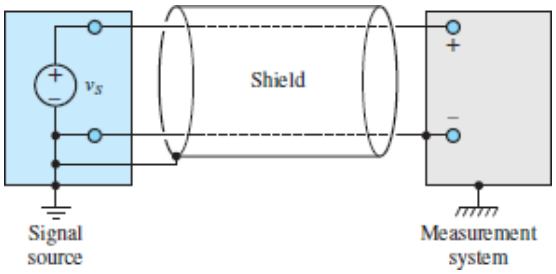


Figure 7.15 Shielding

Four important rules of thumb to reduce measurement noise are

1. Use large conducting ground planes to reduce the resistance from one point to another along the common reference ground.
2. Connect loads with potentially large currents close to the ultimate return point of a ground plane so as to minimize the current through most of the ground plane.
3. Use grounded shielding, such as provided in coax cabling, between a source and the measurement system.
4. Use twisted-pair wiring or other means to keep the enclosed area of all wire loops as small as possible.

7.3 SIGNAL CONDITIONING

A properly wired, grounded, and shielded sensor connection is a necessary first stage of any well-designed measurement system. The next stage consists of any **signal conditioning** that may be required to manipulate the sensor output into a form appropriate for the intended use. Very often, the sensor output is meant to be fed into a digital computer, as illustrated in Figure 7.1. In this case, it is important

to condition the signal so that it is compatible with the process of data acquisition. Two of the most important signal conditioning functions are *amplification* and *filtering*. Both are discussed in this section.

Instrumentation Amplifiers

An **instrumentation amplifier (IA)** is a differential amplifier with very high input impedance, low bias current, and programmable gain that finds widespread application when low-level signals with large common-mode components are to be amplified in noisy environments. This situation occurs frequently when a low-level transducer signal needs to be preamplified, prior to further signal conditioning (e.g., filtering). Instrumentation amplifiers were briefly introduced in Chapter 6 as an extension of the differential amplifier. The IA introduced there consisted of two stages, the first composed of two noninverting amplifiers, the second of a differential amplifier. Although this design is useful and is sometimes employed in practice, it suffers from a few drawbacks, most notably the requirement for precisely matched resistors and source impedances to obtain the maximum possible cancellation of the common-mode signal. If the resistors are not matched exactly, the common-mode rejection ratio of the amplifier is significantly reduced.

Assume the amplifier of Figure 7.16 has the following characteristics:

$$R_2 = R_2' \qquad R_F = R_F' \qquad R' = R + \Delta R$$

where ΔR is the mismatch between R and R'. The closed-loop gain of the input-stage noninverting amplifiers (see Example 6.2) is

$$G = \frac{v_b'}{v_b} = \frac{v_a'}{v_a} = 1 + \frac{2R_2}{R_1} \tag{7.4}$$

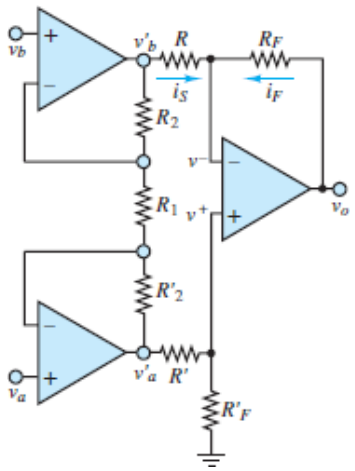


Figure 7.16 Discrete op-amp instrumentation amplifier

Page 484

The voltage at the noninverting terminal is

$$v^+ = \frac{R_F}{R_F + R + \Delta R} v_a' \tag{7.5}$$

Since the inverting-terminal voltage is $v^- = v^+$, the feedback current is

$$i_F = \frac{v_{out} - v^-}{R_F} = \frac{v_{out} - [R_F/(R_F + R + \Delta R)]v_a'}{R_F} \tag{7.6}$$

The source current is

$$i_S = \frac{v_b' - v^-}{R} = \frac{v_b' - [R_F/(R_F + R + \Delta R)]v_a'}{R} \tag{7.7}$$

Assume that input currents of each op-amp are negligible and apply KCL at the inverting node to find $i_F = -i_S$ and obtain the following expression.

$$\frac{v_o}{R_F} = \frac{v_a'}{R_F + R + \Delta R} - \frac{v_b'}{R} + \frac{R_F}{R}\frac{v_a'}{R_F + R + \Delta R}$$

$$= \left(1 + \frac{R_F}{R}\right)\frac{v_a'}{R_F + R + \Delta R} - \frac{v_b'}{R}$$

The output voltage may be computed as:

$$v_o = \frac{R_F}{R}\left[\frac{R + R_F}{R + R_F + \Delta R} v_a' - v_b'\right] \tag{7.8}$$

Note that if $\Delta R = 0$, then $v_o = (R_F/R)(v_a' - v_b')$. However, because of the resistor mismatch, there is a corresponding mismatch between the gains for the two differential signal components. Further—and more important—if the original signals v_a and v_b contained both differential-mode and common-mode components such that

$$v_a = v_{CM} + \frac{v_{DM}}{2} \qquad v_b = v_{CM} - \frac{v_{DM}}{2} \tag{7.9}$$

then:

$$v_a' = G(2v_{CM} + v_{DM})/2 \qquad v_b' = G(2v_{CM} - v_{DM})/2 \tag{7.10}$$

The common-mode components do not cancel in the amplifier output because of the gain mismatch, and the output of the amplifier is

$$v_o = R_F\left(\frac{R+R_F}{2R}\right)\left[\frac{G(2v_{CM}+v_{DM})}{R_F+R+\Delta R}\right] - \frac{R_F}{2R}G(2v_{CM}-v_{DM}) \tag{7.11}$$

resulting in the output voltage of

$$v_o = v_{o,DM} + v_{o,CM} \tag{7.12}$$

with

$$\begin{aligned} v_{o,DM} &= R_F\left(\frac{R+R_F}{2R}\right)\left(\frac{Gv_{DM}}{R_F+R+\Delta R}\right) + \frac{R_F}{2R}Gv_{DM} \\ &= \frac{R_F}{R}\,G\frac{v_{DM}}{2}\left(\frac{2R_F+2R+\Delta R}{R_F+R+\Delta R}\right) \end{aligned} \tag{7.13}$$

Page 485

and

$$\begin{aligned} v_{o,CM} &= R_F\left(\frac{R+R_F}{R}\right)\left(\frac{Gv_{CM}}{R_F+R+\Delta R}\right) - \frac{R_F}{R}Gv_{CM} \\ &= \frac{R_F}{R}Gv_{CM}\left(\frac{-\Delta R}{R_F+R+\Delta R}\right) \end{aligned} \tag{7.14}$$

Notice that as $\Delta R \to 0$, $v_{oCM} \to 0$ and $v_{oDM} \to (R_F/R)\,Gv_{DM}$. The common-mode rejection ratio (CMRR; see Chapter 6), in units of decibels, is defined by:

$$\begin{aligned} \text{CMRR}_{dB} &= 20\log\left|\frac{A_{DM}}{A_{CM}}\right| = 20\log\left|\frac{A_{DM}}{v_{o,CM}/v_{CM}}\right| \\ &= 20\log\left|\frac{A_{DM}}{\frac{R_F}{R}G\left(\frac{-\Delta R}{R_F+R+\Delta R}\right)}\right| \end{aligned} \tag{7.15}$$

where A_{DM} is the *differential gain* (which is usually assumed equal to the nominal design value). Since the common-mode gain v_{oCM}/v_{CM} should ideally be zero, the theoretical CMRR for the instrumentation amplifier with perfectly matched resistors is infinite. In fact, even a small mismatch in the resistors would

dramatically reduce the CMRR, as two of the Check Your Understanding exercises at the end of this section illustrate. Even with resistors having 1 percent tolerance, the maximum CMRR that could be attained for typical values of resistors and an overall gain of 1,000 would be only 60 dB. In many practical applications, a requirement for a CMRR of 100 or 120 dB is not uncommon, and these would demand resistors of 0.01 percent tolerance. It should be evident, then, that the "discrete" design of the IA, employing three op-amps and discrete resistors, will not be adequate for the more demanding instrumentation applications.

The general expression for the CMRR of the instrumentation amplifier of Figure 7.16, without assuming any of the resistors are matched, except for R_2 and R'_2, is

$$\mathrm{CMRR_{dB}} = 20\log\left|\frac{A_{DM}}{A_{CM}}\right| = 20\log\left|\frac{(R_F/R)(1+2R_2/R_1)}{\frac{R_F}{R}\left[\frac{R'_F}{R_F}\left(\frac{R_F+R}{R'_F+R'}\right)-1\right]}\right| \tag{7.16}$$

and it can easily be shown that the CMRR is infinite if the resistors are perfectly matched.

Many of the problems encountered in the design of instrumentation amplifiers using discrete components can be dealt with effectively through a single *monolithic integrated circuit,* where the resistors can be carefully matched by appropriate fabrication techniques. The functional structure of an IC instrumentation amplifier is depicted in Figure 7.17. Specifications for a common IC instrumentation amplifier (and a more accurate circuit description) are shown in Figure 7.18. Among the features worth mentioning here are the programmable gains, which the user can set by suitably connecting one or more of the resistors labeled R_1 to the appropriate connection. Note that the user may also choose to Page 486connect additional resistors to control the amplifier gain, without adversely affecting the amplifier's performance, since R_1 requires no matching. In addition to the pin connection that permits programmable gains, two additional pins, **sense** and **reference**, are provided to the user for the purpose of referencing the output voltage to a signal other than ground, by means of the reference terminal, or for further amplifying the output current (e.g., with a transistor stage), by connecting the sense terminal to the output of the current amplifier.

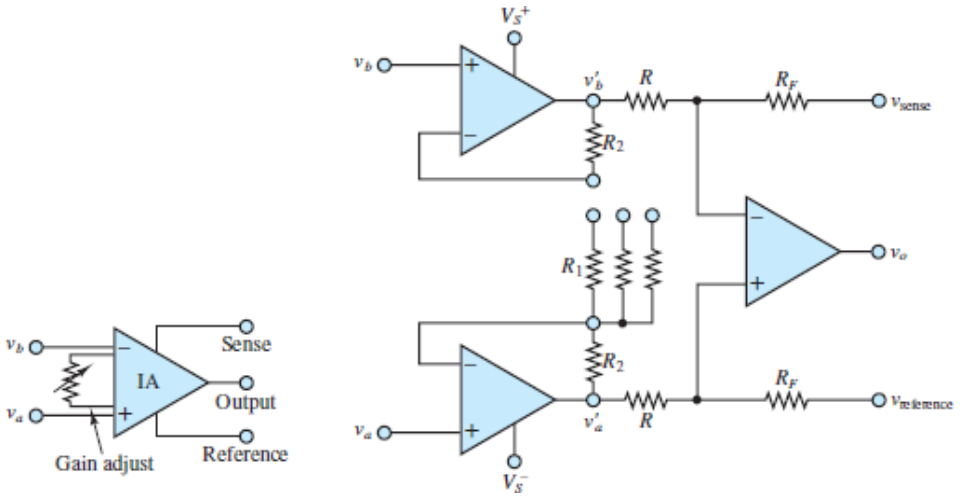


Figure 7.17 IC instrumentation amplifier

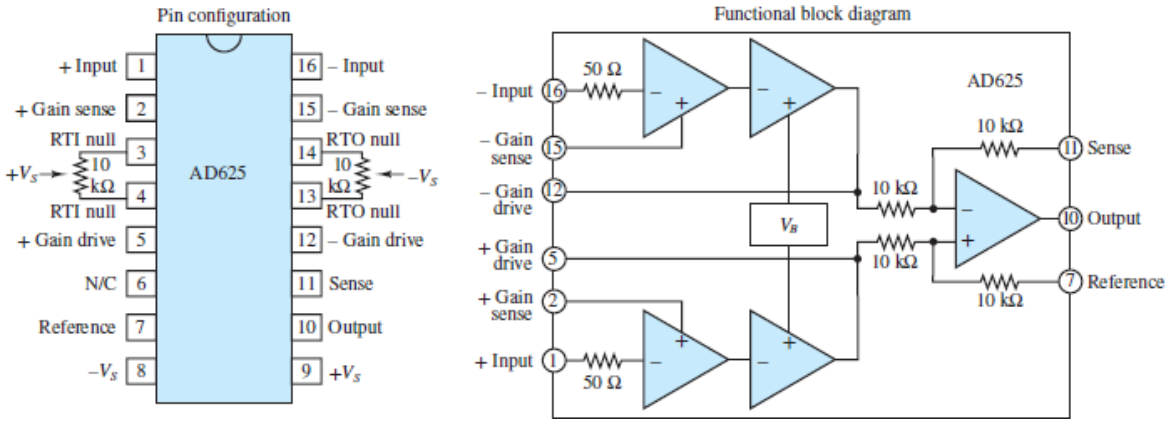


Figure 7.18 AD625 instrumentation amplifier

EXAMPLE 7.1 Common-Mode Gain and Rejection Ratio

Problem

Compute the common-mode gain and common-mode rejection ratio for the instrumentation amplifier of Figure 7.16.

Solution

Known Quantities: Amplifier nominal closed-loop gain; resistance values; resistor tolerance.

Schematics, Diagrams, Circuits, and Given Data: $G = 10$; $R_F = 10$ kΩ; $R = 1$ kΩ; $\Delta R = 20$ Ω.

Find: v_{ocom}/v_{CM}, $CMRR_{dB}$

Analysis: The common-mode gain is equal to the ratio of the common-mode output signal to the common-mode input; from equation 7.14:

$$\frac{v_{o,\,com}}{v_{com}} = \frac{R_F}{R}G\left(\frac{-\Delta R}{R + R_F + \Delta R}\right) = 10(10)\left(\frac{-0.02}{11.02}\right) = -0.1815$$

The CMRR (in units of decibels) can be computed from equation 7.15, where

$$A_{DM} = G \times \frac{R_F}{R} = 100$$

and therefore,

$$CMRR_{dB} = 20\log\left|\frac{A_{DM}}{A_{CM}}\right|_{dB} = 20\log\left|\frac{A_{DM}}{v_{o,\,CM}/v_{CM}}\right| = 20\log\left|\frac{A_{DM}}{\frac{R_F}{R}G\left(\frac{-\Delta R}{R + R_F + \Delta R}\right)}\right|$$

$$= 20\log\left|\frac{100}{\frac{10}{1}(10)\left(\frac{-0.02}{11.02}\right)}\right| = 54.8\ \text{dB}$$

Comments: Note that, in general, it is difficult to determine exactly the level of resistor mismatch ΔR in an instrumentation amplifier.

LO

EXAMPLE 7.2 Instrumentation Amplifier Gain Configuration Using Internal Resistors

Problem

Determine the possible input-stage gains that can be configured using the choice of resistor values given for the instrumentation amplifier (IA) of Figure 7.17.

Solution

Known Quantities: IA resistor values.

Find: G, for different resistor combinations.

Schematics, Diagrams, Circuits, and Given Data: $R_F = R = 10$ kΩ; $R_2 = 20$ kΩ; $R_1 = 80.2, 201, 404$ Ω.

Analysis: Recall that the gain of the input stage (for each of the differential inputs) can be calculated according to equation 7.4:

$$G = 1 + \frac{2R_2}{R_1}$$

Page 488

Thus, by connecting each of the three resistors, we can obtain gains

$$G_1 = 1 + \frac{40{,}000}{80.2} = 500 \qquad G_2 = 1 + \frac{40{,}000}{201} = 200 \qquad G_3 = 1 + \frac{40{,}000}{404} = 100$$

It is also possible to obtain additional input-stage gains by connecting resistors in parallel:

$$80.2\|201 = 57.3\,\Omega (G_4 \approx 700) \qquad 80.2\|404 = 66.9\,\Omega (G_5 \approx 600)$$
$$404\|201 = 134.2\,\Omega (G_6 \approx 300)$$

Comments: The use of resistors supplied with the IA package is designed to reduce the uncertainty introduced by the use of external resistors since the value of the internally supplied resistors can be controlled more precisely.

CHECK YOUR UNDERSTANDING

Use the definition of the common-mode rejection ratio given in equation 7.16 to compute the CMRR (in decibels) of the amplifier of Example 7.1 if $R_F/R = 100$ and $G = 10$, and if $\Delta R = 5$ percent of R. Assume $R = 1$ kΩ and $R_F = 100$ kΩ.

Repeat for a 1 percent variation in R.

Repeat for a 0.01 percent variation in R.

Calculate the mismatch in differential gains for the 5 percent resistance mismatch.

Answer: 66 dB; 80 dB; 120 dB; −6.1 dB

CHECK YOUR UNDERSTANDING

Calculate the mismatch in gains for the differential components for the 1 percent resistance mismatch of the previous Check Your Understanding.

What value of resistance R_1 would permit a gain of 1,000 for the IA of Example 7.2?

Answer: −20.1 dB; 40 Ω

7.4 ANALOG-TO-DIGITAL AND DIGITAL-TO-ANALOG CONVERSION

To take advantage of the capabilities of a microprocessor, it is necessary to suitably interface signals from external devices with a microcontroller. Depending on the nature of the signal, either an analog or a digital interface circuit will be required. The advantages in memory storage, programming flexibility, and computational power afforded by microcontrollers are such that the instrumentation designer almost always chooses to convert an analog signal to an equivalent digital representation, to exploit these capabilities. In many cases, the data converted from analog to digital form remain in digital form for ease of storage or for further processing. In some instances it is necessary to convert the data back to analog form. The latter condition arises frequently in the context of control system design, Page 489where an analog measurement is converted to digital form and processed by a digital computer to generate a control action (e.g., raising or lowering the temperature of a process, or exerting a force or a torque); in such cases, the output of the digital computer is converted back to analog form so that a continuous signal becomes available to the actuators. Figure 7.19

illustrates the general appearance of a digital measuring instrument and of a digital controller acting on a plant or process.

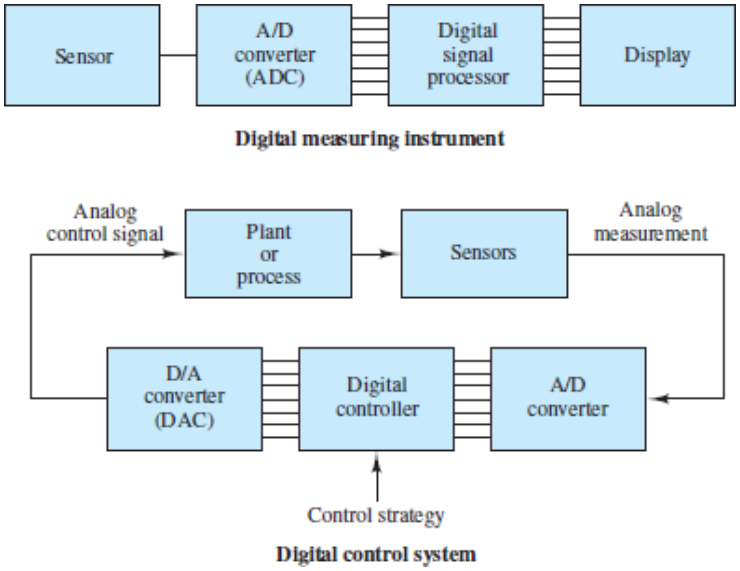


Figure 7.19 Block diagrams of a digital measuring instrument and a digital control system

It is useful and often necessary to understand how the digital-to-analog converter (DAC) and analog-to-digital converter (ADC) blocks of Figure 7.19 function. Discrete circuit examples of these blocks illustrate their fundamental aspects; however, it is uncommon (and impractical) to design such circuits using discrete components. The performance and ease of use of IC packages make them the preferred choice in virtually all applications.

Digital-to-Analog Converters

A **digital-to-analog converter (DAC)** converts a binary word to an analog output voltage (or current). The binary word is represented in terms of 1s and 0s, where typically (but not necessarily) 1s correspond to a 5-V level and 0s to a 0-V signal. As an example, consider a 4-bit binary word representing a positive (or unsigned) integer number

$$B = (b_3 b_2 b_1 b_0)_2 = (b_3 \cdot 2^3 + b_2 \cdot 2^2 + b_1 \cdot 2^1 + b_0 \cdot 2^0)_{10} \tag{7.17}$$

The analog voltage corresponding to the digital word B would be

$$v_a = (8b_3 + 4b_2 + 2b_1 + b_0)\delta v \tag{7.18}$$

where δv is the smallest *step size* by which v_a can increment. This step size will occur whenever the least significant bit (LSB) b_0 changes and is the smallest increment the digital number can make.

Page 490

The step size is determined on the basis of each given application, and it is usually determined on the basis of the number of bits in the digital word to be converted to an analog voltage. The maximum value v_a can attain is

$$\begin{aligned} v_{a\max} &= (2^{n-1} + 2^{n-2} + \cdots + 2^1 + 2^0)\delta v \\ &= (2^n - 1)\delta v \end{aligned} \tag{7.19}$$

It is relatively simple to construct a DAC using a summing amplifier. Consider the circuit shown in Figure 7.20, where each bit is represented by a switch. When the switch is closed, the bit takes a value of 1; when the switch is open, the bit has a value of 0. Thus, the output of the DAC is proportional to the word $bn_{-1}bn_{-2}\cdots b_1b_0$.

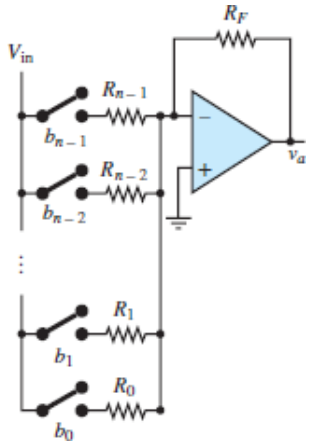


Figure 7.20 An *n*-bit digital-to-analog converter

A property of the summing amplifier is that the sum of the currents at the inverting node is zero, yielding the relationship

$$v_a = -\sum_{0}^{n-1} \frac{R_F}{R_i} b_i V_{in} \tag{7.20}$$

where R_i is the resistor associated with each bit and b_i is the decimal value of the ith bit (that is, $b_i = 2^0$, $b_1 = 2^1$, and so on). If the resistors R_i are selected such that

$$R_i = \frac{R_0}{2^i} \tag{7.21}$$

the result is weighted gains for each bit such that the output voltage is

$$v_a = -\frac{R_F}{R_0}(2^{n-1}b_{n-1} + \cdots + 2^1 b_1 + 2^0 b_0)V_{\text{in}} \tag{7.22}$$

Notice that the analog output voltage is proportional to the decimal representation of the binary word. As an illustration, consider the case of a 4-bit word; a reasonable choice for R_0 might be R_0 = 10 kΩ, yielding a resistor network consisting of 10-, 5-, 2.5-, and 1.25-kΩ resistors, as shown in Figure 7.21. The largest decimal value of a 4-bit word is $2^4 - 1 = 15$, and so it is reasonable to divide this range into steps of 1 V (that is, δv = 1 V). Thus, the full-scale value of v_a is 15 V

$$0 \le v_a \le 15 \text{ V}$$

with R_F selected according to the following expression:

$$R_F = \frac{\delta v R_0}{V_{\text{in}}} = \frac{1 \cdot 10^4}{5} = 2 \text{ k}\Omega$$

The corresponding 4-bit DAC is shown in Figure 7.21.

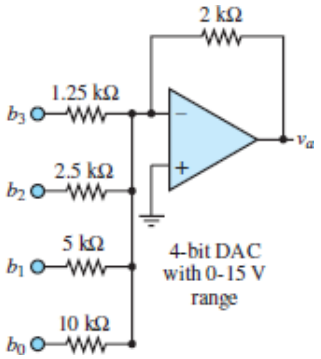


Figure 7.21 A 4-bit DAC

The DAC transfer characteristic is such that the analog output voltage v_a has a steplike appearance because of the discrete nature of the binary signal. The coarseness of the "staircase" can be adjusted by adjusting the number of bits in the binary representation.

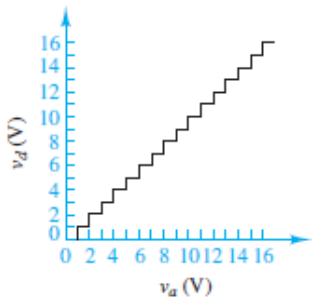
The practical design of a DAC is generally not carried out in terms of discrete components because of problems such as the accuracy required of the resistor value. Many of the problems associated with this approach can be solved by designing the complete DAC circuit as an integrated circuit. The specifications stated by the IC manufacturer include the **resolution**, that is, the minimum nonzero voltage; Page 491the **full-scale accuracy**; the **output range**; the **output settling time**; the **power supply requirements**; and the **power dissipation**.

Analog-to-Digital Converters

An **analog-to-digital converter (ADC)** converts an analog input to a binary word and is also available as a single IC. This section illustrates the essential features of four types of ADCs: the tracking ADC, which utilizes a DAC to perform the conversion; the integrating ADC; the flash ADC; and the successive-approximation ADC. The *sample-and-hold amplifier* is also introduced.

Quantization

The process of converting an analog voltage (or current) to digital form requires that the analog signal be quantized and encoded in binary form. The process of **quantization** consists of subdividing the range of the signal into a finite number of intervals; usually, one employs $2^n - 1$ intervals, where n is the number of bits available for the corresponding binary word. Following this quantization, a binary word is assigned to each interval (i.e., to each range of voltages or currents); the binary word is then the digital representation of any voltage (current) that falls within that interval. Note that the smaller the interval, the more accurate the digital representation is. However, some error is necessarily always present in the conversion process; this error is usually referred to as **quantization error**. Let v_a represent the analog voltage and v_d its quantized counterpart, as shown in Figure 7.22 for an analog voltage in the range of 0 to 16 V. In the figure, the analog voltage v_a takes on a value of $v_d = 0$ whenever it is in the range of 0 to 1 V; for $1 \leq v_a < 2$, the corresponding value is $v_d = 1$; for $2 \leq v_a < 3$, $v_d = 2$; and so on, until for $15 \leq v_a < 16$, we have $v_d = 15$. Note that each 1-V analog interval corresponds to a unique binary word. In this example, a 4-bit word is used to represent the analog voltage, although the representation is only accurate to 1 V. As the number of bits increases, the quantized voltage is closer and closer to the original analog signal.



Quantized voltage	Binary representation			
v_d	b_3	b_2	b_1	b_0
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
⋮		⋮		
14	1	1	1	0
15	1	1	1	1

Figure 7.22 A digital voltage representation of an analog voltage

Tracking ADC

Although not the most efficient in all applications, the **tracking ADC** is an easy starting point to illustrate the operation of an ADC, in that it is based on the DAC presented in the previous section. The tracking ADC, shown in Figure 7.23, compares the analog input signal with the output of a DAC; the comparator output determines whether the DAC output is larger or smaller than the analog input to be converted to binary form. If the DAC output is smaller, then the comparator output will cause an up-down counter to count up until it reaches a level close to the analog signal; if the DAC output is larger than the analog signal, then the counter is forced to count down. Note that the rate at which the up-down counter is incremented is determined by the external clock and that the binary counter output corresponds to the binary representation of the analog signal. A feature of the tracking ADC is that it follows ("tracks") the analog signal by changing 1 bit at a time.

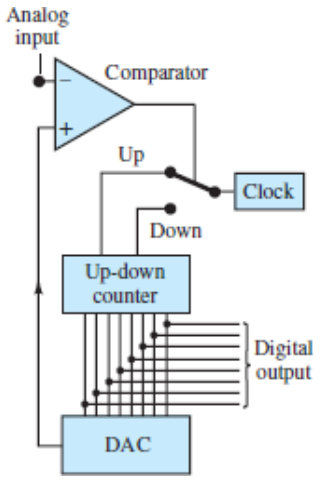


Figure 7.23 Tracking ADC

Integrating ADC

The **integrating ADC** operates by charging and discharging a capacitor, according to the following principle: If one can ensure that the capacitor charges (discharges) Page 492linearly, then the time it will take for the capacitor to discharge is linearly related to the amplitude of the voltage that has charged the capacitor. In practice, to limit the time it takes to perform a conversion, the capacitor is not required to charge fully. Rather, a clock is used to allow the input (analog) voltage to charge the capacitor for a short time, determined by a fixed number of clock pulses. Then the capacitor is allowed to discharge through a known circuit, and the corresponding clock count is incremented until the capacitor is fully discharged. The latter condition is verified by a comparator, as shown in Figure 7.24. The clock count accumulated during the discharge time is proportional to the analog voltage.

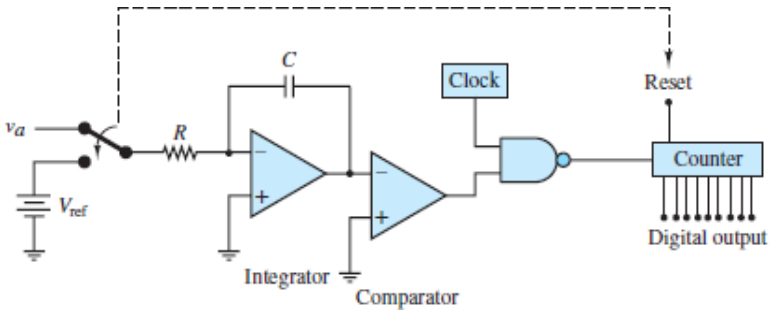


Figure 7.24 Integrating ADC

In Figure 7.24, the switch causes the counter to reset when it is connected to the reference voltage V_{ref}. The reference voltage is used to provide a known, linear discharge characteristic through the capacitor. When the comparator detects that the output of the integrator is equal to zero, it switches state and disables the NAND gate, thus stopping the count. The binary counter output is now the digital counterpart of the voltage v_a.

Successive-Approximation ADC

Successive-approximation ADCs are the most commonly used. A block diagram of the successive-approximation ADC is shown in Figure 7.25(a). This type of ADC uses a single comparator, and its performance depends strongly on the accuracy of the DAC used in the circuit. The analog output of a high-speed DAC is compared against the analog input signal. The digital result of the comparison, that is, the output of the comparator [*C* in Figure 7.25(a)] is used to control the contents of a digital buffer that both drives the DAC and provides the digital output word. The digital word corresponding to the output of the ADC is obtained by using *n* bit-by-bit comparisons, where *n* is the length of the binary word.

Flash ADC

The **flash ADC** is fully parallel and is used for high-speed conversion. A resistive divider network of 2^n resistors divides the known voltage range into that many equal increments. A network of $2^n - 1$ comparators then compares the unknown voltage with that array of test voltages. All comparators with inputs exceeding the unknown are *on*; all others are *off*. This comparator code can be converted to conventional binary by a digital priority encoder circuit. For example, assume that the 3-bit flash ADC of Figure 7.25(b) is set up with $V_{ref} = 8$ V. An input of 6.2 V Page 493is provided. Numbering the seven comparators from the top of Figure 7.25(b), the state of each is given in Table 7.3.

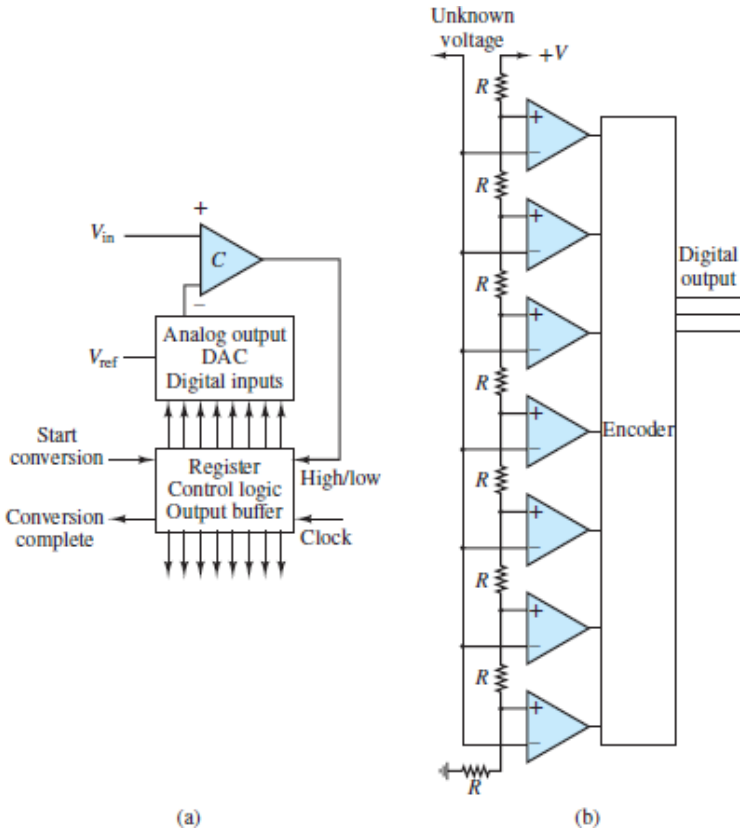


Figure 7.25 (a) Block diagram of 8-bit successive-approximation ADC; (b) a 3-bit flash ADC

Table 7.3 State of comparators in a 3-bit flash ADC

Comparator	Input on + line (V)	Input on − line (V)	Output
1	7	6.2	H
2	6	6.2	L
3	5	6.2	L
4	4	6.2	L
5	3	6.2	L
6	2	6.2	L
7	1	6.2	L

Sample-and-Hold Amplifier

These four different techniques for converting an analog voltage to its digital counterpart require a certain amount of time to perform the A/D conversion. This ADC **conversion time** is an important specification of an ADC device. A natural question at this point would be: If the analog voltage changes during the Page

494analog-to-digital conversion and the conversion process itself takes a finite time, how fast can the analog input signal change while still allowing the ADC to provide a meaningful digital representation of the analog input? To resolve the uncertainty generated by the finite ADC conversion time of any practical converter, it is necessary to use a sample-and-hold amplifier. The objective of such an amplifier is to "freeze" the value of the analog waveform for a time sufficient for the ADC to complete its task.

A typical sample-and-hold amplifier is shown in Figure 7.26. A MOSFET analog switch (see chapter 10) is used to "sample" the analog waveform. When a voltage pulse is provided to the sample input of the MOSFET switch, the MOSFET acts as a short-circuit for the duration of the sampling pulse. While the MOSFET conducts, the analog voltage v_a charges the "hold" capacitor C at a fast rate through the small "on" resistance of the MOSFET. Because the MOSFET acts as a short-circuit for the duration of the sampling pulse, the charging (RC) time constant is very small, and the capacitor charges very quickly. When the sampling pulse is over, the MOSFET analog switch returns to its nonconducting state, and the capacitor holds the sampled voltage without discharging, thanks to the extremely high input impedance of the voltage-follower (buffer) stage. Thus, v_{SH} is the sampled-and-held value of v_a at any given sampling time.

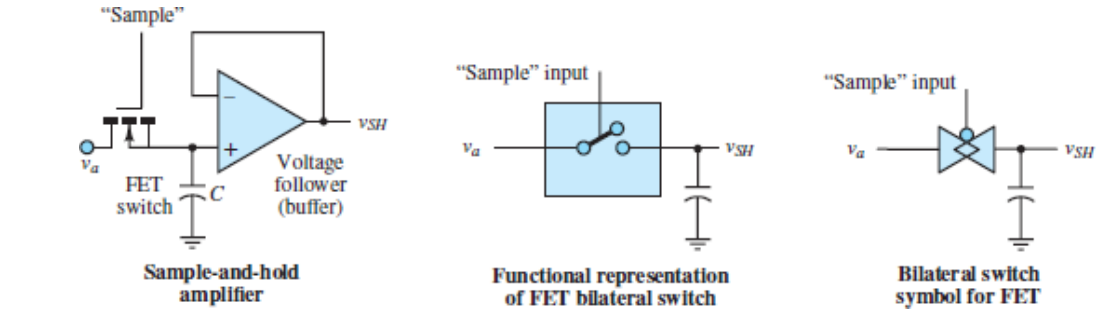


Figure 7.26 Description of the sample-and-hold process

The resolution of an ADC is a very important feature in selecting a specific ADC for a given application. Instrumentation manufacturers typically refer to **ADC resolution** as the maximum analog voltage range divided by 2^n, where n is the number of bits in the ADC. For example, an 8-bit ADC with an analog voltage range of ±15 V will have a resolution of $30/2^8 = 30/256 = 117.2$ mV. Example 7.8 illustrates this calculation in a practical application.

The appearance of the output of a typical sample-and-hold circuit is shown in Figure 7.27, together with the analog signal to be sampled. The time interval between samples, or **sampling interval**, $t_n - t_{n-1}$ allows the ADC to perform the conversion and make the digital version of the sampled signal available, say, to a microcontroller or to another data acquisition and storage system. The sampling interval needs to be at least as long as the A/D conversion time, of course; but it is reasonable to ask how frequently one needs to sample a signal to preserve its fundamental properties, that is, the basic shape Page 495of the waveform. One might instinctively be tempted to respond that it is best to sample as frequently as possible, within the limitations of the ADC, so as to capture all the features of the analog signal. In fact, this is not necessarily the best strategy. Fortunately, an entire body of knowledge exists with regard to sampling theory, which enables the practicing engineer to select the best sampling rate for any given application. The most fundamental sampling theorem is the **Nyquist sampling criterion**.

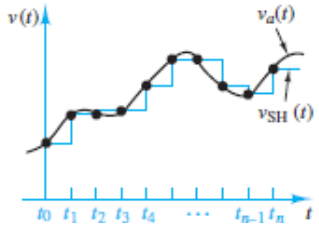


Figure 7.27 Sampled data

LO

The Nyquist criterion states that *the sample rate should be selected to be at least twice the highest-frequency component present in the signal.*

Thus, if we were sampling an audio signal (say, music), we would have to sample at a frequency of at least 40 kHz (twice the highest audible frequency, 20 kHz). In practice, it is advisable to select sampling frequencies substantially greater than the Nyquist rate; a good rule of thumb is 5 to 10 times greater. Example 7.8 illustrates how the designer might take the Nyquist criterion into account in designing a practical A/D conversion circuit.

Data Acquisition Systems

The basic building blocks of a data acquisition system are shown in Figure 7.28. A typical system employs an *analog multiplexer* to process several different input signals. A bank of bilateral analog MOSFET switches (see chapter 10), such as the one described in the sample-and-hold amplifier, enables the selection of the input signal to be sampled and converted to digital form. Control logic, employing standard gates and counters, is used to select the desired *channel* (input signal) and to trigger the sampling circuit and the ADC. When the conversion is completed, the ADC sends an appropriate *end-of-conversion* signal to the control logic, thereby enabling the next channel to be sampled.

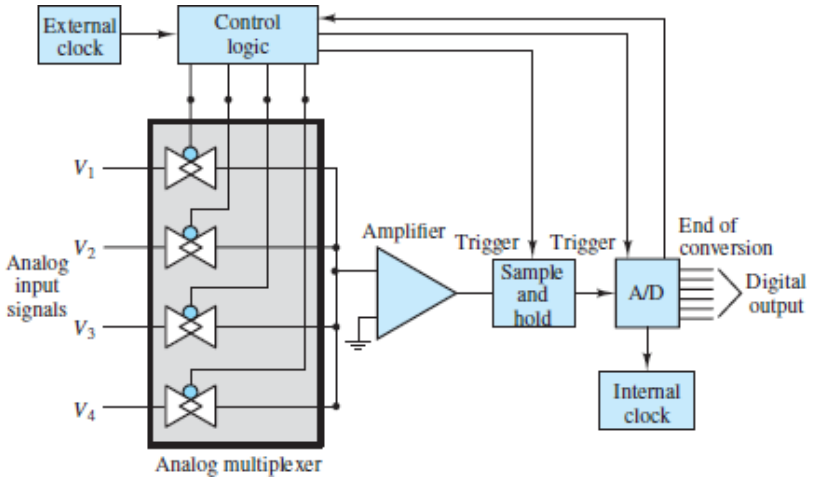


Figure 7.28 Data acquisition system

Page 496

In the block diagram of Figure 7.28, four analog inputs are shown; if these were to be sampled at regular intervals, the sequence of events would appear as depicted in Figure 7.29. Notice that the effective sampling rate for each channel is one-fourth the actual external clock rate; thus, it is important to ensure that the sampling rate for each individual channel satisfies the Nyquist criterion. Further, although each sample is held for four consecutive cycles of the external clock, the ADC can use only one cycle of the external clock to complete the conversion since its services will be required by the next channel during the next clock cycle. The internal clock that times the ADC must be sufficiently fast to allow for a complete conversion of any sample within the design range.

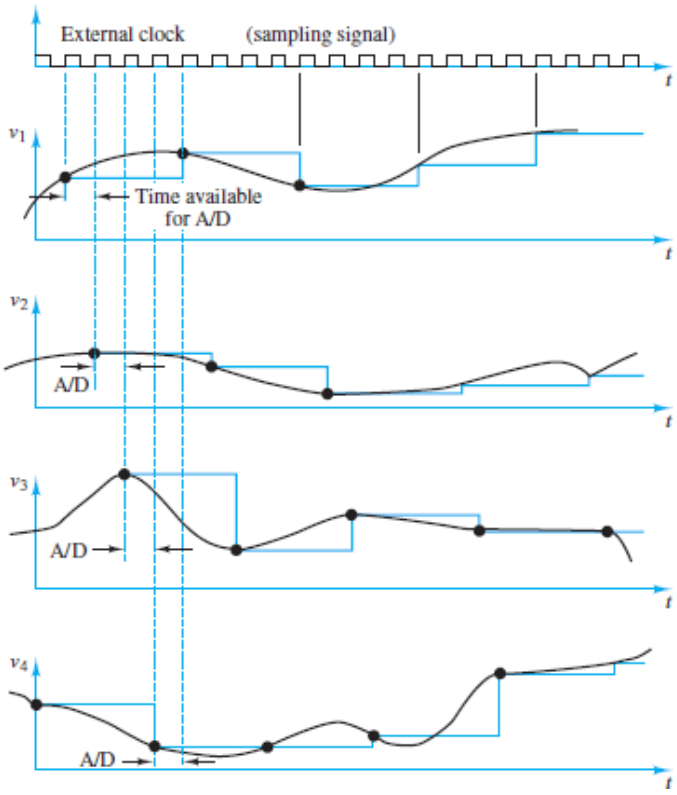


Figure 7.29 Multiplexed sampled data

EXAMPLE 7.3 DAC Resolution

Problem

Determine the *resolution* of an 8-bit DAC sampling a 12-V range.

Solution

Known Quantities: Maximum analog voltage.

Find: Resolution δv.

Page 497

Schematics, Diagrams, Circuits, and Given Data: $v_{a,\,\max} = 12$ V.

Analysis: Using equation 7.19, compute

$$\delta v = \frac{v_{a,\max} - v_{a,\min}}{2^8 - 1} = \frac{12 - 0}{2^8 - 1} = 47.1 \text{ mV}$$

Comments: Note that the resolution is dependent not only on the number of bits, but also on the analog voltage range (12 V in this case).

EXAMPLE 7.4 Determining the Required Number of Bits in a DAC

Problem

Find an expression for the required precision (number of bits) in a DAC, using the definitions of *range* and *resolution.*

Solution

Known Quantities: Range and resolution of DAC. Voltage level corresponding to logic 1.

Find: Number of DAC bits required.

Schematics, Diagrams, Circuits, and Given Data:

Range: the analog voltage range of the DAC = $v_{a\ \max} - v_{a\ \min}$

Resolution: the minimum step size δv

V_{in} = voltage level corresponding to logic 1

0 V = voltage level corresponding to logic 0

Analysis: The maximum amplitude of the DAC analog voltage output is obtained when all bits are set to 1. Using equation 7.22, determine $v_{a,\ \max}$:

$$v_{a,\max} = V_{\text{in}} \frac{R_F}{R_0} (2^n - 1)$$

The minimum analog voltage output is realized when all bits are set to logic 0. In this case, since the voltage level associated with a logic 0 is 0 V, $v_{a\ \min} = 0$. Thus, the range of this DAC is $v_{a\ \max} - v_{a\ \min} = v_{a\ \max}$.

The resolution was defined in Example 7.3 as

$$\delta v = \frac{v_{a,\max} - v_{a,\min}}{2^n - 1}$$

Knowing both range and resolution, and that $\log 2^n = n \log 2$, solve for the number of bits n.

$$n = \frac{\log[(v_{a,\max} - v_{a,\min})/\delta v + 1]}{\log 2} = \frac{\log(\text{range}/\text{resolution} + 1)}{\log 2}$$

Since n must be an integer, the result of the above expression will be rounded up to the nearest integer. For example, if we require a 10-V range DAC with a resolution of 10 mV, we can compute the required number of bits to be

$$n = \frac{\log(10/10^{-2} + 1)}{\log 2} = 9.97 \rightarrow 10 \text{ bits}$$

Page 498

EXAMPLE 7.5 Using DAC Device Data Sheets

Problem

Use a datasheet for the AD7524 8-bit D/A converter to answer the following questions:

1. What is the best (smallest) resolution attainable for a range of 10 V?
2. What is the maximum allowable conversion frequency of this DAC?

Solution

Known Quantities: Desired range of DAC.

Find: Resolution and maximum conversion frequency.

Schematics, Diagrams, Circuits, and Given Data: Range = 10 V. DAC specifications found in device data sheet.

Assumptions: The DAC is operated at full-scale range.

Analysis:

1. The AD7524 is an 8-bit converter. Thus, the best resolution that can be obtained (see equation 7.19) is

$$\delta v = \frac{v_{a,\max} - v_{a,\min}}{2^n - 1} = \frac{10}{2^8 - 1} = 39.2 \text{ mV}$$

2. The maximum frequency of the DAC depends on the *settling time*. This is defined as the time required for the output to settle to within one-half of the least significant bit of its final value. Only one conversion can be performed during the settling time. The settling time is dependent on the voltage range, and for the 10-V range indicated in this problem it is equal to $T_S = 1$ μs. The corresponding maximum *sampling frequency* is $F_S = 1/T_S = 1$ MHz.

EXAMPLE 7.6 Flash ADC

Problem

How many comparators are needed in a 4-bit flash ADC?

Solution

Known Quantities: ADC resolution.

Find: Number of comparators required.

Analysis: The number of comparators needed is $2^n - 1 = 15$.

Comments: The flash ADC has the advantage of high speed because it can simultaneously determine the value of each bit, thanks to the parallel comparators. However, because of the large number of comparators, flash ADCs tend to be expensive.

EXAMPLE 7.7 Sample-and-Hold Amplifier

Problem

Use a datasheet for the AD585 sample-and-hold amplifier to answer the following questions:

1. What is the acquisition time of the AD585?
2. How could the acquisition time be reduced?

Solution

Known Quantities: AD585 device data sheets.

Find: Acquisition time.

Schematics, Diagrams, Circuits, and Given Data: DAC specifications are found in the device data sheet. *Definition:* The *acquisition time T* is the time required for the output of the sample-and-hold amplifier to reach its final value, within a specified error bound, after the amplifier has switched from the *sample mode* to the *hold mode.* The time *T* includes the switch delay time, the slewing interval, and the amplifier settling time.

Analysis:

1. From the data sheets, the acquisition time for the AD585 is 3 μs.
2. This acquisition time could be reduced by reducing the value of the holding capacitor C_H.

EXAMPLE 7.8 Performance Analysis of an Integrated-Circuit ADC

Problem

Use a datasheet for the AD574 12-bit A/D converter to answer the following questions:

1. What is the accuracy (in volts) of the AD574?
2. What is the highest-frequency signal that can be converted by this ADC without violating the Nyquist criterion?

Solution

Known Quantities: ADC supply voltage; input voltage range.

Find: ADC accuracy; maximum signal frequency for undistorted A/D conversion.

Schematics, Diagrams, Circuits, and Given Data: $V_{CC} = 15$ V; $0 \le V_{in} \le 15$ V. ADC specifications are found in device data sheet.

Analysis:

1. The precision of the AD574 is determined by its LSB. For a range of 0 to 15 V, the resolution of the LSB is

$$\frac{V_{in,max} - V_{in,min}}{2^n - 1} = \frac{15}{2^{12} - 1} \times (\pm 1 \text{ bit}) = \pm 3.66 \text{ mV}$$

Page 500

2. The data sheet states that the maximum guaranteed conversion time of the ADC is 35 μs; therefore, the highest conversion frequency for this ADC is

$$f_{max} = \frac{1}{35 \times 10^{-6}} = 28.57 \text{ kHz}$$

   The Nyquist criterion states that the maximum signal frequency that can be sampled without distortion due to aliasing is one-half of the sampling frequency. Thus, the maximum signal frequency that can be acquired by this ADC is approximately 14 kHz.

Comments: In practice, it is a good idea to *oversample* by a certain amount. A reasonable rule of thumb is to oversample by a factor of 2 to 5. If the signal is oversampled by a factor of 2, then the maximum signal frequency would be 7 kHz, in this example.

One way to ensure that the signal being sampled is limited to a 7-kHz bandwidth is to prefilter the signal with a low-pass filter having a cutoff frequency at or below 7 kHz.

CHECK YOUR UNDERSTANDING

If the maximum analog voltage $V_{a\max}$ of a 12-bit digital-to-analog converter (DAC) is 15 V, find the smallest step size δv by which v_a can increment.

Answer: 3.66 mV

CHECK YOUR UNDERSTANDING

Find the minimum number of bits required in a DAC if the range of the DAC is from 0.5 to 15 V and the resolution of the DAC is 20 mV.

Answer: 10 bits

CHECK YOUR UNDERSTANDING

In Example 7.8, if the maximum conversion time were 50 μs, what would be the highest-frequency signal allowed by the Nyquist criterion?

Answer: $f_{\max} = 10$ kHz

7.5 COMPARATOR AND TIMING CIRCUITS

Timing and comparator circuits find frequent application in instrumentation systems. The aim of this section is to introduce the foundations of op-amp comparators and multivibrators, and of an integrated circuit timer.

The Op-Amp Comparator

The prototype of op-amp switching circuits is the op-amp comparator of Figure 7.30. This circuit *does not employ feedback.* Thus, when $R_o = 0$:

$$v_o = A(v^+ - v^-) \tag{7.23}$$

Because of the large open-loop gain of the op-amp ($A > 10^5$), any small difference ε between input voltages will cause large outputs. In particular, for ε on the order of a few tens of microvolts, the op-amp will go into saturation at either extreme, according to the voltage supply values and the polarity of the voltage difference (recall the discussion of the op-amp voltage supply limitations in Section 6.6). For example, for ε equal to 1 mV and for an open-loop gain $A = 10^5$, the ideal op-amp output would equal 100 V, and yet, in practice, the op-amp would saturate at the voltage supply limits. Clearly, almost any difference between input voltages will cause the output to saturate toward either supply voltage, depending on the polarity of ε.

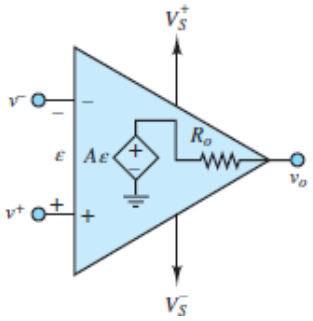


Figure 7.30 Op-amp in open-loop mode

One can take advantage of this property to generate switching waveforms. Consider, for example, the circuit of Figure 7.31 in which a sinusoidal voltage source $v_{in}(t)$ of peak amplitude V is connected to the noninverting input. In this circuit, in which the inverting terminal has been connected to ground, the differential input voltage is given by

$$\varepsilon = V \cos \omega t \tag{7.24}$$

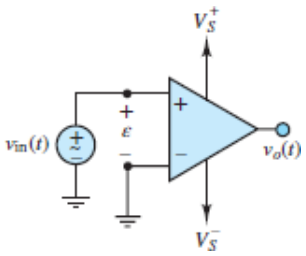


Figure 7.31 Noninverting op-amp comparator

and will be positive during the positive half-cycle of the sinusoid and negative during the negative half-cycle. Thus, the output will saturate toward V_S^+ or V_S^-, depending on the polarity of ε: the circuit is, in effect, *comparing* $v_{in}(t)$ and ground, producing a positive v_o when $v_{in}(t)$ is positive and a negative v_o when $v_{in}(t)$ is negative, independent of the amplitude of $v_{in}(t)$ (provided, of course, that the peak amplitude of the sinusoidal input is at least 1 mV or so). The circuit just described is therefore called a **comparator** and in effect it performs a binary decision, determining whether $v_{in}(t) > 0$ or $v_{in}(t) < 0$. The comparator is perhaps the simplest form of an analog-to-digital converter, that is, a circuit that converts a continuous waveform to discrete values. The comparator output consists of only two discrete levels: greater than and less than a reference voltage.

The input and output waveforms of the comparator are shown in Figure 7.32, where it is assumed that $V = 1$ V and that the saturation voltage corresponding to the ±15-V supplies is approximately ±13.5 V. This circuit is termed a **noninverting comparator** because a positive voltage differential ε gives rise to a positive output voltage. It should be evident that it is also possible to construct an inverting comparator by connecting the noninverting terminal to ground and connecting the input to the inverting terminal. Figure 7.33 depicts the waveforms for the **inverting comparator**. The analysis of any comparator circuit is summarized by the following relationship:

LO

$$\begin{array}{lll} \varepsilon > 0 & \Rightarrow & v_o = V_{sat}^+ \\ \varepsilon < 0 & \Rightarrow & v_o = V_{sat}^- \end{array} \quad \text{Operation of op-amp comparator} \tag{7.25}$$

where V_{sat} is the saturation voltage for the op-amp (somewhat lower than the supply voltage, as discussed in Chapter 6). Typical values of supply voltages for practical op-amps are ±5 to ±24 V.

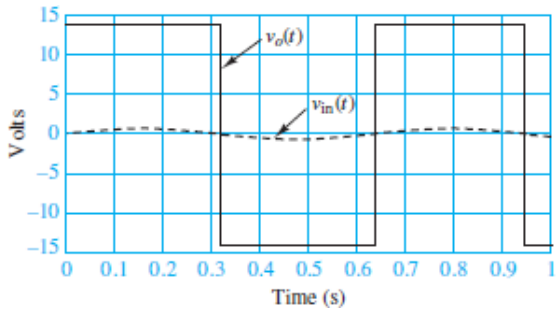


Figure 7.32 Input and output of noninverting comparator

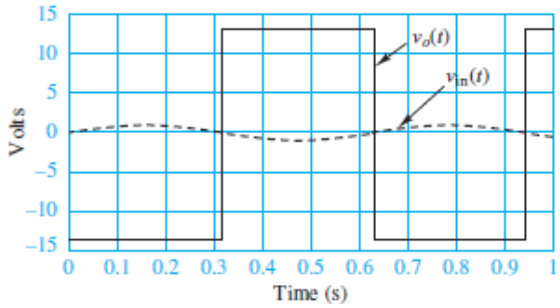


Figure 7.33 Input and output of inverting comparator

A simple modification of the comparator circuit just described consists of connecting a fixed reference voltage to one of the input terminals; the effect of the reference voltage is to raise or lower the voltage level at which the comparator will switch from one extreme to the other.

Another useful interpretation of the op-amp comparator can be obtained by considering its **input-output transfer characteristic**. Figure 7.34 displays a plot of v_o versus v_{in} for a noninverting zero-reference (no offset) comparator. This circuit is often called a **zero-crossing comparator** because the output voltage goes through a transition (V_{sat} to $-V_{sat}$, or vice versa) whenever the input voltage changes sign. Figure 7.35 displays the transfer characteristic for a comparator of the inverting type with a nonzero reference voltage.

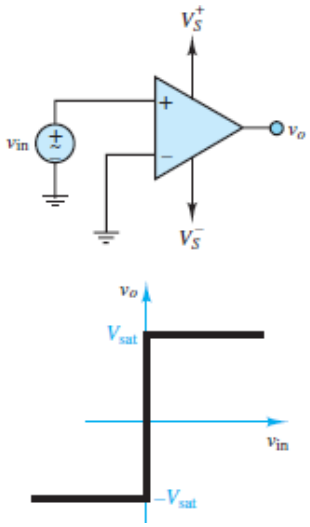


Figure 7.34 Transfer characteristic of zero-crossing comparator

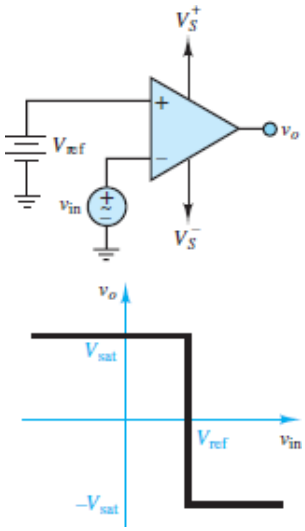


Figure 7.35 Transfer characteristic of inverting comparator with offset

When converting an analog signal to a binary representation, it is often necessary to use voltage levels other than $\pm V_{sat}$, such as 0 and 5 V. The voltage transfer characteristic can be modified by connecting a Zener diode between the output of the op-amp and the noninverting input, in the configuration sometimes Page 503called a **level** or **Zener clamp**. The circuit shown in Figure 7.36 relies on a reversed-biased Zener diode to hold a constant voltage V_Z across its terminals, from cathode to anode, as discussed in Chapter 8. When the diode is forward-biased, on the other hand, the voltage across the Zener diode terminals equals the offset voltage $V\gamma$, from anode to cathode. An additional advantage of

the level clamp is that it reduces the switching time. Input and output waveforms for this particular Zener-clamped comparator are shown in Figure 7.37, for the case of a sinusoidal $v_{in}(t)$ of peak amplitude 1 V and Zener voltage equal to 5 V. More practical variations of the Zener-clamped comparator exist, some of which employ negative feedback.

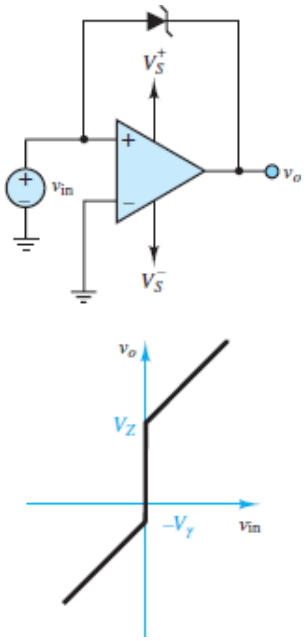


Figure 7.36 Level-clamped comparator

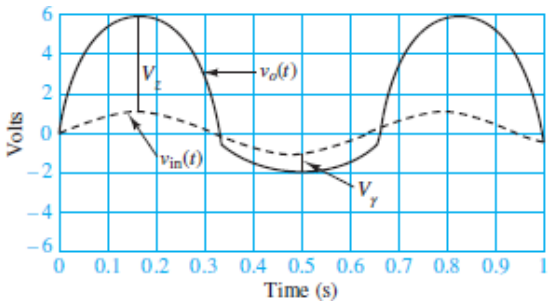


Figure 7.37 Zener-clamped comparator waveforms

Although the Zener-clamped circuit illustrates a specific issue of interest in the design of comparator circuits, namely, the need to establish desired reference output voltages other than the supply saturation voltages, this type of circuit is rarely employed in practice. Special-purpose integrated-circuit packages are available that are designed specifically to serve as comparators. These can typically accept relatively large inputs and have provision for selecting the

desired reference voltage levels (or, sometimes, are internally clamped to a specified voltage range). A representative product is the LM311, which provides an open-collector output, as shown in Figure 7.38. The open-collector output allows the user to connect the output transistor to any supply voltage of choice by means of an external pull-up resistor, thus completing the output circuit. The actual value of the resistor is not critical; values between a few hundred and a few thousand ohms are typical.

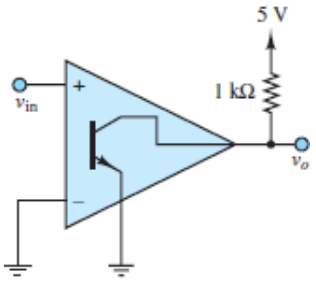


Figure 7.38 Open-collector comparator output with representative external supply connection

The Schmitt Trigger

One of the typical applications of the op-amp comparator is the detection of an input voltage exceeding a threshold level. The desired threshold is then represented by a DC reference V_{ref} connected to the noninverting input, and the input voltage source is connected to the inverting input, as in Figure 7.35. Under ideal conditions, for noise-free signals, and with an infinite slew rate for the op-amp, the operation of such a circuit would be as depicted in Figure 7.39.

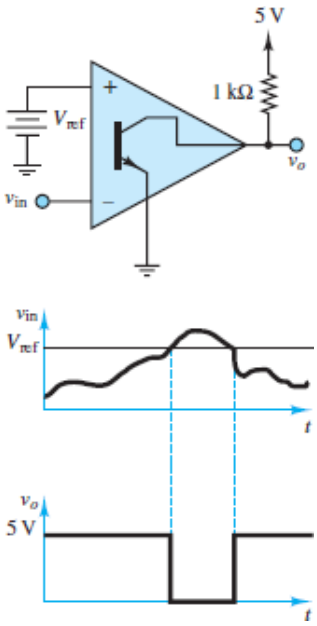


Figure 7.39 Waveforms for inverting comparator with offset

Two improvements can be made to better the switching speed of the comparator and to enable correct operation in the presence of noisy signals. If the input to the comparator is changing slowly, the comparator will not switch instantaneously since its open-loop gain is not infinite and, more important, its slew rate limits the switching speed. In fact, commercially available comparators have slew rates that are typically much lower than those of conventional op-amps. Further, Page 504in the presence of noisy inputs, a conventional comparator is inadequate because the input signal could cross the reference voltage level repeatedly and cause multiple triggering. Figure 7.40 depicts the latter occurrence.

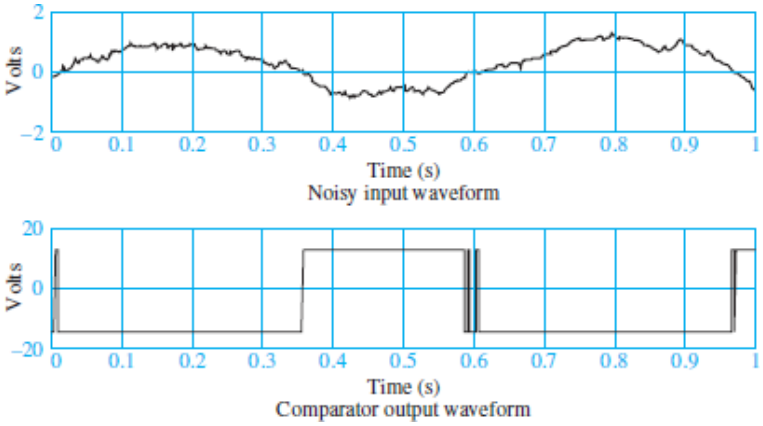


Figure 7.40 Comparator response to noisy inputs

One effective way of improving the performance of the comparator is by introducing positive feedback. Positive feedback can increase the switching speed of the comparator and provide noise immunity at the same time. Figure 7.41 depicts a frequently used comparator circuit known as a **Schmitt trigger** in which the output has been tied back to the *noninverting* input (thus the terminology *positive* feedback) by means of a resistive voltage divider. The effect of this positive-feedback connection is to provide a reference voltage at the noninverting input equal to a fraction of the comparator output voltage; since the comparator output is equal to either the positive or the negative saturation voltage $\pm V_{sat}$, the reference voltage at the noninverting input can be either positive or negative.

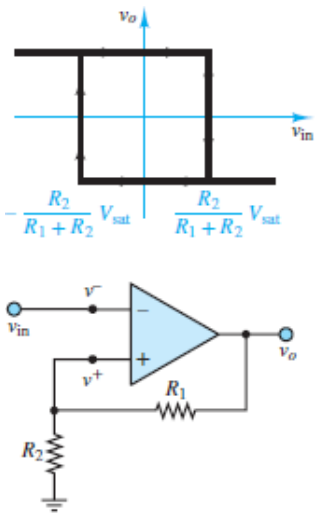


Figure 7.41 Transfer characteristic of the Schmitt trigger

Consider, first, the case when the comparator output is $v_o = +V_{sat}$. It follows that

$$v^+ = \frac{R_2}{R_2 + R_1} V_{sat} \tag{7.26}$$

and therefore the differential input voltage is

$$\varepsilon = v^+ - v^- = \frac{R_2}{R_2 + R_1} V_{sat} - v_{in} \tag{7.27}$$

The output voltage v_o will switch from the positive to the negative saturation state when the differential voltage ε becomes negative; that is, the condition for v_o to switch states is

$$v_{\text{in}} > \frac{R_2}{R_2 + R_1} V_{\text{sat}} \qquad (v_o: V_{\text{sat}} \rightarrow -V_{\text{sat}}) \tag{7.28}$$

Thus, v_o will not transition when v_{in} crosses zero; instead, v_{o} transitions at a *positive* voltage that is determined by V_{sat}, R_1 and R_2.

Page 505

Consider, now, the case when the comparator output is $v_o = -V_{\text{sat}}$. Then

$$v^+ = -\frac{R_2}{R_2 + R_1} V_{\text{sat}} \tag{7.29}$$

and therefore

$$\varepsilon = v^+ - v^- = -\frac{R_2}{R_2 + R_1} V_{\text{sat}} - v_{\text{in}} \tag{7.30}$$

The output voltage v_o will switch from the negative to the positive saturation state when the differential voltage ε becomes positive; the condition for v_o to switch states is now

$$v_{\text{in}} < -\frac{R_2}{R_2 + R_1} V_{\text{sat}} \qquad (v_o: -V_{\text{sat}} \rightarrow V_{\text{sat}}) \tag{7.31}$$

Thus, v_{o} will not transition when v_{in} crosses zero; instead, v_{o} transitions at a *negative* voltage that is determined by V_{sat}, R_1 and R_2. Figure 7.41 depicts the effect of the different thresholds on the voltage transfer characteristic, showing the switching action by means of arrows.

If it is desired to switch about voltages not centered at zero, a DC offset voltage can be superposed upon the noninverting terminal, as shown in Figure 7.42. Now the noninverting terminal voltage is

$$v^+ = \frac{R_2}{R_2 + R_1} v_o + V_{\text{ref}} \frac{R_1}{R_2 + R_1} \tag{7.32}$$

and the switching levels for the Schmitt trigger are

$$v_{\text{in}} > \frac{R_2}{R_2 + R_1} V_{\text{sat}} + V_{\text{ref}} \frac{R_1}{R_2 + R_1} \qquad (v_o: V_{\text{sat}} \rightarrow -V_{\text{sat}}) \tag{7.33}$$

for the negative-going transition and

$$v_{in} < -\frac{R_2}{R_2 + R_1} V_{sat} + V_{ref} \frac{R_1}{R_2 + R_1} \qquad (v_o: -V_{sat} \rightarrow V_{sat}) \tag{7.34}$$

for the positive-going transition. In effect, the Schmitt trigger provides a noise rejection range equal to $\pm[R_2/(R_2 + R_1)]V_{sat}$ within which v_o will not switch. Thus, if the noise amplitude is contained within this range, the Schmitt trigger will prevent multiple triggering. Figure 7.43 depicts the response of a Schmitt trigger with appropriate switching thresholds to a noisy waveform.

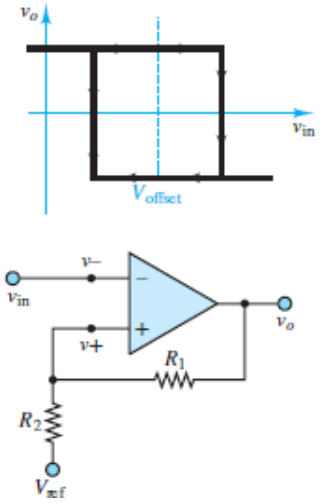


Figure 7.42 Schmitt trigger with offset voltage $V_{offset} = V_{ref} R_1/(R_1 + R_2)$

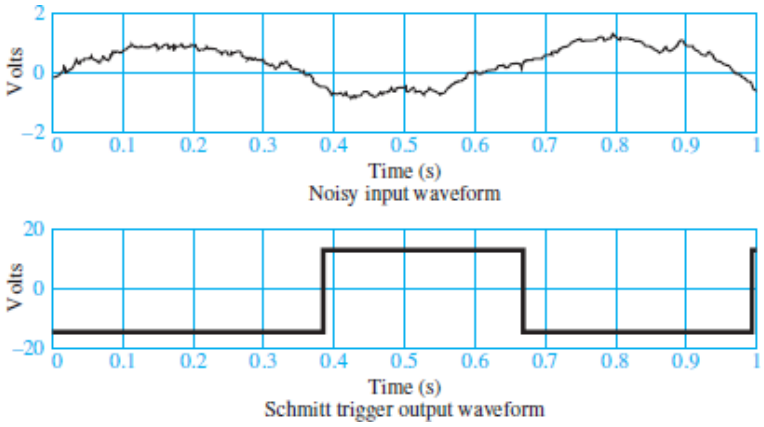


Figure 7.43 Schmitt trigger response to noisy waveforms

Multivibrators

Timing circuits

Numerous applications require timing functions, such as the generation of a fixed frequency clock waveform, which is simply a pulse train with a known period. Another is the one-shot in which a pulse of known duration and amplitude is Page 506generated. These two timing functions are known as **astable** and **monostable** modes, respectively, of a so-called *multivibrator* circuit.

Monostable multivibrators are usually employed in IC package form. An IC one-shot can generate voltage pulses when triggered by the **rising** or a **falling edge** of an external input voltage. Various input connections are usually provided for selecting the preferred triggering mode. Multivibrators make use of a time constant that is usually set by an external *RC* network to determine the output pulse duration. Figure 7.44 shows the response of a one-shot to a triggering signal for the four conditions that may be attained with a typical one-shot.

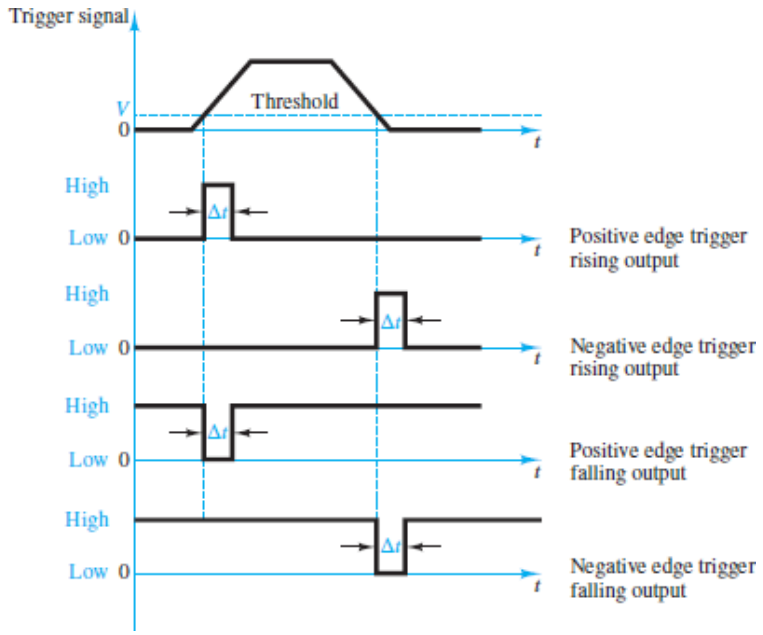


Figure 7.44 IC monostable multivibrator waveforms

Page 507

A typical IC one-shot circuit based on the 74123 is displayed in Figure 7.45. The 74123 is a **dual one-shot**, meaning that the package contains two monostable multivibrators, which can be used independently. The outputs of the one-shot are indicated by the symbols Q_1, $\bar{Q}_1$, Q_2, and $\bar{Q}_2$, where the overbar indicates the

complement of the output. For example, if Q_1 corresponds to a positive-going output pulse, $\bar{Q}_1$ indicates a negative-going output pulse of equal duration.

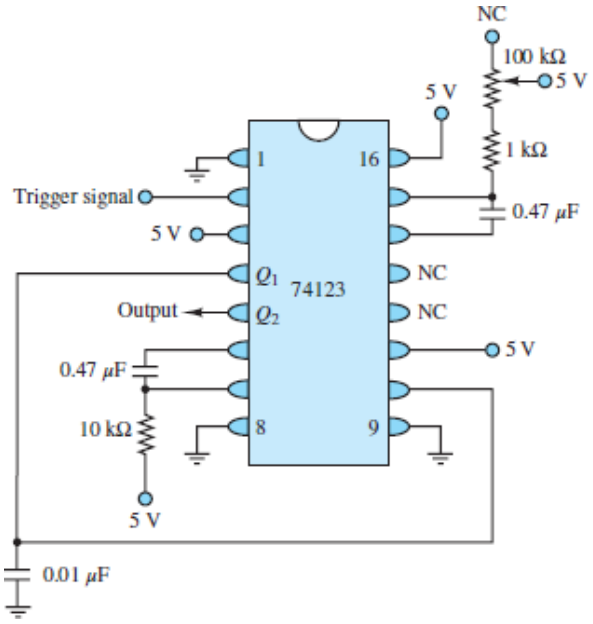


Figure 7.45 Dual one-shot circuit

Timer ICs: The NE555

The NE555 is a multivibrator IC capable of operating in either a monostable or astable mode. In the monostable mode the time delay or pulse duration is determined by an external *RC* network. In the astable mode the pulse train frequency is typically determined by two external resistors and one capacitor. Figure 7.46 depicts typical circuits for monostable and astable operation of the NE555. The threshold and trigger pins are used to set the voltage levels at which the transitions of the NE555 output occur. For the monostable circuit, the pulse width is

$$T = 1.1R_1C \tag{7.35}$$

For the astable circuit, the positive and negative pulse widths are

$$T_+ = 0.69(R_1 + R_2)C \tag{7.36}$$

$$T_{-} = 0.69R_2C \tag{7.37}$$

Page 508

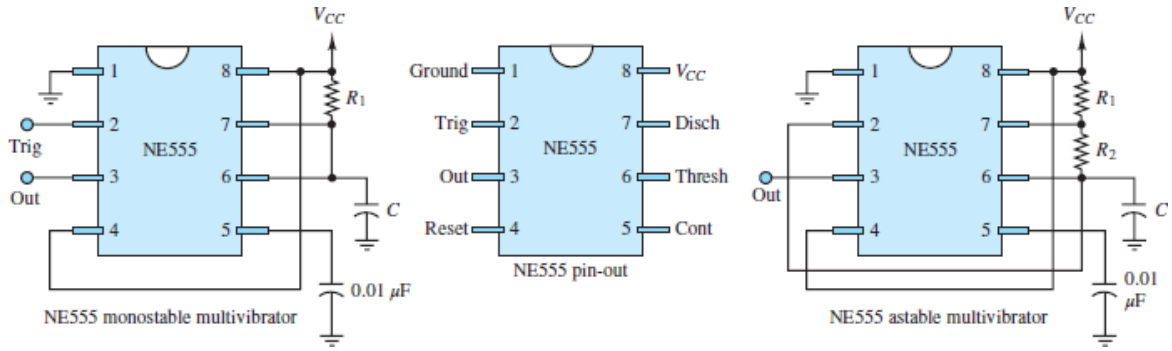


Figure 7.46 NE555 timer

EXAMPLE 7.9 Comparator With Offset

Problem

Sketch the input and output waveforms of the comparator with offset shown in Figure 7.47.

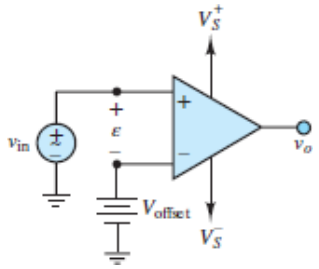


Figure 7.47 Comparator with offset

Solution

Known Quantities: Input voltage, voltage offset.

Find: Output voltage $v_o(t)$.

Schematics, Diagrams, Circuits, and Given Data: $v_{in}(t) = \sin \omega t$; $V_{offset} = 0.6$ V.

Analysis: First, compute the differential voltage across the inputs of the op-amp:

$$\varepsilon = v_{in} - V_{offset}$$

Then, use equation 7.25 to determine the switching conditions for the comparator:

$$v_{in} > V_{offset} \quad \Rightarrow \quad v_o = V_{sat}^+$$
$$v_{in} < V_{offset} \quad \Rightarrow \quad v_o = V_{sat}^-$$

Thus, the comparator will switch whenever the sinusoidal voltage rises above or falls below the reference voltage. Figure 7.48 depicts the appearance of the comparator output voltage. Note that the comparator output waveform is no longer a symmetric square wave.

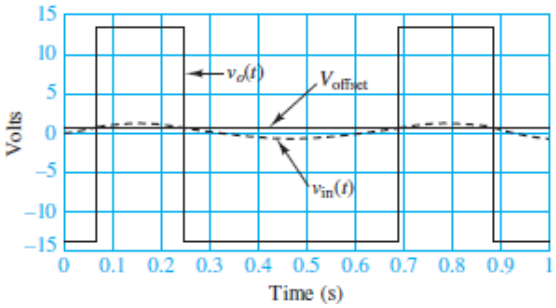


Figure 7.48 Waveforms of comparator with offset

Comments: Since it is often not practical to use a separate external reference voltage source, a potentiometer can be used between the supply voltages to achieve V_{ref} by voltage division.

Page 509

EXAMPLE 7.10 Analysis and Design of Schmitt Trigger

Problem

Find the required resistor values for the Schmitt trigger circuit shown in Figure 7.49.

Solution

Known Quantities: Supply voltages and supply saturation voltages; reference voltage (offset); noise amplitude.

Find: R_1, R_2, R_3.

Schematics, Diagrams, Circuits, and Given Data: $|V_S| = 18$ V; $|V_{sat}| = 16.5$ V; $V_{offset} = 2$ V.

.

Assumptions: $|v_{noise}| = 100$ mV..

Analysis: To avoid using a separate independent reference voltage source, the resistor R_3 is included to establish an offset voltage V_{offset} by voltage division at the noninverting terminal. From the circuit of Figure 7.49 the noninverting terminal voltage is expressed using the principle of superposition.

$$v^+ = \frac{R_2 \| R_3}{R_1 + R_2 \| R_3} v_o + \frac{R_1 \| R_2}{R_1 \| R_2 + R_3} V_S^+$$

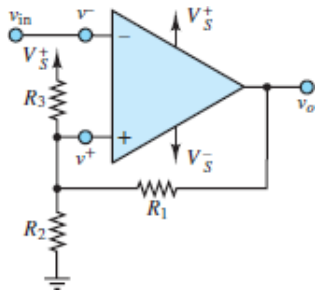


Figure 7.49 Schmitt trigger

The required noise protection level is $\Delta V = \pm 100$ mV, *which is one-half of the width of the transfer characteristic symmetrically placed about Voffset in Figure 7.42. Thus:*

$$\frac{\Delta V}{2} = \frac{R_2 \| R_3}{R_1 + R_2 \| R_3} V_{sat} = \frac{R_2 \| R_3}{R_1 + R_2 \| R_3} \times 16.5 = 0.1 \text{ V}$$

The offset voltage itself is related to the resistor network by

$$V_{offset} = \frac{R_1 \| R_2}{R_3 + R_1 \| R_2} V_S^+$$

Figure 7.50 depicts the ±100-mV noise protection band around the 2 V offset voltage. The two equations for ΔV and V_{offset} can be solved for R_2 and R_3 once a value is picked for R_1. Notice that when $R_1 \gg R_2$ and v_o is set to zero R_2 and R_3 will act like a simple series voltage divider. Likewise, when $R_3 \gg R_2$ and v_s^+ is set to zero, R_1 and R_2 will act like a simple series voltage divider. ($R_a \;/\!/\; R_b \approx R_a$ when $R_b \gg R_a$). These ideas suggest that $R_1 \gg R_3 \gg R_2$ should satisfy the requirements for ΔV and V_{offset}.

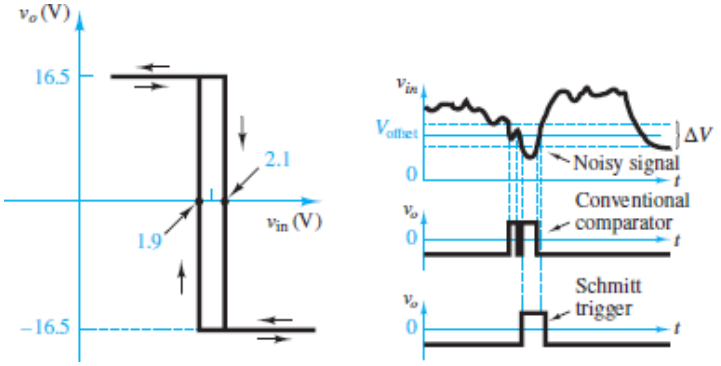


Figure 7.50 Schmitt trigger waveforms and transfer characteristics

Page 510

After some algebraic manipulations the following relationships can be derived.

$$\frac{R_2}{R_1} = \frac{b}{1-a} \quad \text{and} \quad \frac{R_3}{R_1} = \frac{b}{a}$$

where

$$a = \frac{V_{\text{sat}}}{V_S^+}\,\frac{V_{\text{offset}}}{V_{\text{sat}} - \Delta V/2} \quad \text{and} \quad b = \frac{\Delta V/2}{V_{\text{sat}} - \Delta V/2}$$

Assume R_1 = 100.0 kΩ and plug in values for v_{sat}, v_S^+, v_{offset} and ΔV to find $R_3 \approx 2.73$ kΩ and $R_2 \approx 342$ Ω. Using standard discrete resistor values this circuit can be implemented by selecting R_3 = 2.7 kΩ and R_2 = 330 Ω. The transfer characteristic of the comparator and the associated waveforms are shown in Figure 7.50. In practice, some adjustment of the resistor values is necessary to account for the impact of any specific op amp.

Comments: In Figure 7.50 the Schmitt trigger output is compared to that of a comparator without noise protection. For the input used in this example, the conventional comparator is triggered twice in the presence of noise.

EXAMPLE 7.11 Analysis of the 555 Timer

Problem

Calculate the component values required to obtain a 0.421-ms pulse using the 555 timer monostable configuration of Figure 7.46.

Solution

Known Quantities: Desired pulse duration T.

Find: Values of R_1 and C.

Schematics, Diagrams, Circuits, and Given Data: $T = 0.421$ ms.

Assumptions: Assume a value for C.

Analysis: Using equation 7.35,

$$T = 1.1R_1C$$

And assuming $C = 1\ \mu$F, calculate

$$0.421 \times 10^{-3} = 1.1\,R_1 \times 10^{-6}$$

or

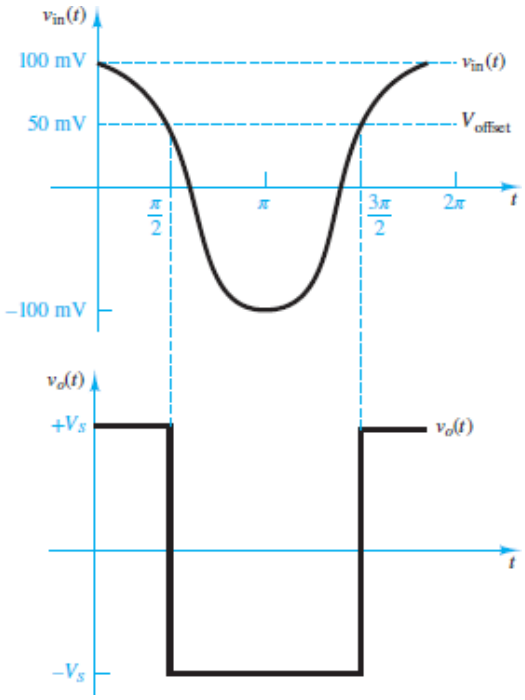
$$R_1 = 382.73\ \Omega$$

Comments: Any reasonable combination of R_1 and C values can yield the desired design value of T. Thus, the component selection shown in this example is not unique.

CHECK YOUR UNDERSTANDING

For the comparator circuit of Figure 7.47, sketch the waveforms $v_o(t)$ and $v_{\text{in}}(t)$ if $v_{\text{in}}(t) = 0.1 \cos \omega t$ and $V_{\text{offset}} = 50$ mV. Assume that $|V_S| = 15$ V.

Page 511



CHECK YOUR UNDERSTANDING

Derive the expressions for the switching thresholds of the Schmitt trigger of Figure 7.41.

Answer:

Real op-amp $\Rightarrow$

$$\left.\begin{aligned} V^+ - V^- &< 0 \;\leftarrow\; v_o = +V_s \\ V^+ - V^- &> 0 \;\leftarrow\; v_o = -V_s \end{aligned}\right\}$$

$$\left.\begin{aligned} V^+ &= R_2 \cdot I \\ I &= \frac{v_o}{R_1 + R_2} \end{aligned}\right\} \Rightarrow \boxed{V^+ = v_o \frac{R_2}{R_1 + R_2}}$$

Considering $V^2 = V_{in}$, write:

$$\left.\begin{aligned} v_o \cdot \frac{R_2}{R_1 + R_2} - V_{in} &> 0 \;\leftarrow\; v_o = -V_s \\ v_o \cdot \frac{R_2}{R_1 + R_2} - V_{in} &< 0 \;\leftarrow\; v_o = +V_s \end{aligned}\right\}$$

$$\text{At } t = 0 \;\leftarrow\; v_o = 0 \;\leftarrow\; V^+ = 0,$$

$$\left.\begin{aligned} \text{If } -V_{in} &< 0 \leftarrow v_o = +V_s \\ \text{If } -V_{in} &> 0 \leftarrow v_o = -V_s \end{aligned}\right\}$$

$$\text{At } t > 0 \;\leftarrow\; \boxed{v_o = \pm V_s} \;\leftarrow\; \left.\begin{aligned} \text{If } V_{in} &< v_o \cdot \frac{R_2}{R_1 + R_2} \leftarrow v_o = +V_s \\ \text{If } V_{in} &> v_o \cdot \frac{R_2}{R_1 + R_2} \leftarrow v_o = -V_s \end{aligned}\right\}$$

Page 512

Conclusion

In this chapter, you have learned a number of important facts about instrumentation systems.

Measurements and instrumentation are among the most important areas of electrical engineering because virtually all engineering disciplines require the ability to perform measurements of some kind.

A measurement system consists of three essential elements: a sensor, signal conditioning circuits, and recording or display devices. The last are often based on digital computers.

Sensors are devices that convert a change in a physical variable to a corresponding change in an electrical variable, typically a voltage. A broad range of sensors exist to measure virtually all physical phenomena. Proper wiring,

grounding, and shielding techniques are required to minimize undesired interference and noise.

Often, sensor outputs need to be conditioned before further processing can take place. The most common signal conditioning circuits are instrumentation amplifiers and active filters.

If the conditioned sensor signals are to be recorded in digital form by a computer, it is necessary to perform an analog-to-digital conversion process; timing and comparator circuits are also often used in this context.

Upon completing this chapter you should have mastered the following learning objectives:

1. Be familiar with major classes of sensors.
2. Know how to properly ground circuits and be familiar with noise shielding and noise reduction methods.
3. Design signal conditioning amplifiers and filters.
4. Understand the principles of A/D and D/A conversion and know how to select the specifications of an A/D or D/A system.
5. Know how to analyze and design simple comparator and timing circuits, and how to use other common integrated circuits.

HOMEWORK PROBLEMS

Section 7.1: Measurement Systems and Transducers

7.1 Most motorcycles have engine speed tachometers, as well as speedometers, as part of their instrumentation. What differences, if any, are there between the two in terms of transducers?

7.2 Explain the differences between the engineering specifications you would write for a transducer to measure the frequency of an audible sound wave and a transducer to measure the frequency of a visible light wave.

7.3 A measurement of interest in the summer is the temperature-humidity index, consisting of the sum of the temperature and the relative humidity percentage. How would you measure this? Sketch a simple schematic diagram.

7.4 Consider a capacitive displacement transducer as shown in Figure P7.4. Its capacitance is

$$C = \frac{0.255A}{d} \text{ F}$$

where A = cross-sectional area of the transducer plate (in^2) and d = air-gap length (in). Determine the change in voltage Δv_0 when the air gap changes from 0.01 to 0.015 in.

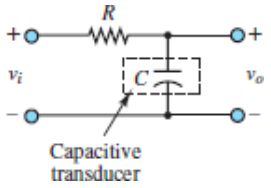


Figure P7.4

Page 513

7.5 The circuit of Figure P7.5 may be used for operation of a photodiode. The voltage V_D is a reverse-bias voltage large enough to make the diode current i_D proportional to the incident light intensity H. Under this condition, i_D/H = 0.5 μA-m^2/W.

a. Show that the output voltage v_o varies linearly with H.

b. If H = 1,500 W/m^2, V_D = 7.5 V, and an output voltage of 1 V is desired, determine an appropriate value for R_o.

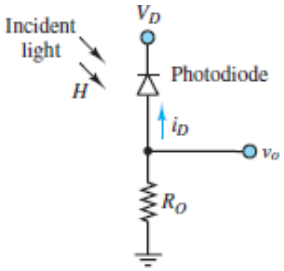


Figure P7.5

7.6 A material constant G is equal to 0.055 V-m/N for quartz in compressive stress and 0.22 V-m/N for polyvinylidene fluoride in axial stress.

a. A force sensor uses a piezoelectric quartz crystal as the sensing element. The quartz element is 0.25 in thick and has a rectangular cross section of 0.09 in^2. The sensing element is compressed, and the

output voltage is measured across the thickness. What is the output of the sensor in volts per newton?

b. A polyvinylidene fluoride film is used as a piezoelectric load sensor. The film is 30 μm thick, 1.5 cm wide, and 2.5 cm in the axial direction. It is stretched by the load in the axial direction, and the output voltage is measured across the thickness. What is the output of the sensor in volts per newton?

7.7 Let b be the damping constant of the mounting structure of a machine as pictured in Figure P7.7. It must be determined experimentally. First, the spring constant K is determined by measuring the resultant displacement under a static load. The mass m is directly measured. Finally, the damping ratio ζ is measured using an impact test. The damping constant is given by $b = 2\zeta\sqrt{Km}$. If the allowable levels of error in the measurements of K, m, and ζ are ±5 percent, ±2 percent, and ±10 percent, respectively, estimate a percentage error limit for b.

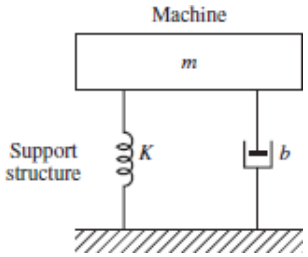


Figure P7.7

7.8 The quality control system in a plant that makes acoustical ceiling tile uses a proximity sensor to measure the thickness of the wet pulp layer every 2 ft along the sheet, and the roller speed is adjusted based on the last 20 measurements. Briefly, the speed is adjusted unless the probability that the mean thickness lies within ±2 percent of the sample mean exceeds 0.99. A typical set of measurements (in millimeters) is as follows:

8.2, 9.8, 9.92, 10.1, 9.98, 10.2, 10.2, 10.16, 10.0, 9.94, 9.9, 9.8, 10.1, 10.0, 10.2, 10.3, 9.94, 10.14, 10.22, 9.8

Would the speed of the rollers be adjusted based on these measurements? Justify your answer.

7.9 Discuss and contrast the following terms:

a. Measurement accuracy.

b. Instrument accuracy.

c. Measurement error.

d. Precision.

7.10 Four sets of measurements were taken on the same response variable of a process using four different sensors. The true value of the response was known to be constant. The four sets of data are shown in Figure P7.10. Rank these data sets (and hence the sensors) with respect to

a. Precision.

b. Accuracy.

Page 514

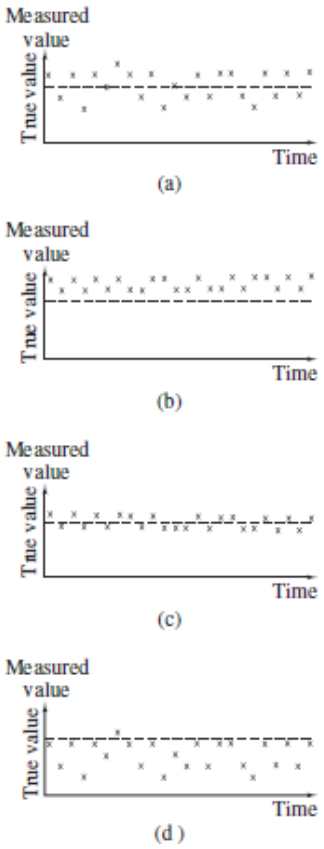


Figure P7.10

Section 7.3: Signal Conditioning

7.11 For the instrumentation amplifier (IA) of Figure P7.11, find the gain of the input stage if $R_1 = 1$ kΩ and $R_2 = 5$ kΩ.

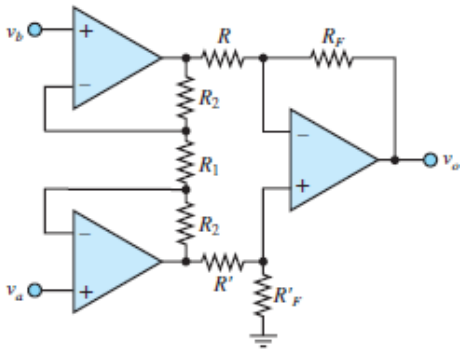


Figure P7.11

7.12 Consider the IA of Figure P7.11. Let $R_1 = 1$ kΩ. What value of R_2 should be used to make the gain of the input stage equal 50?

7.13 Consider the IA of Figure P7.11. Let $R_2 = 10$ kΩ. What value of R_1 will yield an input-stage gain of 16?

7.14 For the IA of Figure P7.11, find the gain of the input stage if $R_1 = 1$ kΩ and $R_2 = 10$ kΩ.

7.15 For the IA of Figure P7.11, find the gain of the input stage if $R_1 = 1.5$ kΩ and $R_2 = 80$ kΩ.

7.16 Find the differential gain for the IA of Figure P7.11 if $R_2 = 5$ kΩ, $R_1 = R' = R = 1$ kΩ, and $R_F = 10$ kΩ.

7.17 Refer to equations 7.4–7.16 and suppose, for the circuit of Figure P7.11, that $R_F = 200$ kΩ, $R = 1$ kΩ, and $\Delta R = 0.02R$. Calculate the CMRR of the IA. Express your result in decibels.

7.18 Given the IA of Figure P7.11, with the component values of Problem 7.17, calculate the mismatch in gains for the differential components. Express your result in decibels.

7.19 Given $R_F = 10$ kΩ and $R_1 = 2$ kΩ for the IA of Figure 7.11, find R and R_2 so that a differential gain of 900 can be achieved.

Section 7.4: Analog-to-Digital and Digital-to-Analog Conversion

7.20 List two advantages of digital signal processing over analog signal processing.

7.21 Discuss the role of a multiplexer in a data acquisition system.

7.22 Discuss the purpose of using sample-and-hold circuits in data acquisition systems.

7.23 The circuit shown in Figure P7.23 represents a sample-and-hold circuit, such as might be used in a successive-approximation ADC. Assume that the NMOS is turned *on* (i.e., acting as a short-circuit) when v_G is high and *off* (i.e., acting as an open-circuit) when v_G is low. Explain the operation of the circuit.

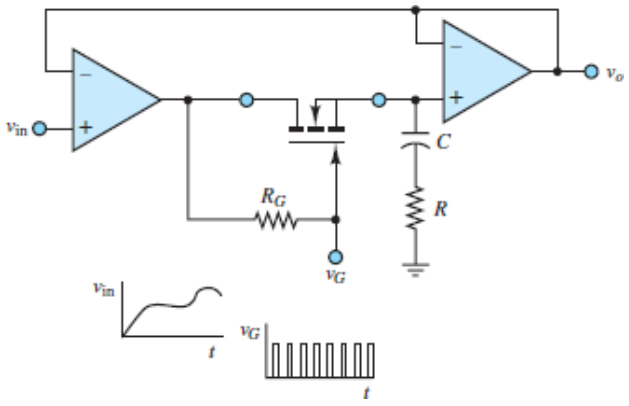


Figure P7.23

Page 515

7.24 For the circuit shown in Figure P7.23, let v_{in} be a 1-kHz sinusoidal signal with 0° phase angle, 0 VDC offset, and 20-V peak-to-peak amplitude. Let v_G be a rectangular pulse train, with pulse width 10 μs and period 100 μs with the leading edge of the first pulse at $t = 0$.

a. Sketch v_o if the *RC* circuit has a time constant equal to 20 μs.

b. Sketch v_o if the *RC* circuit has a time constant equal to 1 ms.

7.25 The unsigned decimal number 12_{10} is inputted to a 4-bit DAC. (See Figure 7.20.) Given that $R_F = R_0/15$, logic 0 corresponds to 0 V, and logic 1 corresponds to 4.5 V,

a. What is the output of the DAC?

b. What is the maximum voltage that can be outputted from the DAC?

c. What is the resolution over the range 0 to 4.5 V?

d. Find the number of bits required in the DAC if an improved resolution of 20 mV is desired.

7.26 The unsigned decimal number 215_{10} is inputted to an 8-bit DAC. (See Figure 7.20.) Given that $R_F = R_0/255$, logic 0 corresponds to 0 V, and logic 1 corresponds to 10 V,

a. What is the output of the DAC?

b. What is the maximum voltage that can be outputted from the DAC?

c. What is the resolution over the range 0 to 10 V?

d. Find the number of bits required in the DAC if an improved resolution of 3 mV is desired.

7.27 The circuit shown in Figure P7.27 represents a simple 4-bit DAC. Each switch is controlled by the corresponding bit of the digital number—if the bit is 1, the switch is up; if the bit is 0, the switch is down. Let the digital number be represented by $b_3b_2b_1b_0$. Determine an expression relating v_o to the binary input bits.

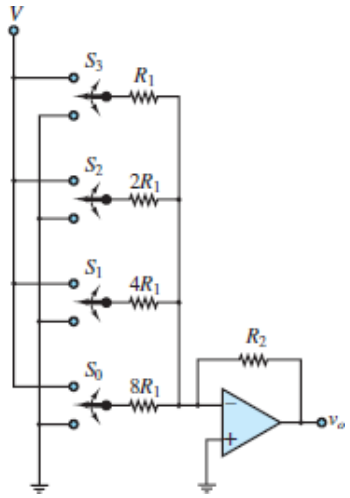


Figure P7.27

7.28 The unsigned decimal number 98_{10} is inputted to an 8-bit DAC. (See Figure 7.20.) Given that $R_F = R_0/255$, logic 0 corresponds to 0 V, and logic 1 corresponds to 4.5 V,

a. What is the output of the DAC?

b. What is the maximum voltage that can be outputted from the DAC?

c. What is the resolution over the range 0 to 4.5 V?

d. Find the number of bits required in the DAC if an improved resolution of 0.5 mV is desired.

7.29 For the DAC circuit shown in Figure P7.29 (using an ideal op-amp), what value of R_F will give an output range of $-10 \leq v_o < 0$ V? Assume that logic 0 = 0 V and logic 1 = 5 V.

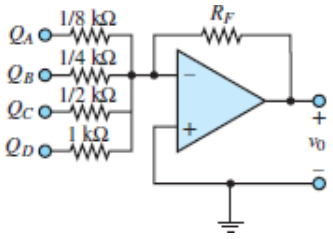


Figure P7.29

7.30 Explain how to redesign the circuit of Figure P 7.27 so that the overall circuit is a noninverting device.

7.31 The circuit of Figure P7.31 has been suggested as a means of implementing the NMOS switches needed for the 4-bit DAC of Figure P7.27. Assume the NMOS transistors act as short- and open-circuits when their gate inputs are logic 1 and 0, respectively. Explain how the circuit works. $k = 0, \cdots, 3$.

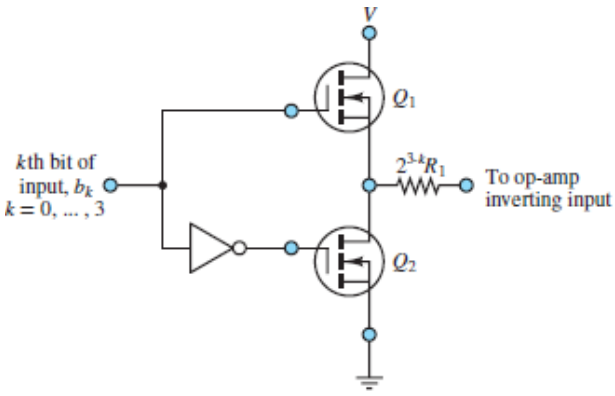


Figure P7.31

7.32 The unsigned decimal number 345_{10} is inputted to a 12-bit DAC. (See Figure 7.20.) Given that $R_F = R_0/4{,}095$, logic 1 corresponds to 10 V, and logic 0 corresponds to 0 V,

a. What is the output of the DAC?

b. What is the maximum voltage that can be outputted from the DAC?

Page 516

c. What is the resolution over the range 0 to 10 V?

d. Find the number of bits required in the DAC if an improved resolution of 0.5 mV is desired.

7.33 For the DAC circuit shown in Figure P7.29 (using an ideal op-amp), what value of R_F will give an output range of $-15 \leq v_o \leq 0$ V?

7.34 Using the model of Figure P7.27, design a 4-bit DAC whose output is given by

$$v_o = -\frac{1}{10}(8b_3 + 4b_2 + 2b_1 + b_0)\,V$$

7.35 A data acquisition system uses a DAC with a range of ±15 V and a resolution of 0.01 V. How many bits must be present in the DAC?

7.36 A data acquisition system uses a DAC with a range of ±10 V and a resolution of 0.04 V. How many bits must be present in the DAC?

7.37 A data acquisition system uses a DAC with a range of −10 to +15 V and a resolution of 0.004 V. How many bits must be present in the DAC?

7.38 A DAC is to be used to deliver velocity commands to a motor. The maximum velocity is to be 2,500 r/min, and the minimum nonzero velocity is to be 1 r/min. How many bits are required in the DAC? What will the resolution be?

7.39 Assume the full-scale value of the analog input voltage to a particular ADC is 10 V.

a. If this is a 3-bit device, what is the resolution of the output?

b. If this is an 8-bit device, what is its resolution?

c. Make a general comment about the relationship between the number of bits and the resolution of an ADC.

7.40 The voltage range of feedback signals from a process is −5 to +15 V, and a resolution of 0.05 percent of the voltage range is required. How many bits are required for the DAC?

7.41 Eight channels of analog information are being used by a computer to close eight control loops. Assume that all analog signals have identical frequency content and are multiplexed into a single ADC. The ADC requires 100 μs per conversion. The closed-loop software requires 500 μs of computation and output time for four of the loops, and for the other four it requires 250 μs. What is the maximum frequency content that the analog signal can have according to the Nyquist criterion?

7.42 A rotary potentiometer is to be used as a remote rotational displacement sensor. The maximum displacement to be measured is 180°, and the potentiometer is rated for 10 V and 270° of rotation.

a. What voltage increment must be resolved by an ADC to resolve an angular displacement of 0.5°? How many bits would be required in the ADC for full-range detection?

b. The ADC requires a 10-V input voltage for full-scale binary output. If an amplifier is placed between the potentiometer and the ADC, what amplifier gain should be used to take advantage of the full range of the ADC?

7.43 Suppose it is desired to digitize a 250-kHz analog signal to 10 bits using a successive-approximation ADC. Estimate the maximum permissible conversion time for the ADC.

7.44 A torque sensor has been mounted on a farm tractor engine. The voltage produced by the torque sensor is to be sampled by an ADC. The rotational speed of the crankshaft is 800 r/min. Because of speed fluctuations caused by the reciprocating action of the engine, frequency content is present in the torque signal at twice the shaft rotation frequency. What is the minimum sampling period that can be used to ensure that the Nyquist criterion is satisfied?

7.45 The output voltage of an aircraft altimeter is to be sampled using an ADC. The sensor outputs 0 V at 0-m altitude and outputs 10 V at 10,000-m altitude. If the allowable error in sensing $(\pm\frac{1}{2}\,\text{LSB})$ is 10 m, find the minimum number of bits required for the ADC.

7.46 Consider a circuit that generates interrupts at fixed time intervals. Such a device is called a *real-time clock* and is used in control applications to

establish the sample period as T seconds for control algorithms. Show how this can be done with a square wave (clock) that has a period equal to the desired time interval between interrupts.

7.47 Find the minimum number of bits required to digitize an analog signal with a resolution of

a. 5 percent

b. 2 percent

c. 1 percent

Page 517

Section 7.5: Comparator and Timing Circuits

7.48 A useful application that exploits the open-loop characteristics of op-amps is known as a comparator. One particularly simple example, known as a window comparator, is shown in Figure P7.48(a) and (b). Show that $v_o = 0$ whenever $V_{low} < v_{in} < V_{high}$ and that $v_o = +V$ otherwise.

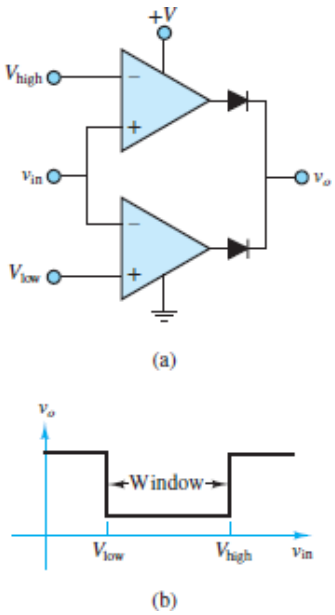


Figure P7.48

7.49 Design a Schmitt trigger to operate in the presence of noise with peak amplitude = ±150 mV. The circuit is to switch around the reference value −1 V. Assume an op-amp with ±10-V supplies (V_{sat} = 8.5 V).

7.50 In the circuit of Figure P7.50, $R_1 = 100\ \Omega$, $R_2 = 56\ \text{k}\Omega$, $R_i = R_1 /\!/ R_2$, and v_{in} is a 1-V peak-to-peak sine wave. Assuming that the supply voltages are ±15 V, determine the threshold voltages (positive and negative v^+) and draw the output waveform. (See Example 6.15 to understand the role of R_i.)

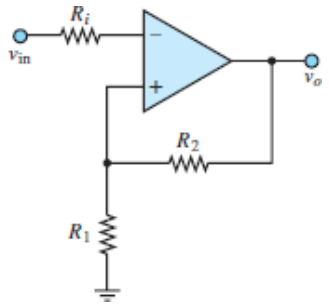


Figure P7.50

7.51 The circuit in Figure P7.51 shows how a Schmitt trigger might be constructed with an op-amp. Explain the operation of this circuit.

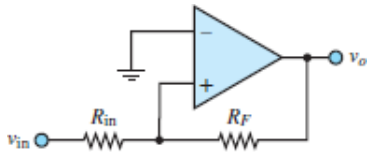


Figure P7.51

7.52 Consider the circuit of Figure P7.51. Let the op-amp be an LM741 with ±15-V bias supplies, and suppose R_F is chosen to be 104 kΩ. Assume v_{in} is a 1-kHz sinusoidal signal with 1-V amplitude.

a. Determine the appropriate value for R_{in} if the output is to be high whenever $|v_{\text{in}}| \geq 0.25\ \text{V}$.

b. Sketch the input and output waveforms.

7.53 For the circuit shown in Figure P7.53,

a. Draw the output waveform for v_{in}, a 4-V peak-to-peak sine wave at 100 Hz and $V_{\text{ref}} = 2$ V.

b. Draw the output waveform for v_{in}, a 4-V peak-to-peak sine wave at 100 Hz and $V_{\text{ref}} = -2$ V.

Note that the silicon diodes placed at the input ensure that the differential voltage does not exceed the diode offset voltage, $V_\gamma = 0.7\ V$. (See chapter 8.)

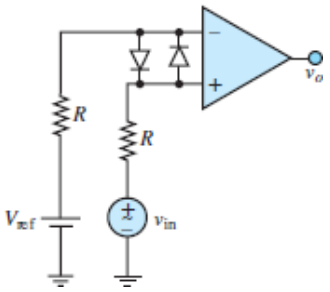


Figure P7.53

7.54 Figure P7.54 shows a simple *go-no go* detector application of a comparator.

a. Explain how the circuit works.

b. Design a circuit (i.e., choose proper values for the resistors) such that the green LED will turn on when v_{in} exceeds 5 V and the red LED will be on whenever v_{in} is less than 5 V. Assume only 15-V supplies are available.

Page 518

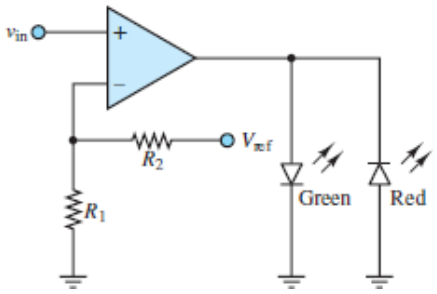


Figure P7.54

7.55 For the circuit of Figure P7.55, v_{in} is a 100-mV-peak sine wave at 5 kHz, R = 10 kΩ, and D_1 and D_2 are 6.2-V Zener diodes. Draw the output voltage waveform.

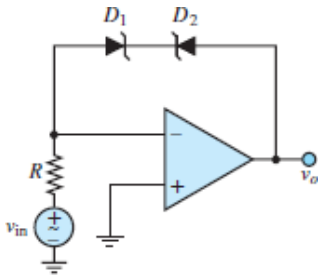


Figure P7.55

7.56 An op-amp multivibrator can be constructed by first adding a resistor R_3 between the output and inverting terminals of the Schmitt trigger shown in Figure 7.41. A capacitor C is then added between the inverting terminal and reference. In this way, v_{in} tracks the capacitor voltage as the capacitor charges and discharges. Show that the period of oscillation of an op-amp astable multivibrator is given by the expression

$$T = 2R_3C\log_e\left(\frac{2R_2}{R_1} + 1\right)$$

7.57 Use the data sheets for the 74123 monostable multivibrator to analyze the connection shown in Figure 7.45 in the text. Draw a timing diagram indicating the approximate duration of each pulse, assuming that the trigger signal consists of a positive-going transition.

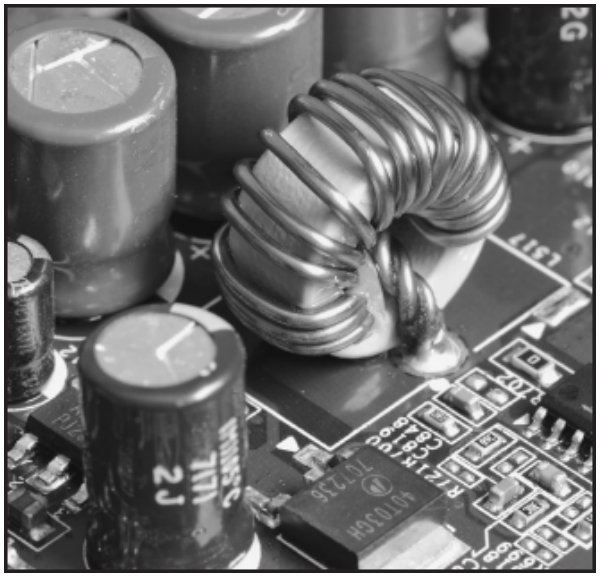
7.58 In the monostable multivibrator (one-shot) configuration of the NE555 timer shown in Figure 7.46 in the text, assume R_1 = 10 kΩ and that the output pulse width is T = 10 ms. Determine the capacitance C.

[1]See the Focus on Measurements box, "Resistance Strain Gauges."

[2]See the Focus on Measurements box "Magnetic Reluctance Position Sensors" in Chapter 14.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

PART III
ANALOG ELECTRONICS



CobraCZ/Shutterstock

Page 521

C H A P T E R
8

SEMICONDUCTORS AND DIODES

Marvelous advances have taken place in the field of **solid state electronics** ever since the invention of the *diode* and *transistor.* Modern analog and digital electronic systems are possible because these discrete electronic elements have been integrated into complex devices and systems. Although discrete electronic elements have been replaced in many applications by *integrated circuits* (e.g., operational amplifiers), it is nonetheless important to understand how these elements function. The aim of Part III of this textbook is to explore the behavior and applications of diodes, transistors, and other *electronic devices.*

This chapter explains the workings of the semiconductor diode, a device that finds use in many practical circuits used in electric power systems and in high- and low-power electronic circuits. While the *i-v* characteristic of a diode is inherently nonlinear, simple *linear models* can be used to approximate the diode characteristic and thus produce *linear circuits* that can be analyzed using the analytical tools developed in earlier chapters.

Page 522

Learning Objectives

Students will learn to...

1. Understand the basic principles underlying the physics of semiconductor devices in general and of the *pn* junction in particular. Become familiar with the forward-bias exponential diode equation and a typical diode *i-v* characteristic *Sections 8.1–8.2*.
2. Use linear large-signal models of the semiconductor diode in simple circuits *Section 8.3*.
3. Linearize the forward-bias exponential diode model in the neighborhood of an operating point to analyze the impact of small variations in the diode voltage or the diode current. *Section 8.4*.
4. Study practical full-wave rectifier circuits and learn to analyze and determine the practical specifications of a rectifier by using large-signal diode models. *Section 8.5*.
5. Understand the basic operation of Zener diodes as voltage references and use simple circuit models to analyze elementary voltage regulators. *Section 8.6*.
6. Understand the basic principle of operation of photodiodes, including solar cells photosensors, and light-emitting diodes. *Section 8.7*.

8.1 ELECTRICAL CONDUCTION IN SEMICONDUCTOR DEVICES

Elemental[1] or intrinsic **semiconductors** are those elements, specifically silicon and germanium, from group IV of the periodic table whose conductivity is much weaker than that of a typical conductor but significantly stronger than that of a typical insulator. For example, typical conductivities of copper (a good conductor) and glass (a common insulator) are 5.96×10^7 S/m and 10^{-13} S/m, respectively. By comparison, silicon and germanium, both semiconductors, have conductivities on the order of 10^{-3} S/m and 10^0 S/m, respectively. Another important property of silicon and germanium is that their conductivities *increase* with temperature, whereas the conductivity of most conductors (e.g., metals) decreases with temperature. It is important to note that most of the group IV elements are *not* semiconductors; tin and lead are metals whose conductivity is large and decreases with temperature.

Conducting materials have enough weakly bonded electrons in the outer conduction band that a modest electric field can easily produce a significant current. By contrast, the outer-band electrons in a semiconducting material are held by **covalent bonds** such that much stronger electric fields are needed to liberate them.

Figure 8.1 depicts the lattice arrangement for a pure silicon (Si) matrix. At sufficiently high temperatures, thermal energy causes the atoms in the lattice to vibrate; when sufficient kinetic energy is present, some of the valence electrons break their bonds with the lattice structure and become available as conduction electrons. These **free electrons** enable current flow in the semiconductor. As the temperature increases, more valence electrons are liberated, which explains why the conductivity of a semiconductor increases with temperature.

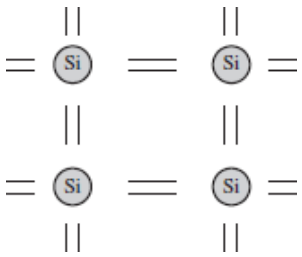


Figure 8.1 Lattice structure of silicon, with four valence electrons

However, the free valence electrons are not the only charge carriers present in a semiconductor. Whenever a free electron is liberated from the lattice, a corresponding Page 523net positive charge or **hole** within the lattice is also created as depicted by Figure 8.2. Holes act as positive charge carriers within a semiconducting material but with a different **mobility**—the ease with which charge carriers move through the lattice—than free electrons. Free electrons move far more easily around the lattice than holes. These two charge carriers also move in opposite directions when subjected to an external electric field, as illustrated in Figure 8.3.

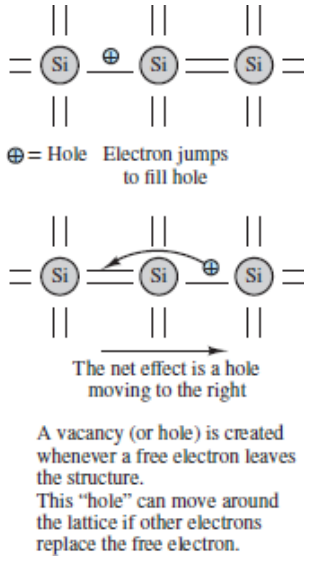
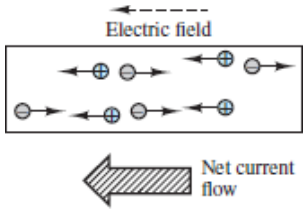


Figure 8.2 Free electrons and "holes" in the lattice structure



An external electric field forces holes to migrate to the left and free electrons to the right. The net current flow is to the left.

Figure 8.3 Current in a semiconductor

Occasionally, a free electron traveling in the immediate neighborhood of a hole will recombine with it to form a covalent bond. The result is two lost charge carriers. This additional phenomenon of **recombination** is proportional to the number of free electrons and holes and reduces the number of charge carriers in a semiconductor. However, in spite of recombination, at any given temperature a number of free electrons and holes will be available for conduction. The number of available charge carriers is called the **intrinsic concentration** n_i . The most commonly reported expression for n_i is

$$n_i \propto T^{1.5} e^{-E_g/2kT} \tag{8.1}$$

where T is temperature in K; E_2 is the bandgap energy, which for silicon is 1.12 eV; and k is Boltzmann's constant 8.62×10^{-5} eV/K. At $T = 300$ K, n_2 is approximately 1.5×10^{10} carriers /cm^3. Note the strong dependence on temperature.[2]

As noted, pure semiconductors are not particularly good conductors. To enhance the concentration of charge carriers and thus the conductivity, a semiconductor can be **doped**, whereby either *trivalent* (group III) or *pentavalent* (group V) impurities are added to the crystalline structure of the semiconductor.[3] Trivalent impurities, such as boron and gallium, add holes to the semiconductor's lattice and are known as *acceptors;* pentavalent impurities, such as phosphorus and arsenic, add free electrons, as depicted in Figure 8.4, and are known as *donors*.

An additional free electron is created when Si is "doped" with a group V element.

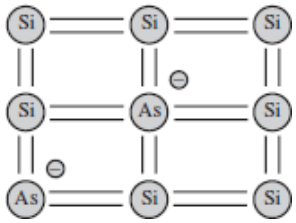


Figure 8.4 Doped semiconductor

Free electrons are the *majority charge carrier*, and holes are the *minority charge carrier* in semiconductors doped with donor elements. These materials are called ***n*-type semiconductors**. Likewise, holes are the majority charge carrier, and free electrons are the minority charge carrier in semiconductors doped with acceptor elements. These materials are called ***p*-type semiconductors**. In thermal equilibrium, the concentration of free electrons n (negative) is related to the concentration of holes p (positive) by:

$$pn = n_i^2 \tag{8.2}$$

In a doped semiconductor, the concentration of donated atoms is usually much greater than the intrinsic concentration of the semiconductor. In this case, the concentration of majority charge carriers is approximately the same as the concentration of donated atoms, which is determined by the doping process and is not a function of temperature. However, the concentration of minority charge carriers is determined by temperature and is usually much less than the intrinsic concentration of the semiconductor. For example, in an n-type material, the concentration of free Page 524electrons n_n is approximately equal to the concentration of donor atoms n_D . Since $p_n n_n = n_i^2$, the result is

$$n_n \approx n_D \gg n_i \quad \text{and} \quad p_n = \frac{n_i^2}{n_n} \approx \frac{n_i^2}{n_D} \ll n_i \quad n\text{-type} \tag{8.3}$$

Likewise, in a p-type material where n_A is the concentration of acceptor atoms:

$$p_p \approx n_A \gg n_i \quad \text{and} \quad n_p = \frac{n_i^2}{p_p} \approx \frac{n_i^2}{n_A} \ll n_i \quad p\text{-type} \tag{8.4}$$

In the previous two equations, the subscripts i, n, and p indicate whether the material is intrinsic (pure) semiconductor, n-type, or p-type, respectively.

It is important to keep in mind that doped n- and p-type materials are *electrically neutral* because the donor and acceptor elements are themselves electrically neutral. The material type simply indicates the nature of the *mobile* majority charge carriers present in the conduction band of the material lattice.

8.2 THE *PN* JUNCTION AND THE SEMICONDUCTOR DIODE

A simple section of n- or p-type material is not particularly useful for the construction of electronic circuits. However, when sections of n- and p-type material

are brought in contact to form a ***pn* junction**, a **diode** is formed. Diodes have a number of interesting and useful properties that are due entirely to the nature of the *pn* junction.

Figure 8.5 depicts an idealized *pn* junction. The difference in concentrations of free electrons in the *n*-type material compared to the *p*-type material results in a *diffusion* of free electrons from right to left across the junction. Likewise, the difference in concentration of holes across the junction results in diffusion of holes from left to right. In both cases, the **diffusion current** I_2 is directed left to right because a positive current is defined as either positive holes moving left to right or negative free electrons moving right to left.

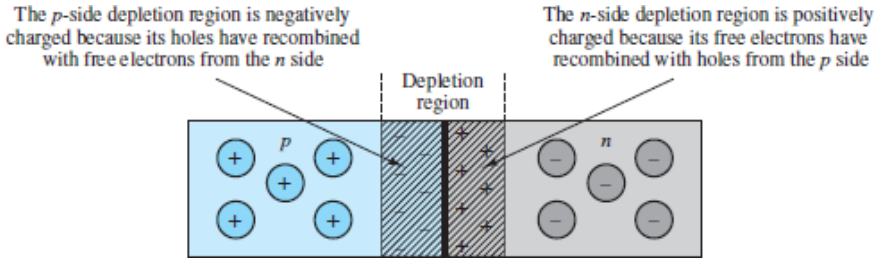


Figure 8.5 A *pn* junction

As free electrons leave the *n*-type material and enter the *p*-type material they tend to recombine with holes. Likewise, as holes leave the *p*-type material and enter the *n*-type material they tend to recombine with free electrons. Once free electrons and holes recombine they are no longer mobile, but held in place in the material lattice by covalent bonds. At first, most of the recombinations occur close to the junction. However, as time passes, more and more of the mobile charges near the junction have recombined such that diffusing mobile charges must travel further from the junction to encounter a partner with which to recombine. Thus, this diffusion process results in recombinations on both sides of the junction and, as the process continues, an expanding **depletion region** wherein virtually no mobile charge carriers remain. This region is *electrically charged* because mobile Page 525charge carriers that have recombined to form the region have no electrical counterpart in the lattice where they have become fixed. In Figure 8.5 this result is depicted by the negatively charged *p*-type region to the left of the junction and the positively charged *n*-type region to the right of the junction.

Once the depletion region begins to form, the resulting net charge separation produces an electric field pointing from the positively charged *n*-type to the negatively charged *p*-type portions of the depletion region. This electric field slows the ongoing diffusion of majority charge carriers by establishing a **potential barrier** or **contact potential** across the depletion region. This potential depends upon the

semiconductor material (about 0.6 to 0.7 V for silicon) and is also known as the *offset voltage* V_γ .

In addition to the diffusion current associated with majority charge carriers, an oppositely directed **drift current** I_S associated with minority charge carriers is established across the depletion region. Specifically, free electrons and holes are thermally generated in the *p*- and *n*-type materials, respectively. Any of these minority carriers that manage to reach the depletion region are swept across it by the electric field. Note that both components of the drift current contribute to a positive current from right to left because a positive current is defined as either positive holes moving right to left or negative free electrons moving left to right.

Figure 8.6 depicts the presence of both a diffusion current and drift current across the depletion region. Its equilibrium width is reached when the average net drift current exactly offsets the average net diffusion current. Recall that the magnitude of the diffusion current is largely determined by the concentration of the donor and acceptor elements while the magnitude of the drift current is highly temperature dependent. Thus, the equilibrium width of the depletion region depends upon both temperature and the doping process.

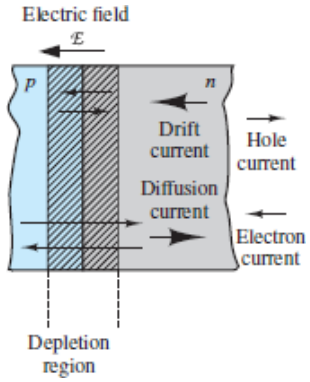


Figure 8.6 Drift and diffusion currents in a *pn* junction

Now consider the case shown in Figure 8.7(a) where a battery has been connected across a *pn* junction in the **reverse-biased** direction. Assume that suitable contacts between the battery and the *p*- and *n*-type materials are established. The reverse-bias orientation of the battery widens the depletion region and increases the potential barrier across it such that the majority carrier diffusion current decreases. On the other hand, the minority carrier drift current increases such that there is now a small (on the order of nano-amperes) nonzero current I_0 directed from the *n*- to *p*-type region. I_0 is small because it is comprised of minority carriers. Thus, when reverse-biased, the diode current i_D is

$$i_D = -I_0 = I_S \qquad \text{Reverse-biased diode current} \tag{8.5}$$

where I_S is known as the **reverse saturation current**.

When the *pn* junction is forward-biased as in Figure 8.7(b), the depletion region is narrowed and the potential barrier across it is lowered such that the majority carrier diffusion current increases. As the forward-biased diode voltage v_D is increased the diffusion current I_d increases exponentially:

$$I_d = I_0 e^{q_e v_D/kT} = I_0 e^{v_D/V_T} \tag{8.6}$$

where $q_e = 1.6 \times 10^{-19}$ C is the elementary charge, k = 1.38 times 10−23 J/K is the Boltzmann constant, T is the material temperature (in K), and $V_T = kT/q_e$ is the **thermal voltage**. At room temperature, $V_T \approx 25$ mV. The net diode current under forward bias is

$$i_D = I_d - I_0 = I_0(e^{v_D/V_T} - 1) \qquad \text{Diode equation} \tag{8.7}$$

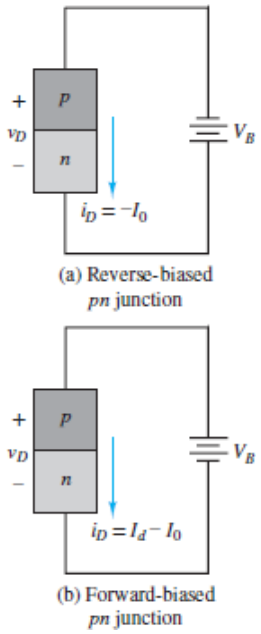


Figure 8.7 Forward- and reverse-biased *pn* junctions

Page 526

Figure 8.8 depicts the diode *i-v* characteristic described by the diode equation for a fairly typical silicon diode for $v_D > 0$. Since I_0 is typically very small (10^{-9} to 10^{-15}

A), the diode equation is often approximated by:

$$i_D = I_0 e^{v_D/V_T} \tag{8.8}$$

This expression is a good approximation for a silicon diode at room temperature when v_D is greater than a few tenths of a volt.

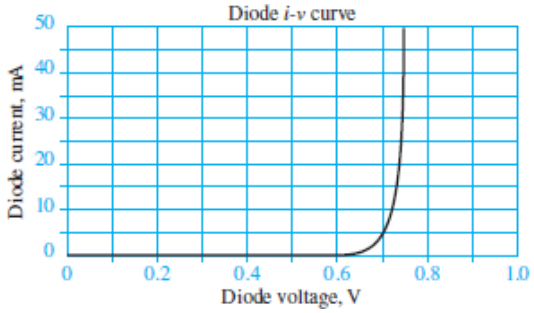


Figure 8.8 Typical diode *i-v* characteristic curve

The ability of the *pn* junction to conduct significant current only in the forward-biased direction allows it to function in electric circuits much like a check valve functions in mechanical circuits. A generic *pn* junction and the diode circuit symbol are shown in Figure 8.9. Notice that the triangle shape suggests the direction of forward-biased current. Positive current i_2 passes from the **anode** to the **cathode**, where the term *cathode* always refers to the source of electrons (negative charge carriers) whether used in reference to a diode or battery.[4]

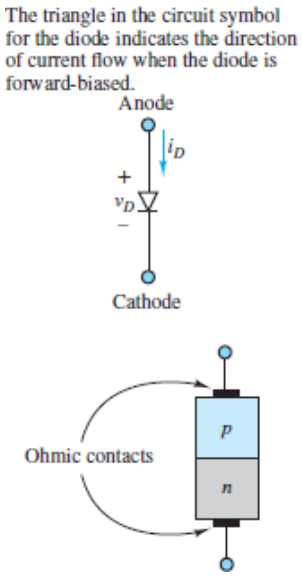


Figure 8.9 Diode circuit symbol

Figure 8.10 shows the complete *i-v* characteristic of a diode. Note that the diode current is approximately zero when $v_D < 0$ unless v_D is sufficiently large and negative (reverse-biased) such that **reverse breakdown** occurs. When $v_D < -V_Z$, the diode conducts current *in the reverse-biased direction.* Two effects contribute to this reverse-biased current: the *Zener effect* and *avalanche breakdown.* In silicon diodes, the Zener effect tends to dominate when $V_Z < 5.6$ V while avalanche breakdown tends to dominate at larger, more negative diode voltages.

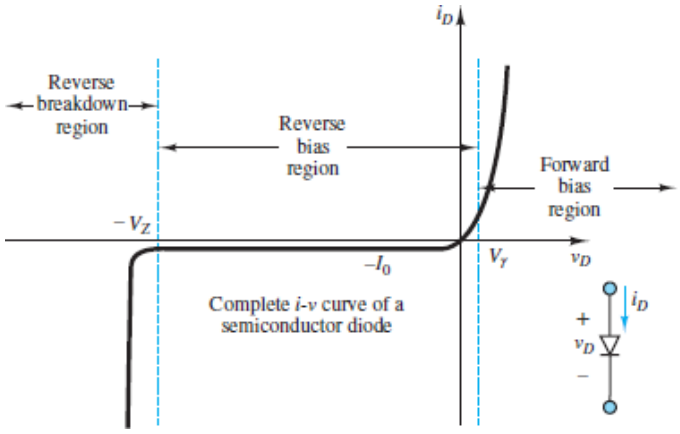


Figure 8.10 The diode *i-v* characteristic

The root causes of these two effects, while similar, are not the same. The Zener effect is significant when the depletion region is designed to be heavily doped but very thin such that for a given potential difference v_2, the electric field is large

enough to sever covalent bonds in the depletion region and generate pairs of free electrons and holes, which are then swept away by the electric field, thus creating a current. Avalanche breakdown occurs when the potential difference v_D is large enough that the kinetic energy of minority charge carriers is sufficient to break covalent bonds during collisions. These collisions may liberate free electrons and holes, which, again, are swept away by the electric field. The process by which energy is imparted to new charge carriers is called *impact ionization*. These new charge carriers may also have enough energy to energize other low-energy electrons, such that a sufficiently large reverse-biased diode voltage may initiate an avalanche of liberated charge carriers.

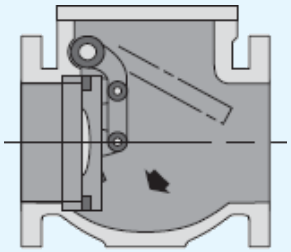
In **Zener breakdown** the high concentration of charge carriers provides the means for a substantial reverse-biased current to be sustained, at a nearly constant Page 527reverse-biased voltage, the **Zener voltage** V_Z . This effect is very useful in applications where one would like to regulate (hold constant) the voltage across a load. It should also be noted that a typical silicon diode is not designed for use in reverse breakdown, where even a modest current at a large V_Z will likely generate more power than the diode can dissipate through heat transfer. The result could be a melted or burned diode!

MAKE THE CONNECTION

Hydraulic Check Valves

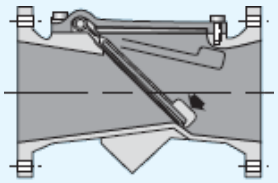
The operation of a diode can be understood intuitively by reference to a very common hydraulic device that finds application whenever one wishes to restrict the flow of a fluid to a single direction and to prevent (check) reverse flow. Hydraulic check valves perform this task in a number of ways. A few examples are illustrated here.

The first figure below depicts a swing check valve. In this design, flow from left to right is permitted, as the greater fluid pressure on the left side of the valve forces the swing door to open. If flow were to reverse, the reversal of fluid pressure (greater pressure on the right) would cause the swing door to shut.



Swing check valve

The second figure below depicts a flapper check valve. The principle is similar to that described above for the swing check valve in that fluid flow is permitted from left to right, but not in the reverse direction. The response of the flapper checkvalve is faster than the swing check valve due to the shorter travel distance of the flapper.



Flapper check valve

Diode circuits are much easier to understand when the behavior of the diode is visualized to be similar to that of a check valve, with the pressure difference across the valve orifice being analogous to the voltage across the diode and the fluid flow rate being analogous to the current through the diode. Charge flows only when the voltage across the diode is positive or *forward-biased*, and no charge flows when the diode voltage is negative or *reverse-biased.*

8.3 LARGE-SIGNAL MODELS FOR THE SEMICONDUCTOR DIODE

From the viewpoint of a *user* of electronic circuits (as opposed to a *designer*), it is often sufficient to characterize a device in terms of its *i-v* characteristic, using either load-line analysis or appropriate circuit models to determine the operating currents and voltages. This section shows how it is possible to use the *i-v* characteristics of the semiconductor diode to construct simple yet useful *circuit models.* Depending on the desired level of detail, it is possible to construct *large-signal models* of the diode, which describe the gross behavior of the device in the presence of relatively large

voltages and currents; or *small-signal models,* which are capable of describing the behavior of the diode in finer detail and, in particular, the response of the diode to small changes in the average diode voltage and current. From the user's standpoint, these circuit models greatly simplify the analysis of diode circuits and make it possible to effectively analyze relatively "difficult" circuits simply by using the familiar circuit analysis tools of Chapter 2. The first two major divisions of this section describe different diode models and the assumptions under which they are obtained, to provide the knowledge you will need to select and use the appropriate model for a given application.

Ideal Diode Model

The simplest large-signal diode model is the **ideal diode**, which approximates a diode as a simple on/off device (like a check valve in hydraulic circuits). The circuit symbol for an ideal diode, its *i-v* approximation, and the *i-v* characteristic of a typical diode are shown in Figure 8.11. An ideal diode behaves as an open-circuit when reverse-biased ($v_D < 0$) and as a short-circuit when forward-biased Page 528($v_D \leq 0$). Due to its simplicity, the ideal diode model can be very useful in circuit analysis.

Ideal diodes are represented by the solid black triangle symbol shown in Figure 8.11.

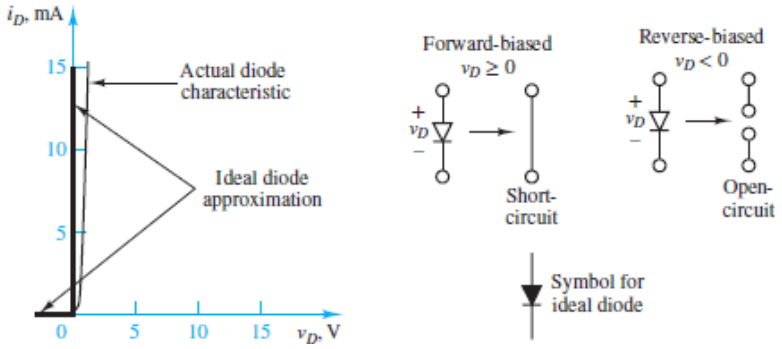


Figure 8.11 Large-signal on/off ideal diode model

A general method for analyzing diode circuits is illustrated using the circuit shown in Figure 8.12, which contains a 1.5-V battery, an ideal diode, and a 1-k Ω resistor. The method is simply to assume that the ideal diode is forward-biased ($v_D \leq 0$) and thus equivalent to a short-circuit, as indicated in Figure 8.13. Under this

assumption, $v_D = 0$ such that the loop current is $i_D = 1.5$ V/1 k Ω = 1.5 mA. Since the resulting direction of the current and the diode voltage are consistent with the assumption of a conducting diode ($v_D \leq 0$, $i_D > 0$), the assumption is correct. If the assumption had resulted in diode current and voltage that contradict the assumption, then the assumption would have been deemed incorrect, and the opposite assumption of a nonconducting diode could be tested, and presumably found to be true.

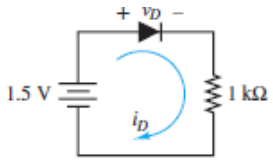


Figure 8.12 Circuit containing ideal diode

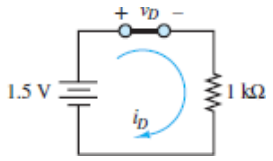


Figure 8.13 Circuit of Figure 8.12, assuming that the ideal diode conducts

To test the opposite assumption, assume the ideal diode is reverse-biased ($v_D < 0$) and thus equivalent to an open-circuit, as shown in Figure 8.14. Since the loop does not form a closed path, the current i_D must be zero and thus Ohm's law requires the voltage across the resistor to also be zero. Then, KVL requires that $v_D = 1.5$ V. However, this result contradicts the assumption that the ideal diode is reverse-biased. Thus, the assumption is deemed incorrect.

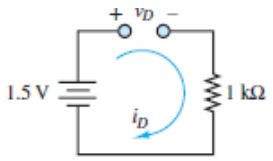


Figure 8.14 Circuit of Figure 8.12, assuming that the ideal diode does not conduct

The method can be applied to more complicated circuits involving multiple diodes by simply testing all the possible combinations of forward- and reverse-biased assumptions for the diodes. In such cases, it is helpful to consider which Page 529combinations are more likely to yield a correct solution and test those combinations first. With practice, such educated guesses should become more and more effective in reducing the number of tests necessary for any particular problem.

It is only necessary to find one set of assumptions that does not result in a contradiction.

Offset Diode Model

While the ideal diode model is useful in approximating the large-scale characteristics of a physical diode, it does not account for the diode offset voltage. A better model is the **offset diode model**, which consists of a battery when the diode if forward-biased, as shown in Figure 8.15, where the battery voltage equals the offset voltage (for silicon diodes $V_\gamma \approx 0.6$ V). The effect of the battery is to shift the forward-biased characteristic of the ideal diode to the right on the voltage axis, as shown in Figure 8.16.

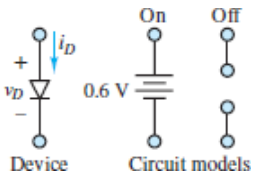


Figure 8.15 Offset diode model forward- and reverse-biased states

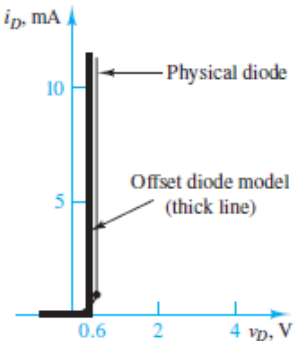


Figure 8.16 Offset diode model *i-v* characteristic

The behavior of a diode in the offset diode model is described as follows:

$$\begin{array}{ll} v_D \geq V_\gamma & \text{Diode} \rightarrow \text{0.6-V battery} \\ v_D < V_\gamma & \text{Diode} \rightarrow \text{open-circuit} \end{array} \qquad \text{Offset diode model} \tag{8.9}$$

LO

FOCUS ON PROBLEM SOLVING

DETERMINING THE CONDUCTION STATE OF IDEAL DIODES

1. Assume a diode conduction state (forward- or reverse-biased) for each diode
2. Replace each diode with an ideal diode (short-circuit if forward-biased, o circuit if reverse-biased).
3. Solve for the diode currents and voltages, using linear circuit analysis.
4. If the entire solution is consistent with the assumptions, then the ir assumptions were correct; if not, at least one of the initial diode conduction assumptions is wrong. Change at least one of the assumed diode conduc states, and solve the new circuit. Continue to iterate this process until a solu is found that is consistent with the assumptions.

LO

EXAMPLE 8.1 Determining the Conduction State of an Ideal Diode

Problem

Determine whether the ideal diode of Figure 8.17 is conducting.

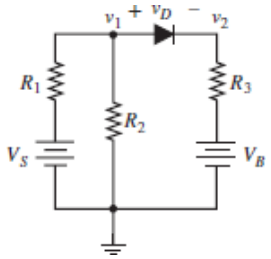


Figure 8.17

Solution

Known Quantities: $V_S = 12$ V; $V_B = 11$ V; $R_1 = 5\ \Omega$; $R_2 = 10\ \Omega$; $R_3 = 10\ \Omega$.

Find: The conduction state of the diode.

Assumptions: Use the ideal diode model.

Analysis: Assume initially that the ideal diode does not conduct, and replace it with an open-circuit, as shown in Figure 8.18. The voltage across R_2 can then be computed by using the voltage divider rule:

$$v_1 = \frac{R_2}{R_1 + R_2} V_S = \frac{10}{5 + 10} 12 = 8 \text{ V}$$

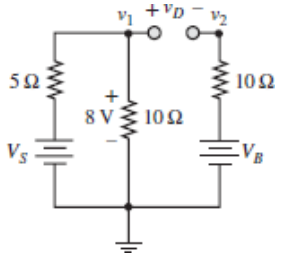


Figure 8.18

Apply KVL to the right-hand-side mesh, in which the current is zero, to obtain

$$v_1 = v_D + V_B \quad \text{or} \quad v_D = 8 - 11 = -3 \text{ V}$$

The result indicates that the diode is reverse-biased and confirms the initial assumption. Thus, the diode is not conducting.

As further illustration, assume that the diode conducts. In this case, the diode is replaced with a short-circuit, as shown in Figure 8.19. The resulting circuit can be solved by node analysis, noting that $v_1 = v_2$ because of the short-circuit.

$$\frac{V_S - v_1}{R_1} = \frac{v_1 - 0}{R_2} + \frac{v_2 - V_B}{R_3}$$

$$\frac{V_S}{R_1} + \frac{V_B}{R_3} = \frac{v_1}{R_1} + \frac{v_1}{R_2} + \frac{v_2}{R_3}$$

$$\frac{12}{5} + \frac{11}{10} = \left(\frac{1}{5} + \frac{1}{10} + \frac{1}{10}\right) v_1$$

$$v_1 = 2.5(2.4 + 1.1) = 8.75 \text{ V}$$

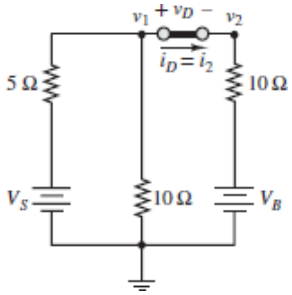


Figure 8.19

With $v_1 = v_2 = 8.75$ V, the current through the diode is

$$i_D = \frac{v_1 - V_B}{R_3} = \frac{8.75 - 11}{10} = -0.225 \text{ A}$$

However, this negative current violates the forward-biased assumption about the diode. Thus, the forward-biased conducting assumption is incorrect.

EXAMPLE 8.2 Determining the Conduction State of an Ideal Diode

Problem

Determine whether the ideal diode of Figure 8.20 is conducting.

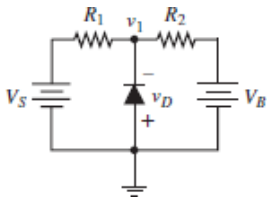


Figure 8.20

Solution

Known Quantities: $V_S = 12$ V; $V_B = 11$ V; $R_1 = 5\ \Omega$; $R_2 = 4\ \Omega$.

Find: The conduction state of the diode.

Assumptions: Use the ideal diode model.

Analysis: Assume initially that the ideal diode does not conduct, and replace it with an open-circuit, as shown in Figure 8.21. The current through the resulting series loop is

$$i = \frac{V_S - V_B}{R_1 + R_2} = \frac{1}{9} \text{ A}$$

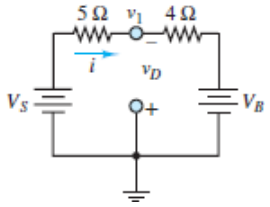


Figure 8.21

The voltage at node v_1 is

$$\frac{12 - v_1}{5} = \frac{v_1 - 11}{4}$$

$$v_1 = 11.44 \text{ V}$$

The result indicates that the diode is strongly reverse-biased, since $v_D = 0 - v_1 = -11.44$ V, and is in accord with the initial assumption. Thus, the diode is not conducting. Note that $v_1 = V_B + 4/9\ (V_S - V_B)$ as expected due to voltage division applied to two resistors in series.

EXAMPLE 8.3 Using the Offset Diode Model

Problem

Use the offset diode model to determine the value of v_1 for which diode D_1 first conducts in the circuit of Figure 8.22.

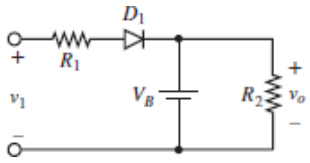


Figure 8.22

Solution

Known Quantities: $V_B = 2$ V; $R_1 = 1$ k Ω; $R_2 = 500$ Ω; $V_\gamma = 0.6$ V.

Find: The lowest value of v_1 for which diode D_1 conducts.

Assumptions: Use the offset diode model.

Analysis: Start by replacing the diode with the offset diode model, as shown in Figure 8.23. If v_1 is negative, the diode will certainly be off. The point at which the diode turns on as v_1 is increased can be determined by analyzing the circuit with the diode assumed to be on. KVL applied around the left mesh yields:

$$v_1 = v_{D1} + 0.6 + 2 \qquad \text{or} \qquad v_{D1} = v_1 - 2.6$$

Thus, the condition required for the diode to conduct is

$$v_1 \geq 2.6 \text{ V} \qquad \text{Diode "on" condition}$$

Comments: The same solution method can be applied using the ideal diode model.

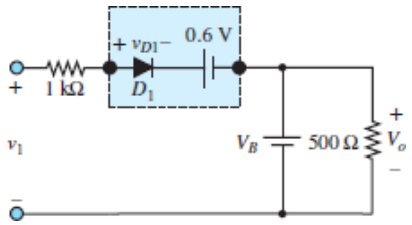


Figure 8.23

Page 532

CHECK YOUR UNDERSTANDING

If the resistor R_2 is replaced with an open-circuit in the circuit of Figure 8.17, will the ideal diode conduct? If the offset diode model is used will the diode conduct?

Answer: Yes, to both questions.

CHECK YOUR UNDERSTANDING

Repeat the analysis of Example 8.2, assuming that the diode is conducting, and show that this assumption leads to inconsistent results.

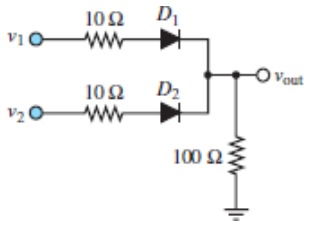
Determine which of the diodes conduct in the circuit shown below for each of the following voltages. Treat the diodes as ideal.

a. $v_1 = 0$ V; $v_2 = 0$ V

b. $v_1 = 5$ V; $v_2 = 5$ V

c. $v_1 = 0$ V; $v_2 = 5$ V

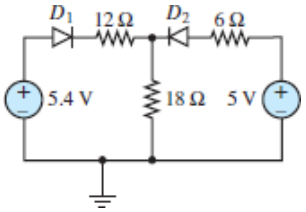
d. $v_1 = 5$ V; $v_2 = 0$ V



Answer: (a) Neither; (b) both; (c) D_2 only; (d) D_1 only

CHECK YOUR UNDERSTANDING

Determine which of the diodes conduct in the circuit shown below. Each diode has an offset voltage of 0.6 V.



Answer: Both diodes conduct.

8.4 SMALL-SIGNAL MODELS FOR THE SEMICONDUCTOR DIODE

As one examines the diode *i-v* characteristic more closely, it becomes apparent that the short-circuit approximation is not adequate to represent the *small-signal behavior* of the diode. The term *small-signal behavior* usually signifies the response of the diode to small time-varying signals that may be superimposed on the average diode current and voltage. Figure 8.8 provides a more detailed view of a silicon diode *i-v* curve. Clearly, the short-circuit approximation is not very accurate when a diode's behavior Page 533is viewed on a finer scale. To a first-order approximation, however, the *i-v* characteristic is linear for voltages greater than the offset voltage. Thus, it seems reasonable to model a conducting diode as a resistor with an offset voltage. Load-line analysis can be exploited to determine the diode **small-signal resistance**, which is related to the slope of its *i-v* characteristic.

Consider the circuit of Figure 8.24, which represents the Thévenin equivalent circuit of an arbitrary linear resistive circuit connected to a diode. KVL yields the *governing equation:*

$$V_S = i_D R_S + v_D \tag{8.10}$$

The *constitutive relation* for the diode is

$$i_D = I_0(e^{v_D/V_T} - 1) \tag{8.11}$$

These two equations in two unknowns cannot be solved analytically since one of the equations is *transcendental;* that is, it contains the unknown v_D in exponential form. Transcendental equations of this type can be solved graphically or numerically. Only a graphical solution is considered here.

Consider a plot of the two preceding equations in the i_D–v_D plane. The diode equation gives rise to the familiar curve of Figure 8.8. The *load-line equation,* obtained from KVL, is the equation of a line with slope $-1/R_S$, open-circuit voltage V_S, and short-circuit current V_S/R_S.

$$i_D = \frac{V_S - v_D}{R_S} = -\frac{1}{R_S}v_D + \frac{V_S}{R_S} \qquad \text{Load-line equation} \tag{8.12}$$

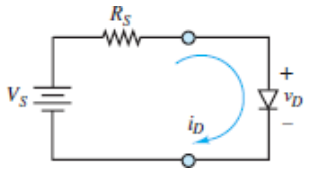


Figure 8.24 Diode circuit used to illustrate load-line analysis

The superposition of these two curves gives rise to the plot of Figure 8.25, where the solution to the two equations is found graphically to be the pair of values (I_Q, V_Q). The intersection of the two curves is called the **quiescent (operating) point**, or ***Q* point**. The voltage $v_D = V_Q$ and the current $i_D = I_Q$ are the actual diode voltage and current when the diode is connected as in the circuit of Figure 8.24. Note that this method is also useful for circuits containing a larger number of elements, where the diode is treated as the load and Thévenin's theorem is used to simplify the remaining (assumed linear) source network.

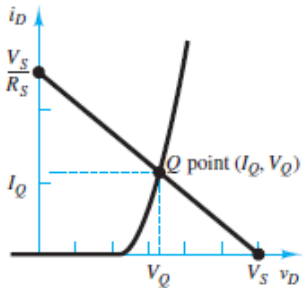


Figure 8.25 Graphical solution of equations 8.13 and 8.14

Piecewise Linear Diode Model

The graphical solution of diode circuits can be somewhat tedious, and its accuracy is limited by the resolution of the graph. However, it does provide insight into the **piecewise linear diode model** in which the diode is treated as an open-circuit in the "off" state and as a linear resistor in series with an offset voltage V_γ in the "on" state. Figure 8.26 provides a graphical illustration of this model. The straight line that approximates the on state of the diode is chosen to be tangent to the operating point Q. Thus, in the neighborhood of the Q point, the diode in this model acts as a linear resistor with slope given by $1/r_D$, where:

$$\frac{1}{r_D} = \left.\frac{\partial i_D}{\partial v_D}\right|_{(I_Q, V_Q)} \qquad \text{Diode incremental resistance} \qquad (8.13)$$

In this context, the diode offset voltage is defined as the intersection of the tangent line at Q with the voltage axis. Thus, rather than represent the diode as a Page

534short-circuit in its forward-biased state, it is treated as a linear resistor r_D to account for changes in i_D due to changes in v_D. The piecewise linear model offers the convenience of a linear representation once the state of the diode is established, and of a more accurate model than either the ideal or the offset diode model. This model is very useful in representing the performance of diodes in circuits where the diode voltage is varying about the DC operating point Q.

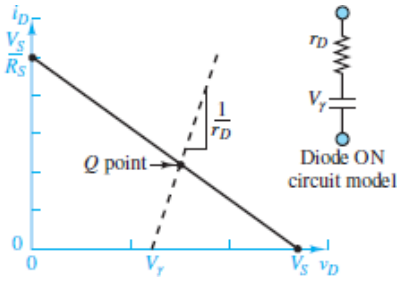


Figure 8.26 Piecewise linear diode model $v_D = V_\gamma + i_D r_D \quad (v_D \geq V_\gamma)$

LO

FOCUS ON PROBLEM SOLVING

DETERMINING THE OPERATING POINT OF A DIODE

1. With the diode as the load, apply Thévenin's or Norton's theorem to simplify (assumed linear) network seen by the diode.
2. Use the result of the simplification to determine the load line (equation 8.12)
3. Use an iterative numerical method to solve the two simultaneous equation two unknowns (the load-line equations and the diode equation) for the d current and voltage.

or

4. Use a graphical method to find the intersection of the diode curve (e.g., fro data sheet) with the load-line curve. The intersection of the two curves is diode operating point Q.

EXAMPLE 8.4 Using Load-Line Analysis and Diode Curves to Determine the Operating Point of a Diode

Problem

Determine the operating point of the 1N914 diode in the circuit of Figure 8.27, and compute the total power output of the 12-V battery.

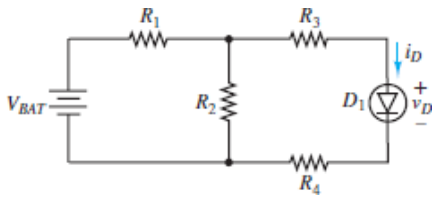


Figure 8.27

Page 535

Solution

Known Quantities: $V_{BAT} = 12$ V; $R_1 = 50\ \Omega$; $R_2 = 10\ \Omega$; $R_3 = 20\ \Omega$; $R_4 = 20\ \Omega$.

Find: The diode operating voltage and current and the power supplied by the battery.

Assumptions: Use the diode nonlinear model, as described by its *i-v* curve (Figure 8.28).

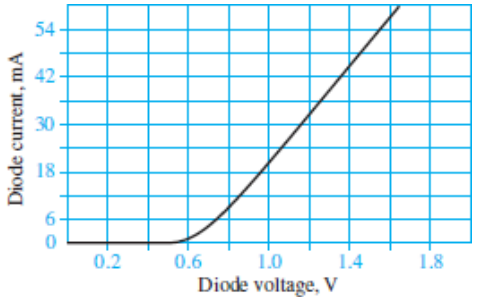


Figure 8.28 The 1N914 diode *i-v* curve

Analysis: Consider the diode in Figure 8.27 to be the load and everything else attached to it as its source network. Replace the source network with its Thévenin equivalent (Figure 8.29) and determine the load line as shown in Figure 8.30. The

Thévenin equivalent resistance and the Thévenin (open-circuit) voltage seen by the diode are

$$R_S = R_3 + R_4 + (R_1 \| R_2) = 20 + 20 + (10 \| 50) = 48.33\ \Omega$$

$$V_S = \frac{R_2}{R_1 + R_2} V_{BAT} = \frac{10}{60} 12 = 2\ \text{V}$$

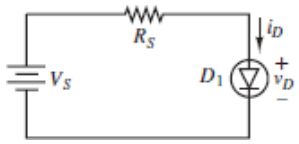


Figure 8.29

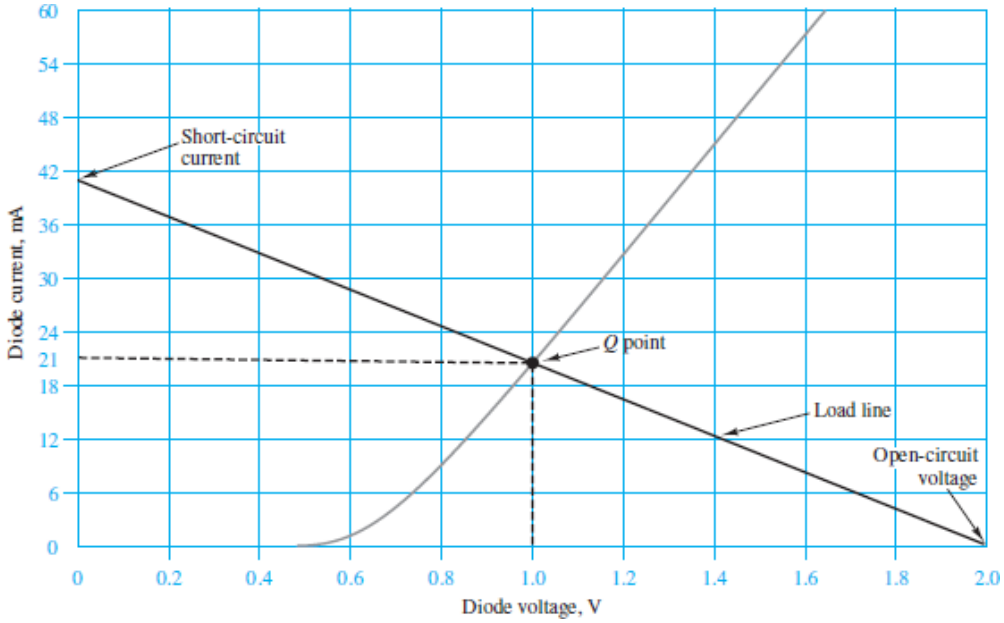


Figure 8.30 Superposition of load line and diode *i-v* characteristic

Page 536

The short-circuit current is $V_S/R_S = 41$ mA. The intersection of the diode curve and the load line is the *quiescent* or *operating point Q* of the diode, which is given by the values $V_Q = 1.0$ V and $I_Q = 21$ mA.

To determine the battery power output, observe that the power supplied by the battery is $P_B = 12 \times I_B$ and that I_B is equal to the current through R_1. Upon further inspection, the battery current must, by KCL, be equal to the sum of the currents through R_2 and the diode. The current through the diode is I_Q . To determine the

current through R_2 , observe that the voltage across R_2 is equal to the sum of the voltages across R_3 , R_4 , and D_1:

$$V_{R2} = I_Q(R_3 + R_4) + V_Q = 0.021 \times 40 + 1 = 1.84 \text{ V}$$

and therefore the current through R_2 is $I_{R2} = V_{R2} / R_2 = 0.184$ A.

Finally:

$$P_B = 12 \times I_B = 12 \times (0.021 + 0.184) = 12 \times 0.205 = 2.46 \text{ W}$$

Comments: Graphical solutions are not the only means of solving the nonlinear equations that result from using a nonlinear diode model. The same equations could be solved numerically by using an iterative nonlinear equation solver.

EXAMPLE 8.5 Computing the Incremental (Small-Signal) Resistance of a Diode

Problem

Determine the incremental resistance of a diode, using the diode equation.

Solution

Known Quantities: $I_0 = 10^{-14}$ A; $V_T = 25$ mV (at $T = 300$ K); $I_Q = 50$ mA.

Find: The diode incremental (small-signal) resistance r_D .

Assumptions: Use the approximate diode equation (equation 8.8).

Analysis: The approximate diode equation is

$$i_D = I_0 e^{v_D/V_T}$$

This expression can be used along with equation 8.13 to compute the incremental resistance:

$$\frac{1}{r_D} = \left.\frac{\partial i_D}{\partial v_D}\right|_{(I_Q, V_Q)} = \frac{I_0}{V_T} e^{V_Q/V_T} = \frac{qI_0}{kT} e^{V_Q/V_T}$$

To calculate the numerical value of the above expression, first compute the quiescent diode voltage corresponding to the quiescent current $I_Q = 50$ mA:

$$V_Q = V_T \log_e \frac{I_Q}{I_0} = \frac{kT}{q} \log_e \frac{I_Q}{I_0} = 0.731 \text{ V}$$

Substitute the numerical value of V_Q in the expression for r_D to obtain:

$$\frac{1}{r_D} = \frac{10^{-14}}{0.025} e^{0.731/0.025} = 2 \text{ S} \qquad \text{or} \qquad r_D = 0.5 \ \Omega$$

Comments: The incremental resistance of a diode at an operating point can be computed for any particular circuit. However, in general, a diode cannot be treated simply as a resistor. Page 537The incremental (small-signal) resistance of the diode is used in the piecewise linear diode model to account for the fact that there is a dependence between diode voltage and current. The incremental resistance will change if the operating point changes since it is, after all, the slope of the *i-v* characteristic at the operating point.

EXAMPLE 8.6 Using the Piecewise Linear Diode Model

Problem

Determine the load voltage v_o in the *rectifier* circuit of Figure 8.31, using a piecewise linear approximation.

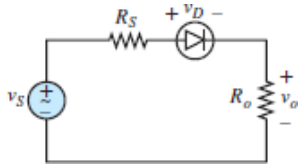


Figure 8.31

Solution

Known Quantities: $v_S(t) = 10 \cos \omega t$; $V_\gamma = 0.6$ V; $r_D = 0.5 \ \Omega$; $R_S = 1 \ \Omega$; $R_o = 10 \ \Omega$.

Find: The load voltage v_o .

Assumptions: Use the piecewise linear diode model (Figure 8.26).

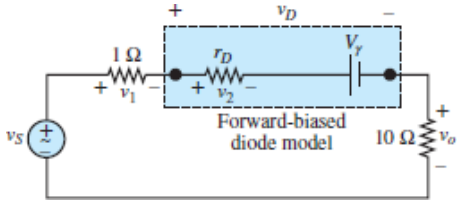


Figure 8.32 Piecewise linear model of forward-biased diode inserted in circuit of Figure 8.31. $v_D = V_\gamma + i_D r_D$

Analysis: Apply KVL to the circuit in Figure 8.32 to determine the requirement for diode conduction. That result and the complementary result for when the diode is not conducting are

$$v_S = v_1 + v_D + v_o = v_1 + v_2 + V_\gamma + v_o \qquad \text{Conducting forward-biased diode}$$

$$v_S = v_D \qquad \text{Nonconducting diode}$$

Observe that when v_S is negative, the diode will be off; it will act as an open-circuit; the voltages v_1 , v_2 , and v_2 will be zero; and $v_D = v_S$. At the onset of conduction the diode is forward-biased but the diode current is still zero. Under this condition v_1 , v_2, and v_o are zero (Ohm's law) such that $v_D = v_S = V_\gamma = 0.6$ V. Thus, the condition for conduction is:

$$v_D = v_S = V_\gamma \approx 0.6 \text{ V} \qquad \text{Onset of conduction}$$

Once the diode conducts, the difference between v_S and V_γ is divided among the three series resistors according to voltage division. Thus:

$$v_o = \begin{cases} 0 & v_S < V_\gamma = 0.6 \text{ V} \\ \dfrac{R_o}{R_S + r_D + R_o}(v_S - V_\gamma) = 8.7 \cos \omega t - 0.52 & v_S \geq 0.6 \text{ V} \end{cases}$$

Page 538

The source and load voltage are plotted in Figure 8.33(a). The *transfer characteristic* of the circuit is shown as a plot of v_o versus v_S in Figure 8.33(b). Notice the offset voltage.

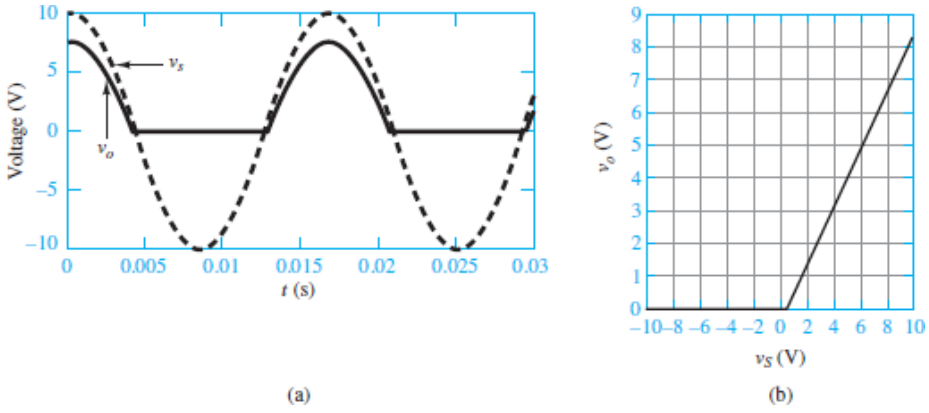
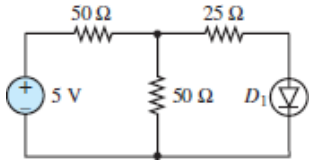


Figure 8.33 (a) Source voltage and rectified load voltage; (b) voltage transfer characteristic

CHECK YOUR UNDERSTANDING

Use load-line analysis to determine the operating point Q of the diode circuit shown below. The diode has the i-v characteristic shown in Figure 8.30. Treat the diode as the load and graph the load line using the short-circuit current V_{TH}/R_{TH} as the ordinate intercept and $-1/R_{TH}$ as the slope of the load line.



Answer: $V_Q = 1.11$ V, $I_Q = 27.7$ mA

CHECK YOUR UNDERSTANDING

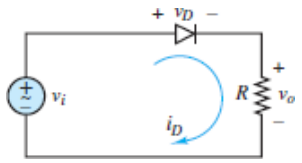
Compute the incremental resistance of the diode of Example 8.5 if the current through the diode is 250 mA.

Answer: $r_D = 0.1\ \Omega$

Page 539

CHECK YOUR UNDERSTANDING

Consider the *half-wave rectifier* circuit shown below where $v_i = 18\cos(t)$ V and $R = 4\ \Omega$. Sketch the output voltage waveform if the piecewise linear diode model is used to represent the diode, with $V_\gamma = 0.6$ V and $r_D = 1\ \Omega$. What is the peak value of the output waveform?



Answer: $v_{o,\,peak} = 13.92$ V

8.5 RECTIFIER CIRCUITS

The need for converting one form of electric energy to another arises frequently in practice. The most readily available form of electric power is alternating current, as generated and delivered by electric power utilities. However, DC power is frequently required for applications ranging from the control of electric motors to the operation of consumer electronic circuits, such as tablet computers and smartphones. An important part of the process of converting an AC signal to direct current is *rectification,* which is the process of converting an electrical signal so that all of its parts have the same sign. Of particular interest is the process of converting an AC signal (e.g., a typical 120-V rms line voltage) with zero average (DC) value to a signal with a nonzero DC value. For example, power supplies use rectification to produce a DC output from the readily available AC line voltage. The basic principle of rectification is well illustrated using ideal diodes, particularly when the magnitude of the AC voltage is large compared to the diode offset voltage V_γ.

This section introduces the following three types of rectifier circuits:

- Half-wave rectifier
- Full-wave rectifier
- Bridge rectifier

The Half-Wave Rectifier

Consider the circuit of Figure 8.34, where an AC source v_i is connected to an ideal diode and a resistive load in a series loop. The diode will conduct only when it is forward-biased ($v_D \leq 0$), which occurs during the positive half-cycle of the sinusoidal voltage. During that interval the ideal diode acts as a short-circuit such that $v_o = v_i$ and $i_D = v_i/R$. During the negative half-cycle of the sinusoid the ideal diode is reverse-biased ($v_D < 0$) and acts as an open-circuit. The loop current i_D is then zero, and, by Ohm's law, the output voltage v_o is also zero. The input voltage v_i and the resulting output voltage v_o are shown in Figure 8.35, where the frequency Page 540is assumed to be $\omega = 2\,\pi f = 2\,\pi$ (60 Hz). Notice that although the input voltage has a zero average (DC) value, the rectified output voltage v_2 has a nonzero average (DC) value, which is computed, in general, as:

$$(v_o)_{avg} = \frac{1}{T}\int_0^T v_o(t)\,dt = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} v_o(t)\,dt \tag{8.14}$$

where T is the period of the output waveform. For example, assume $v_i = 120\sqrt{2}\sin(\omega t)$ V. Then:

$$(v_o)_{avg} = \frac{\omega}{2\pi}\left[\int_0^{\pi/\omega} 120\sqrt{2}\sin(\omega t)\,dt + \int_{\pi/\omega}^{2\pi/\omega} 0\,dt\right] \tag{8.15}$$

$$= \frac{120\sqrt{2}}{\pi} = 54.0 \text{ V}$$

The circuit of Figure 8.34 is known as a **half-wave rectifier**, because only the positive half of the input waveform appears across the output. This result is not particularly satisfying nor efficient since half of the input waveform is lost.

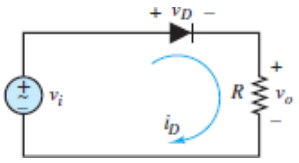


Figure 8.34 Ideal diode acting as a half-wave rectifier

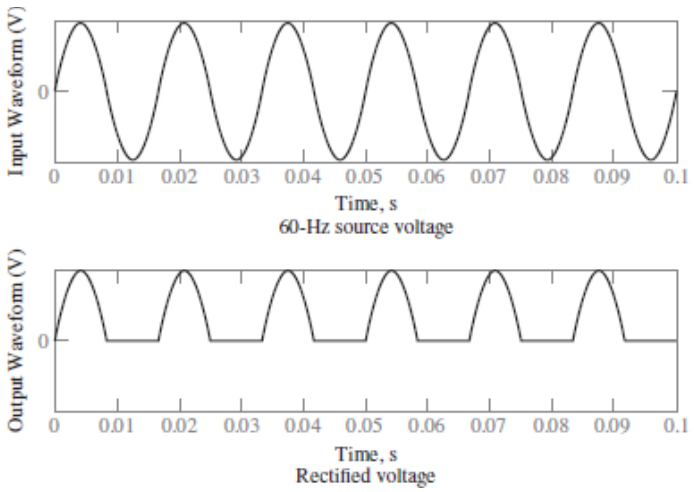


Figure 8.35 Ideal diode half-wave rectifier input and output

The Full-Wave Rectifier

The **full-wave rectifier** shown in Figure 8.36 includes an AC source and a center-tapped transformer with 1:2N turns ratio and offers a substantial improvement in performance over the half-wave rectifier. The purpose of the transformer is to step up ($N > 1$) or step down ($N < 1$) the primary voltage v_S prior to rectification. The voltage amplitude across each half of the secondary coil is Nv_S. In addition to scaling the source voltage, the transformer isolates the rectifier circuit from the AC source voltage since there is no direct electrical connection between the input and output of a transformer.

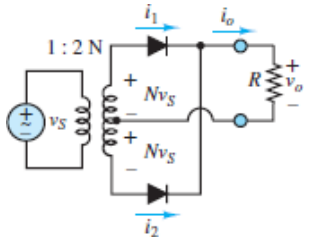


Figure 8.36 Center-tapped AC transformer and a full-wave rectifier with two ideal diodes.

In most applications, the amplitude of the secondary voltage (the input voltage to the rectifier) is much larger than the offset voltage of the diodes. Page 541When this condition is true, the diodes can be approximated as ideal without significantly compromising the result of the analysis. The key to the operation of the full-wave rectifier is to note that as the sign of v_S periodically alternates between positive and negative, the two diodes alternate in turns between forward- and reverse-biased

states. For instance, during the positive half-cycle of v_S, the top diode is forward-biased while the bottom diode is reverse-biased. Alternately, during the negative half-cycle of v_S, the top diode is reverse-biased while the bottom diode is forward-biased. Therefore, the output current i_o satisfies the following two relations:

$$i_o = i_1 = N\frac{v_S}{R} \qquad v_S \geq 0 \tag{8.16}$$

$$i_o = i_2 = -N\frac{v_S}{R} \qquad v_S \leq 0 \tag{8.17}$$

Wow! The direction of i_o does not alternate! It is always positive as shown.

The source voltage, the output voltage, and the currents i_1 and i_2 are shown in Figure 8.37 for a load resistance $R = 1\ \Omega$ and $N = 1$. Notice that the output voltage is exactly the superposition of the output of two half-wave rectifiers 180° out of phase. Thus, the DC output of the full-wave rectifier should be twice that of the half-wave rectifier. This observation can be confirmed by computing the DC value of the full-wave rectifier output.

$$\begin{aligned}(v_o)_{\text{avg}} &= \frac{1}{T}\int_0^T v_o(t)\,dt = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} v_o(t)\,dt \\ &= \frac{\omega}{2\pi}\left[\int_0^{\pi/\omega} v_o(t)\,dt + \int_{\pi/\omega}^{2\pi/\omega} v_o(t)\,dt\right] \\ &= 2\frac{\omega}{2\pi}\left[\int_0^{\pi/\omega} v_o(t)\,dt\right]\end{aligned} \tag{8.18}$$

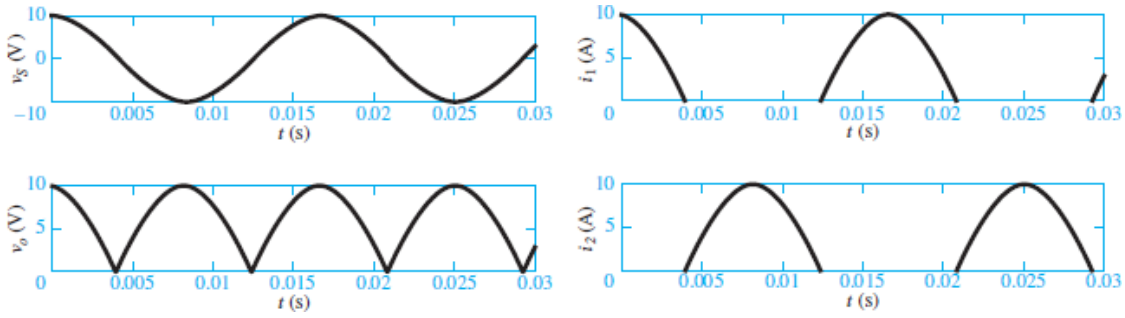


Figure 8.37 Full-wave rectifier current and voltage waveforms ($R = 1\ \Omega$)

Keep in mind that this result is approximate because the impact of the diode offset voltage was ignored by assuming ideal diodes. When the offset voltage is included, there will be periods (typically brief) when both diodes are reverse-biased and the output voltage is zero. The net effect is to reduce the output waveform shown in Figure 8.37 by V_γ . However, those portions of the adjusted waveform that would

otherwise be negative (between 0 and $-V_\gamma$) are, in fact, zero because both diodes are reverse-biased for the brief periods when $-V_\gamma < v_S < V_\gamma$.

Page 542

The Bridge Rectifier

Another rectifier circuit commonly available "off the shelf" as a single *integrated circuit* is the *bridge rectifier,* which employs four diodes in the bridge configuration shown in Figure 8.38.

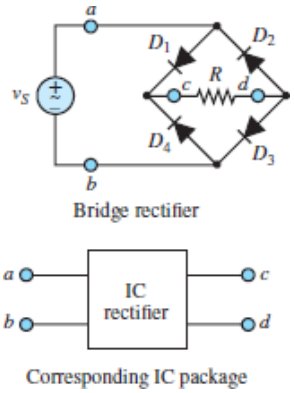
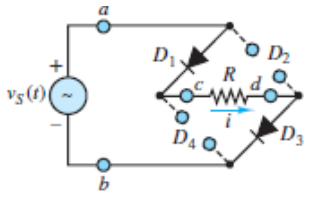
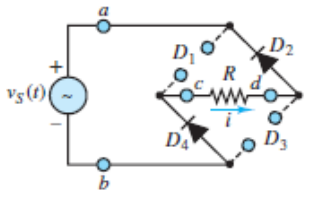


Figure 8.38 Bridge rectifier with four ideal diodes

The analysis of the bridge rectifier is best understood by observing that as the sign of v_S periodically alternates between positive and negative, *pairs* of the four bridge diodes alternate in turns between forward- and reverse-biased states, as shown in Figure 8.39. During the positive half-cycle of v_S, diodes D_1 and D_3 are forward-biased while diodes D_2 and D_4 are reverse-biased. Alternately, during the negative half-cycle of v_S, diodes D_1 and D_3 are reverse-biased while diodes D_2 and D_4 are forward-biased. It is important to note that the current i through R is directed from node c to node d during both half-cycles.



During the positive half-cycle of $v_S(t)$, D_1 and D_3 are forward-biased.



During the negative half-cycle of $v_S(t)$, D_2 and D_4 are forward-biased.

Figure 8.39 Operation of bridge rectifier

The input and rectified output waveforms are shown in Figure 8.40(a) and (b) for the case of ideal diodes and a 30-V peak AC source input. If each diode is assumed to have an offset voltage V_γ = 0.6 V, the effect is to reduce the output waveform by $2V_\gamma$ = 1.2 V, as shown in Figure 8.40(c). The $2V_\gamma$ reduction occurs during both half-cycles. During the positive half-cycle of v_γ, the path from node a to node b contains two forward-biased diodes D_1 and D_3. Alternately, during the negative half-cycle of v_S, the path from node b to node a also contains two forward-biased diodes D_2 and D_4. Each of these forward-biased diodes requires a "toll" of V_γ.

As with the full-wave rectifier, no portion of the rectified output waveform is negative even when reduced by $2V_\gamma$. Instead, during those periods when $-2V_\gamma < v_S < 2V_\gamma$, all four diodes are reverse-biased and the rectified output waveform is zero.

In most practical applications of rectifier circuits, the signal waveform to be rectified is the 60-Hz, 110 V rms line voltage. As shown in Figures 8.37 and 8.40, the fundamental frequency of the rectified output waveform is twice that of the Page 543input waveform. Thus, for a 60-Hz input waveform, the fundamental ripple frequency is 120 Hz or 754 rad/s. A low-pass filter is required such that:

$$\omega_0 \ll \omega_{\text{ripple}} \tag{8.19}$$

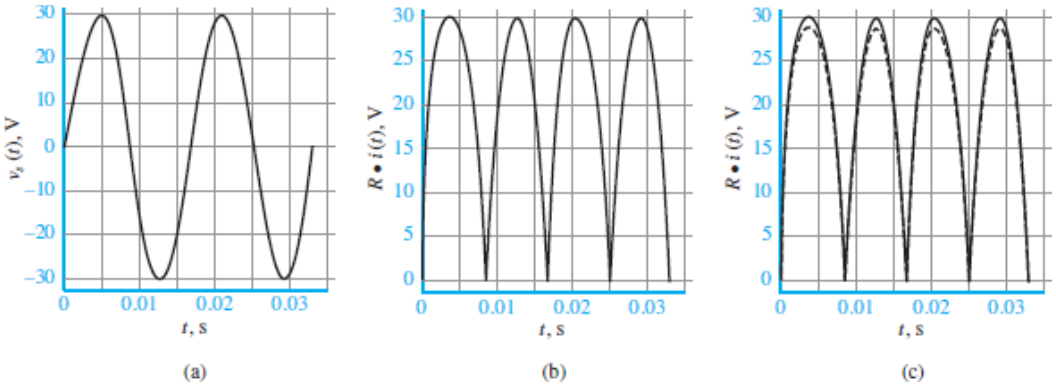


Figure 8.40 (a) Unrectified source voltage; (b) rectified load voltage (ideal diodes); (c) rectified load voltage (ideal and offset diodes)

Figure 8.41 shows the resulting waveforms.

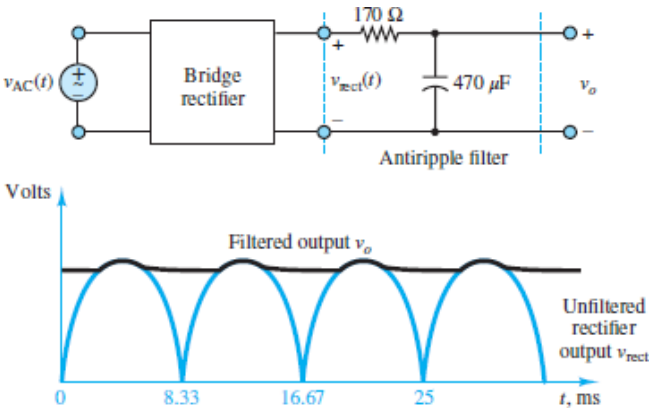


Figure 8.41 Bridge rectifier followed by a low-pass filter, and the resulting waveforms

DC Power Supplies

The rectification of an AC input waveform is just one of four fundamental steps needed to convert an AC input to a practical DC output. In a typical **DC power supply** these steps are, in order:

Step 1: Scale (step up or step down) the amplitude of the AC input waveform. This step is commonly accomplished by a transformer although high-frequency switch-mode circuits can also provide scaling of a DC output.

Step 2: Rectify the scaled AC input waveform. This step may be accomplished by a full-wave or bridge rectifier. Rectification can also be accomplished by more exotic devices, such as gate turnoff thyristors (GTOs) and insulated-gate bipolar transistors (IGBTs).

Step 3: Filter the rectified output waveform to remove remaining AC components known as *ripple*. This step can be accomplished by an *RC* low-pass (antiripple) filter in a simple DC power supply, as shown in Figure 8.41, or by more sophisticated active low-pass filters.

Step 4: Regulate the filtered DC output voltage to maintain the desired DC value for a large range of loads. The Zener diode provides a very inexpensive and simple form of voltage regulation. Linear voltage regulators, which have very good noise characteristics, and switched-mode regulators, which have very high energy efficiency, are available as integrated circuits (e.g., the 78xx linear series).

These steps are represented in the generic depiction of a DC power supply shown in Figure 8.42.

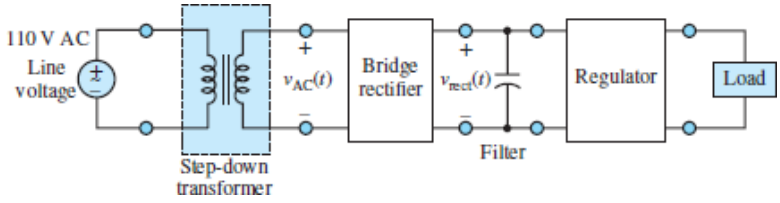


Figure 8.42 DC power supply

EXAMPLE 8.7 Using the Offset Diode Model in a Half-Wave Rectifier Problem

Compute and plot the rectified load voltage v_R in the circuit of Figure 8.43.

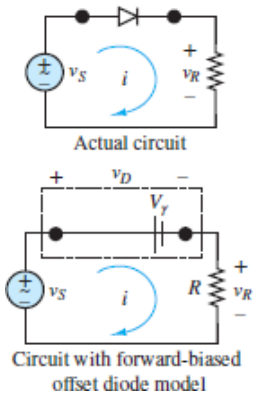


Figure 8.43

Solution

Known Quantities: $v_S(t) = 3 \cos \omega t$; $V_\gamma = 0.6$ V.

Find: An analytical expression for the load voltage.

Assumptions: Use the offset diode model.

Analysis: When the source voltage is greater than $V_\gamma = 0.6$ V, the diode is forward-biased such that it behaves as a short-circuit in series with a small offset voltage drop, as shown in Figure 8.43. The loop current i and the voltage v_R across R are given by:

$$i = \frac{v_S - V_\gamma}{R} \qquad v_R = iR = v_S - V_\gamma$$

Now assume that the diode is reverse-biased and replace it with an open-circuit, as shown in Figure 8.44. Since the current through R is zero, the diode voltage v_D is found from KVL to be:

$$v_D = v_S \qquad i = 0 \qquad \text{when } v_S < V_\gamma \qquad \text{Reverse-bias condition}$$

Page 545

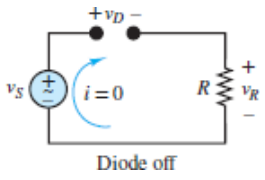


Figure 8.44

Thus, the half-wave rectifier circuit behavior is summarized by:

$$v_R = \begin{cases} 0 & v_S < V_\gamma \\ v_S - 0.6 & v_S \geq V_\gamma \end{cases}$$

The resulting rectified waveform $v_R(t)$ is plotted along with $v_S(t)$ in Figure 8.45. The effect of the offset voltage is to lower the positive portion of the rectified waveform by V_γ. The period $T+$ during which the rectified waveform is positive is slightly shorter than half the period T of the input waveform. The reason for this result is that the diode is not on unless $v_S \leq V_\gamma$. In this example, the onset of conduction occurs when $v_S = V_\gamma = 3\cos(\omega\Delta t)$, such that $T^+ = \mathrm{T}/2 - 2\Delta t$. For ideal diodes, the maximum amplitude of the rectified waveform equals the amplitude of the input waveform and $T^+ = T/2$.

Comments: The rectified waveform is shifted downward by an amount equal to the offset voltage V_γ. The shift is visible in the case of this example because V_γ is a substantial fraction of the source voltage. If the source voltage had peak values of tens or hundreds of volts, such a shift would be negligible, and an ideal diode model would serve just as well.

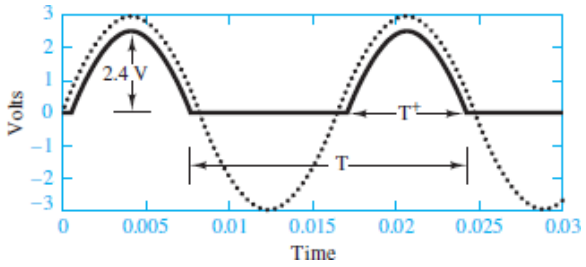


Figure 8.45 Source waveform (…) and rectified waveform (—) for the circuit of Figure 8.43

LO

EXAMPLE 8.8 Half-Wave Rectifiers

Problem

A half-wave rectifier, similar to that in Figure 8.34, is used to provide a DC supply to a 50-Ω load. If the AC source voltage is 20 V rms, find the peak and average current in the load. Assume an ideal diode.

Solution

Known Quantities: Value of circuit elements and source voltage.

Find: Peak and average values of load current in half-wave rectifier circuit.

Schematics, Diagrams, Circuits, and Given Data: $v_S = 20$ V rms, $R = 50\ \Omega$.

Assumptions: Ideal diode.

Analysis: For a forward-biased ideal diode, the peak load voltage is equal to the peak sinusoidal source voltage. Thus, the peak load current is

$$i_{\text{peak}} = \frac{v_{\text{peak}}}{R} = \frac{\sqrt{2}\, v_{\text{rms}}}{R} = 0.567 \text{ A}$$

Page 546

To compute the average current, integrate the half-wave rectified sinusoid:

$$\langle i \rangle = \frac{1}{T}\int_0^T i(t)\,dt = \frac{1}{T}\left[\int_0^{T/2} \frac{v_{\text{peak}}}{R}\sin(\omega t)\,dt + \int_{T/2}^{T} 0\,dt\right]$$

$$= \frac{v_{\text{peak}}}{\pi R} = \frac{\sqrt{2}\, v_{\text{rms}}}{\pi R} = 0.18 \text{ A}$$

EXAMPLE 8.9 Bridge Rectifier

Problem

A bridge rectifier, similar to that in Figure 8.38, is used to produce a 50-V, 5-A DC power supply. What is the smallest allowed value of the load R that will result in a 5-A DC output current? What is the required source voltage v_S (in V rms) to achieve the desired DC output voltage? Assume ideal diodes.

Solution

Known Quantities: Value of circuit elements and source voltage.

Find: Source voltage v_S (in V rms) and the load resistance R.

Schematics, Diagrams, Circuits, and Given Data: $\langle v_o \rangle = 50$ V; $\langle i_o \rangle = 5$ A.

Assumptions: Ideal diodes.

Analysis: The load resistance that will result in an average direct current of 5 A is:

$$R = \frac{\langle v_o \rangle}{\langle i_o \rangle} = \frac{50}{5} = 10\ \Omega$$

which is the lowest value of R for which the DC supply will be able to provide the required current. To compute the required source voltage, we observe that the average load voltage can be found from the expression

$$\langle v_o \rangle = R\langle i_o \rangle = \frac{R}{T}\int_0^T i(t)\,dt = \frac{R}{T}\left[\int_0^{T/2} \frac{v_{\text{peak}}}{R}\sin(\omega t)\,dt\right] \times 2$$
$$= \frac{2v_{\text{peak}}}{\pi} = \frac{2\sqrt{2}\,v_{\text{rms}}}{\pi} = 50\ \text{V}$$

Hence:

$$v_{\text{rms}} = \frac{50\pi}{2\sqrt{2}} = 55.5\ \text{V}$$

CHECK YOUR UNDERSTANDING

Compute the DC value of the rectified waveform for the circuit of Figure 8.34 for $v_i = 52\cos\omega t$ V.

Answer: 16.55 V

Page 547

CHECK YOUR UNDERSTANDING

In Example 8.8, what is the peak current if an offset diode model is used with $V_\gamma = 0.6$ V?

Answer: 0.554 A

CHECK YOUR UNDERSTANDING

Show that the DC output voltage of the full-wave rectifier of Figure 8.36 is $2Nv_{\text{Speak}}/\pi$.

Compute the peak voltage output of the bridge rectifier of Figure 8.38, assuming diodes with 0.6-V offset voltage and a 110-V rms AC supply.

Answer: 154.36 V

8.6 ZENER DIODES AND VOLTAGE REGULATION

For many applications, it is desirable that a DC supply be steady and ripple-free. Voltage regulators are used to ensure that the output of a DC supply is steady and relatively independent of load. The most common device employed in voltage regulation schemes is the Zener diode, which is designed and intended to be used when reverse-biased. The basic mechanism behind the Zener reverse breakdown effect was described in Section 8.2. It is important to recall that the mechanisms behind the Zener and avalanche reverse breakdown effects are different. This difference accounts for the difference in the range of breakdown voltages V_Z within which each effect dominates.

Figure 8.10 shows a generic diode *i*-*v* characteristic, with forward offset voltage V_γ and **reverse breakdown voltage** V_Z. Note the steep slope of the *i*-*v* characteristic near V_Z, which suggests that when $v_D \approx -V_Z$ the diode voltage will change very little for large changes in the diode current. It is exactly this property that makes the Zener diode a useful voltage regulator.

Although the slope of the *i*-*v* characteristic is not constant near $-V_Z$, for the sake of simplicity in introducing the basic principles of voltage regulation this slope will be assumed to be constant such that a Zener diode can be modeled with linear elements when it is reverse-biased near $v_D = -V_Z$.

Like other diodes, a Zener diode has three regions of operation:

1. When $v_D \leq V_\gamma$, the Zener diode is forward-biased and can be analyzed using the piecewise linear model shown in Figure 8.46.
2. When $-V_Z < v_D < V_\gamma$, the Zener diode is reverse-biased but has not reached breakdown. In this region, it can be modeled as an open-circuit.
3. For $v_D \leq -V_Z$, the Zener diode is reverse-biased and breakdown has ensued. In this region, it can be modeled using the piecewise linear model shown in Figure 8.47.

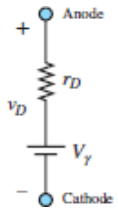


Figure 8.46 Zener diode forward bias model. Note the orientation of anode and cathode. Zener diodes are not, in general, designed for use in forward bias.

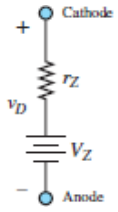


Figure 8.47 Zener diode reverse bias model for voltage regulation. Note the orientation of anode and cathode.

The combined effect of forward and reverse bias may be lumped into a single model with the aid of ideal diodes, as shown in Figure 8.48.

Page 548

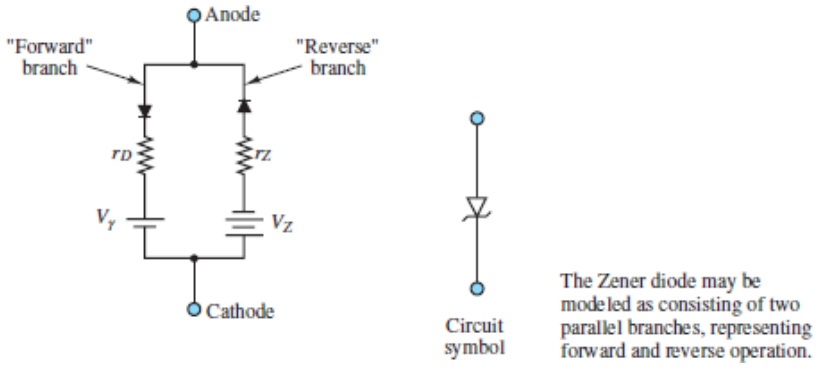


Figure 8.48 Complete model for Zener diode

To illustrate the operation of a Zener diode as a voltage regulator, consider the circuit of Figure 8.49(a), where the unregulated DC source V_S is regulated to the value of the Zener voltage V_Z . Note how the diode must be connected "upside down" to obtain a positive regulated voltage. Also note that when $v_S > V_Z$ the Zener diode is in reverse breakdown. (In practice, it is important that v_S remain greater than V_Z .) The source resistance R_S is essential because it allows the voltage difference $v_S - V_Z$ to be nonzero. If the diode resistance r_Z is small compared to R_S and R, the Zener diode model of Figure 8.47 can be approximated as a battery of strength V_Z , as shown in the simplified circuit of Figure 8.49(b). Typical values of r_Z are on the order of several ohms or smaller.

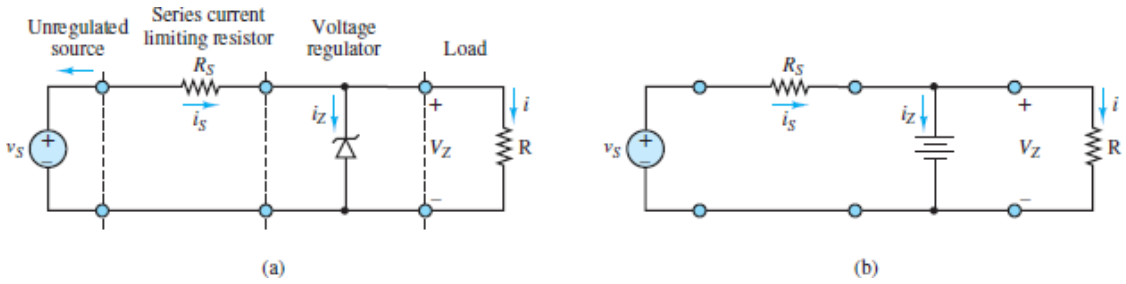


Figure 8.49 (a) A Zener diode voltage regulator; and (b) the simplest equivalent circuit

Three observations are sufficient to understand the operation of this voltage regulator:

1. The load voltage must equal V_Z as long as the Zener diode is in the reverse breakdown mode. Then:

$$i = \frac{V_Z}{R} \tag{8.20}$$

2. The output current is the nearly constant difference between the unregulated supply current i_S and the diode current i_Z:

$$i = i_S - i_Z \tag{8.21}$$

Any current in excess of that required to keep the load at the constant voltage V_Z is sent to ground through the diode. Thus, the Zener diode acts as a sink for any undesired source current.

Page 549

3. The source current is

$$i_S = \frac{v_S - V_Z}{R_S} \tag{8.22}$$

The examples that follow and the end-of-chapter exercises illustrate some of the considerations that arise in the design of a practical voltage regulator. One of these considerations is the power rating of the diode. The power P_Z dissipated by the diode is

$$P_Z = i_Z V_Z \tag{8.23}$$

Since V_Z is more or less constant, the power rating establishes an upper limit on the allowable diode current i_Z . This limit would be exceeded if the supply voltage rises unexpectedly or if the load is removed such that all the supply current sinks through the diode. The possibility of an open-circuit output must be accommodated in the design of a practical voltage regulator.

Another significant limitation occurs when the load resistance is small, thus requiring large amounts of current from the unregulated supply. In this case, the Zener diode is hardly taxed at all in terms of power dissipation, but the unregulated supply may not be able to provide the current required to sustain the load voltage. In this case, regulation fails to take place. Thus, in practice, the range of load resistances for which load voltage regulation may be attained is constrained to a finite interval:

$$R_{\min} \leq R \leq R_{\max} \tag{8.24}$$

where $R_{\max}$ is typically limited by the Zener diode power rating and $R_{\min}$ by the maximum supply current.

EXAMPLE 8.10 Determining the Power Rating of a Zener Diode

Problem

Design a regulator similar to the one depicted in Figure 8.49(a). Determine the minimum acceptable power rating of the Zener diode.

Solution

Known Quantities: $v_S = 24$ V; $V_Z = 12$ V; $R_S = 50\ \Omega$; $R = 250\ \Omega$.

Find: The maximum power dissipated by the Zener diode under worst-case conditions.

Assumptions: Use the piecewise linear Zener diode model (Figure 8.48) with $r_Z = 0$.

Analysis: When the regulator operates according to the intended design specifications, that is, with a 250-Ω load, the source and load currents are

$$i_S = \frac{v_S - V_Z}{R_S} = \frac{12}{50} = 0.24 \text{ A}$$

$$i = \frac{V_Z}{R} = \frac{12}{250} = 0.048 \text{ A}$$

Thus, the Zener current would be

$$i_Z = i_S - i = 0.192 \text{ A}$$

Page 550

corresponding to a nominal power dissipation

$$P_Z = i_Z V_Z = 0.192 \times 12 = 2.304 \text{ W}$$

However, if the load were accidentally (or intentionally) disconnected from the circuit, all the load current would be diverted to flow through the Zener diode. Thus, the *worst-case* Zener current is the source current since the Zener diode would sink all the source current for an open-circuit load:

$$i_{Z\max} = i_S = \frac{v_S - V_Z}{R_S} = \frac{12}{50} = 0.24 \text{ A}$$

Therefore the maximum power dissipation that the Zener diode must sustain is

$$P_{Z\max} = i_{Z\max} V_Z = 2.88 \text{ W}$$

Comments: A safe design would exceed the value of $P_{Z\max}$ computed above. For example, one might select a 3-W Zener diode.

EXAMPLE 8.11 Calculation of Allowed Load Resistances for a Given Zener Regulator

Problem

Calculate the allowable range of load resistances for the Zener regulator of Figure 8.50 such that the diode power rating is not exceeded.

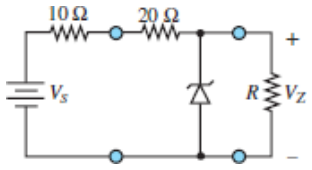


Figure 8.50

Solution

Known Quantities: $V_S = 50$ V; $V_Z = 14$ V; $P_Z = 5$ W.

Find: The smallest and largest values of R for which a load voltage regulation of 14 V is achieved and which do not cause the diode power rating to be exceeded.

Assumptions: Use the piecewise linear Zener diode model of Figure 8.47 with $r_Z = 0$.

Analysis:

1. *Determining the minimum acceptable load resistance.* To determine the minimum acceptable load, we observe that the regulator can at most supply the load with the amount of current that can be provided by the source. Thus, the minimum theoretical resistance can be computed by assuming that all the source current goes to the load and that the load voltage is regulated at the nominal value.

$$R_{\min} = \frac{V_Z}{i_S} = \frac{V_Z}{(V_S - V_Z)/30} = \frac{14}{36/30} = 11.7\ \Omega$$

   If the load required any more current, the source would not be able to supply it. Note that for this value of the load, the Zener diode dissipates zero power because the Zener current is zero.

2. *Determining the maximum acceptable load resistance.* The second constraint we need to invoke is the power rating of the diode. For the stated 5-W rating, the maximum Zener current is

$$i_{Z\max} = \frac{P_Z}{V_Z} = \frac{5}{14} = 0.357 \text{ A}$$

Page 551

Since the source will generate

$$i_{S\max} = \frac{V_S - V_Z}{30} = \frac{50 - 14}{30} = 1.2 \text{ A}$$

the load must not require any less than $1.2 - 0.357 = 0.843$ A; if it required any less current (i.e., if the resistance were too large), the Zener diode would be forced to sink more current than its power rating permits. From this requirement we can compute the maximum allowable load resistance

$$R_{\max} = \frac{V_Z}{i_{S\max} - i_{Z\max}} = \frac{14}{0.843} = 16.6\ \Omega$$

Finally, the range of allowable load resistance is $11.7\ \Omega \le R \le 16.6\ \Omega$.

Comments: This regulator circuit is not practical because it *cannot* operate with an open-circuit load! Typically, Zener diodes are used to regulate source voltages not much greater than the Zener voltage.

EXAMPLE 8.12 Effect of Nonzero Zener Resistance in a Regulator Problem

Calculate the amplitude of the ripple present in the regulator output voltage shown in Figure 8.51. The unregulated supply voltage is depicted in Figure 8.52.

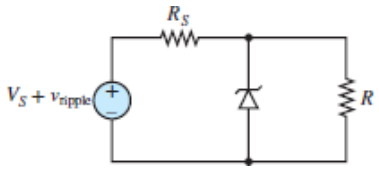


Figure 8.51

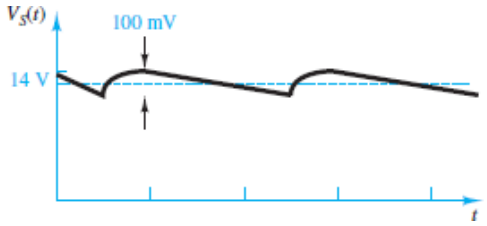


Figure 8.52

Solution

Known Quantities: $V_S = 14$ V; $v_{\text{ripple}} = 100$ mV; $V_Z = 8$ V; $r_Z = 10\ \Omega$; $R_S = 50\ \Omega$; $R = 150\ \Omega$.

Find: Amplitude of ripple component in load voltage.

Assumptions: Use the piecewise linear Zener diode model shown in Figure 8.47.

Analysis: Consider the DC and AC equivalent circuits shown in Figure 8.53.

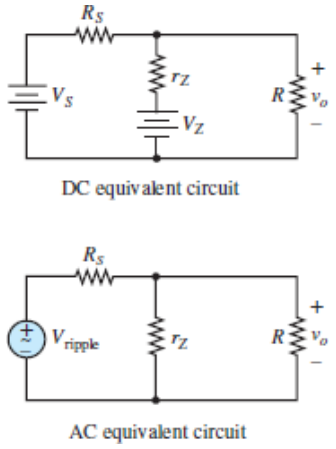


Figure 8.53

1. *DC equivalent circuit.* In the DC equivalent circuit, the output voltage has two components: one due to the unregulated DC supply V_S and one due to the Zener diode V_Z. Apply superposition and voltage division to obtain:

$$V_o = V_S\left(\frac{r_Z \| R}{r_Z \| R + R_S}\right) + V_Z\left(\frac{R_S \| R}{R_S \| R + r_Z}\right) = 2.21 + 6.32 = 8.53 \text{ V}$$

2. *AC equivalent circuit.* The AC component of the output voltage is

$$v_o = v_{\text{ripple}}\left(\frac{r_Z \| R}{r_Z \| R + R_S}\right) = 0.016 \text{ V}$$

Thus, 16 mV of ripple is present in the output voltage, or only 16 percent of the source ripple. Regulation!

Comments: Note that the DC output voltage is affected by the unregulated source voltage; if the unregulated supply were to fluctuate significantly, the regulated voltage would also change. Thus, one of the effects of the Zener resistance is imperfect regulation. If the r_Z is much smaller than both R_S and R, its impact will not be as pronounced.

CHECK YOUR UNDERSTANDING

In Example 8.10, how would the minimum power rating change if the load were reduced to 100 Ω?

Answer: The worst-case power rating would not change.

CHECK YOUR UNDERSTANDING

In Example 8.11, what should the power rating of the Zener diode be to withstand operation with an open-circuit load?

Answer: $P_{Z\,\text{max}} = 16.8$ W

CHECK YOUR UNDERSTANDING

Compute the actual DC load voltage and the percent of ripple reaching the load (relative to the initial 100-mV ripple) for the circuit of Example 8.12 if $r_Z = 1$ Ω.

Answer: 8.06 V, 2 percent

Page 553

FOCUS ON MEASUREMENTS

Diode Peak Detector Circuit for Capacitive Displacement Transducer

Another common application of semiconductor diodes, the *peak detector,* is very similar in appearance to the half-wave rectifier with capacitive filtering as shown in Figure 8.56. One of its more classic applications is in the demodulation of amplitude-modulated (AM) signals.

In Chapter 3, a capacitive displacement transducer was introduced in the two Focus on Measurements boxes, "Capacitive Displacement Transducer and Microphone." It took the form of a parallel-plate capacitor composed of a fixed plate and a movable plate. The capacitance of this variable capacitor was shown to be a function of displacement; that is, it was shown that a movable-plate capacitor can serve as a linear transducer. Recall the expression derived in Chapter 3

$$C = \frac{8.854 \times 10^{-3} A}{x} \quad \text{pF}$$

where C is the capacitance in picofarads, A is the area of the plates in square millimeters, and x is the variable separation distance in millimeters. The nominal plate separation is d. If the capacitor is placed in an AC circuit, its impedance will be determined by the expression

$$Z_C = \frac{1}{j\omega C}$$

so that

$$Z_C = \frac{x}{j\omega(8.854 \times 10^{-3})\text{A}}$$

Thus, at a fixed frequency ω, the impedance of the capacitor will vary linearly with displacement. This property may be exploited in the bridge circuit of Figure 8.54, where a differential-pressure transducer is shown made of two movable-plate capacitors. Such a transducer can be constructed (see Figure 3.7) so that if the capacitance of one capacitor increases as a consequence of a pressure difference across the transducer, the capacitance of the other decreases by a corresponding amount, at least for small effective displacements Δx.

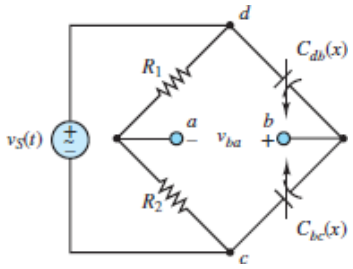


Figure 8.54 Bridge circuit for displacement transducer

Using phasor notation, the output voltage was derived in Chapter 3 to be:

$$\mathbf{V}_{ba}(j\omega) = \mathbf{V}_S(j\omega)\frac{\Delta x}{2d}$$

provided that $R_1 = R_2$. Thus, the output voltage will vary as a scaled version of the input voltage in proportion to the displacement. A typical $v_{ba}(t)$ is displayed in Figure 8.55 for a 0.05-mm "triangle" diaphragm displacement, with $d = 0.5$ mm and $\mathbf{V}_s$ a 50-Hz sinusoid with 1-V amplitude. Clearly, although the output voltage Page 554is a function of the effective displacement Δx, it is not in a convenient form since the displacement is proportional to the amplitude of the sinusoidal peaks.

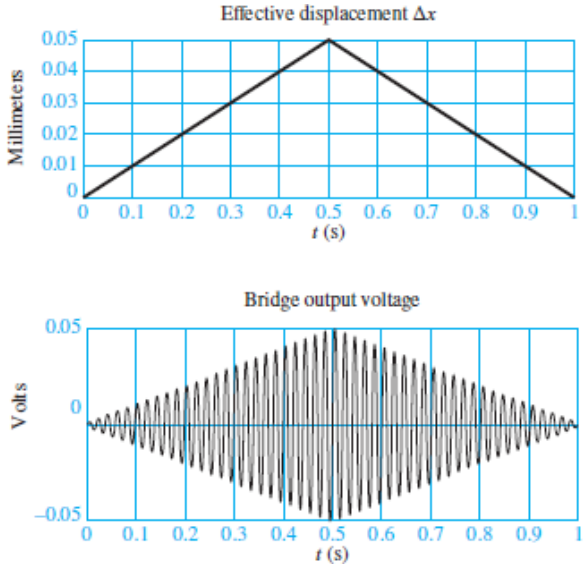


Figure 8.55 Effective displacement Δx and bridge output voltage waveforms

The diode peak detector is a circuit capable of tracking the sinusoidal peaks without exhibiting the oscillations of the bridge output voltage. The peak detector operates by rectifying and filtering the bridge output in a manner similar to that of the circuit of Figure 8.34. The peak detector circuit is shown in Figure 8.56, and the response of a practical peak detector is shown in Figure 8.57. Its operation is based on the rectification property of the diode, coupled with the filtering effect of the shunt capacitor, which acts as a low-pass filter.

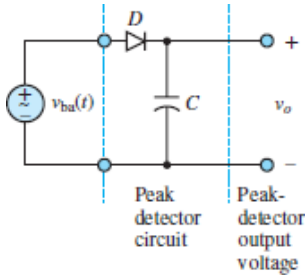


Figure 8.56 Peak detector circuit

From a time domain perspective, when the diode is forward-biased ($v_{ba} \leq V_\gamma$ for an offset diode) the capacitor charges at a rate set by the time constant $R_D C$, where R_D is the forward-biased effective resistance of the diode. When reverse-biased, the diode prevents any significant discharging of the capacitor. Thus, eventually, in steady-state, the capacitor voltage oscillates slightly around the peak voltage as shown in Figure 8.57.

Page 555

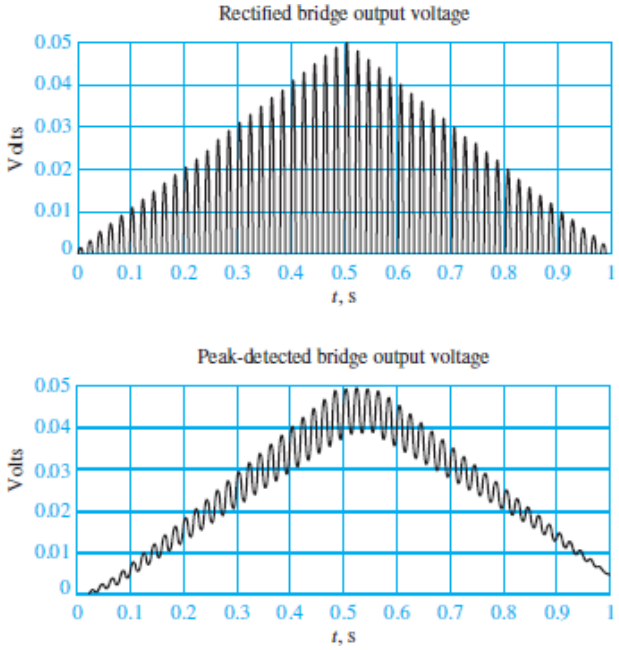


Figure 8.57 Rectified and peak-detected bridge output voltage waveforms

FOCUS ON MEASUREMENTS

Diode Thermometer

Problem:

An interesting application of a diode, based on the diode equation, is an electronic thermometer. The concept is based on the empirical observation that if the current through a diode is nearly constant, the diode voltage is nearly a linear function of temperature, as shown in Figure 8.58(a).

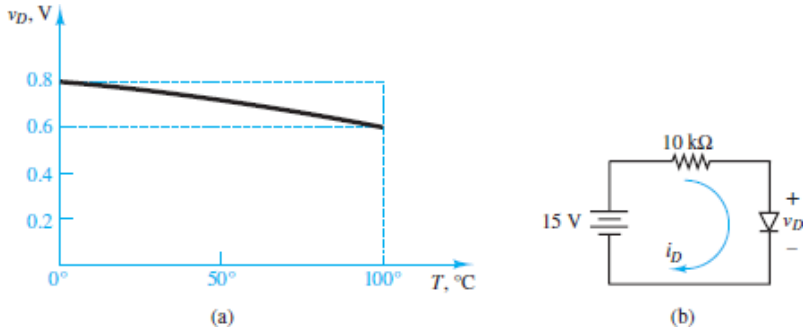


Figure 8.58

Page 556

1. Show that i_D in the circuit of Figure 8.58(b) is nearly constant in the face of variations in the diode voltage v_D. This can be done by computing the percentage change in i_D for a given percentage change in v_D. Assume that v_D changes by 10 percent, from 0.6 to 0.66 V.
2. On the basis of the graph of Figure 8.58(a), write an equation for $v_D(T°)$ of the form

$$v_D = \alpha T° + \beta$$

Solution:

1. With reference to the circuit of Figure 8.58(b), the current i_D is

$$i_D = \frac{15 - v_D}{10} \quad \text{mA}$$

For

$$v_D = 0.8 \text{ V}(0°),\ i_D = 1.42 \text{ mA}$$
$$v_D = 0.7 \text{ V}(50°),\ i_D = 1.43 \text{ mA}$$
$$v_D = 0.6 \text{ V}(100°),\ i_D = 1.44 \text{ mA}$$

The percentage change in v_D over the full scale of the thermometer (assuming the midrange temperature of 50° to be the reference value) is

$$\Delta v_D\% = \pm\frac{0.1\text{ V}}{0.7\text{ V}} \times 100 = \pm 14.3\%$$

The corresponding percentage change in i_D is

$$\Delta i_D\% = \pm\frac{0.01\text{ mA}}{1.43\text{ mA}} \times 100 = \pm 0.7\%$$

Thus, i_D is nearly constant over the range of operation of the diode thermometer.

2. The diode voltage versus temperature equation can be extracted from the graph of Figure 8.58(a):

$$v_D(T) = \frac{(0.8 - 0.6)\text{ V}}{(0 - 100)°\text{C}}\,T + 0.8\text{ V} = -0.002T + 0.8\text{ V}$$

Comments: The graph of Figure 8.58(a) was obtained experimentally by calibrating a commercial diode in both hot water and an ice bath. The circuit of Figure 8.58(b) is rather simple, and one could fairly easily design a better constant-current source; however, this example illustrates that an inexpensive diode can serve quite well as the sensing element in an **electronic thermometer.**

8.7 PHOTODIODES

Another property of semiconductor materials that finds common application in measurement systems is their response to light energy. In appropriately fabricated diodes, called photodiodes, when light reaches the depletion region of a *pn* junction, photons cause hole-electron pairs to be generated by a process called *photoionization.* This effect can be achieved by using a surface material that is transparent to light. As a consequence, the reverse saturation current I_0 depends on the light intensity (i.e., on the number of incident photons), in addition to the other factors mentioned earlier in Section 8.2. In a photodiode, the reverse current is Page 557given by $-(I_0 + I_p)$, where I_p is the additional current generated by photoionization. The result is depicted in the family of solid curves of Figure 8.59, where the diode characteristic is shifted downward by an amount related to the additional current generated by photoionization. Figure 8.59 depicts the appearance of the *i-v* characteristic of a photodiode for various values of I_p, where the *i-v* curve is shifted to lower values for progressively larger values of I_p. The circuit symbol is depicted in Figure 8.60.

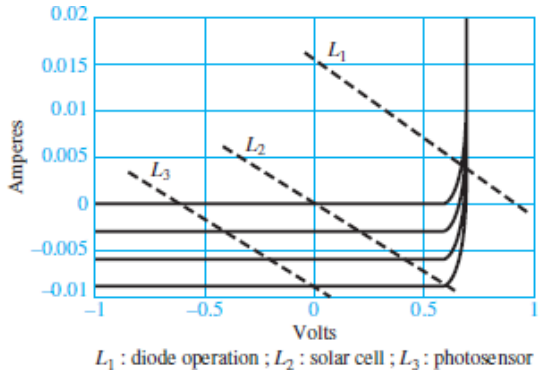


Figure 8.59 Photodiode *i*-*v* curves (—) and three load lines (---)

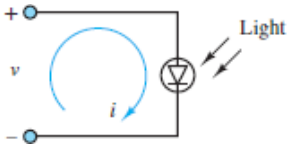


Figure 8.60 Photodiode circuit symbol

Also displayed (as dashed lines) in Figure 8.59 are three load lines, which depict the three modes of operation of a photodiode. Curve L_1 represents the source network of a diode operating under forward bias. Note that the operating point of the device is in the positive i, positive v (first) quadrant of the i-v plane; thus, the diode dissipates positive power in this mode and is therefore a passive device, as we already know. On the other hand, load line L_2 represents the source network of a photodiode operating as a **solar cell**; in this mode, the operating point is in the negative i, positive v, or fourth, quadrant, and therefore the power dissipated by the diode is *negative*. In other words, the photodiode is generating power by converting light energy to electric energy. Note further that the load line intersects the voltage axis at zero, meaning that no supply voltage is required to bias the photodiode in the solar-cell mode. Finally, load line L_3 represents the source network of a photodiode operating as a light sensor: when the diode is reverse-biased, the current flowing through the diode is determined by the light intensity; thus, the diode current changes in response to changes in the incident light intensity.

The operation of the photodiode can also be reversed, in principle, by forward-biasing the diode and causing a significant level of recombination to take place in the depletion region. Some of the energy released is converted to light energy by the emission of photons. Thus, a diode operating in this mode emits light when forward-biased. Photodiodes used in this way are called **light-emitting diodes (LEDs)**; they exhibit a typical forward (offset) voltage of 1.6 to 3.4 V, depending upon the color of the LED. The circuit symbol for the LED is shown in Figure 8.61.

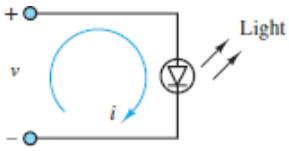


Figure 8.61 Light-emitting diode circuit symbol

Gallium arsenide (GaAs) is one of the more popular substrates for creating LEDs; gallium phosphide (GaP) and the alloy $GaAs_{1-x}P_x$ are also quite common. Table 8.1 lists combinations of materials and dopants used for common LEDs and the colors they emit. The dopants are used to create the necessary *pn* junction.

Page 558

Table 8.1 LED materials and wavelengths

Material	Dopant	Wavelength (nm)	Color
GaAs	Zn	900	Infrared
GaAs	Si	910–1,020	Infrared
GaP	N	570	Green
GaP	N	590	Yellow
GaP	Zn, O	700	Red
$GaAs_{0.6}P_{0.4}$		650	Red
$GaAs_{0.35}P_{0.65}$	N	632	Orange
$GaAs_{0.15}P_{0.85}$	N	589	Yellow

The construction of a typical LED is shown in Figure 8.62, along with the schematic representation for an LED. A shallow *pn* junction is created with electrical contacts made to both *p* and *n* regions. As much of the upper surface of the *p* material is uncovered as possible so that light can leave the device unimpeded. It is important to note that, actually, only a relatively small fraction of the emitted light leaves the device; the majority stays inside the semiconductor. A photon that stays inside the device will eventually collide with an electron in the valence band, and the collision will force the electron into the conduction band, emitting an electron-hole pair and absorbing the photon. To minimize the probability that a photon will be absorbed before it has an opportunity to leave the LED, the depth of the *p*-doped region is left very thin. Also, it is advantageous to have most of the recombinations that emit photons occur as close to the surface of the diode as possible. This is made possible by various doping schemes, but even so, of all the carriers going through the diode, only a small fraction emit photons that are able to leave the semiconductor. An important application of LEDs and photodiodes is shown in Figure 8.63.

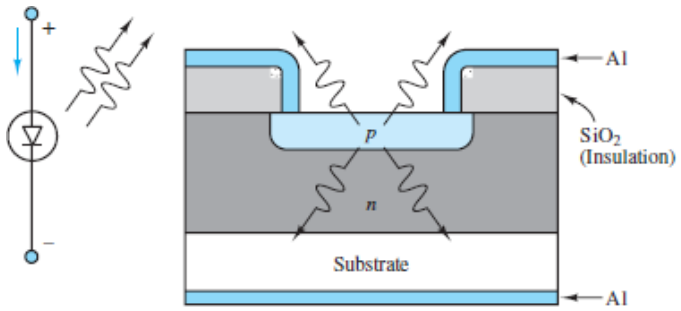


Figure 8.62 Light-emitting diode (LED)

A simple LED circuit is shown in Figure 8.64. From the standpoint of circuit analysis, LED characteristics are very similar to those of the silicon diode, except that the offset voltage is usually quite a bit larger. Typical values of V_γ are in the range of 1.6 to 3.4 V, depending upon the LED color, and operating currents can range from 10 to 100 mA. Manufacturers usually specify an LED's characteristics by giving the rated operating-point current and voltage.

FOCUS ON MEASUREMENTS

Opto-isolators

One of the common applications of photodiodes and LEDs is the **optocoupler**, or **opto-isolator**. This device, which is usually enclosed in a sealed package, uses the

light-to-current and current-to-light conversion property of photodiodes and LEDs to provide signal connection between two circuits without any need for electrical connections. Figure 8.63 depicts the circuit symbol for the opto-isolator.

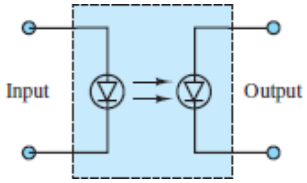


Figure 8.63

Because diodes are nonlinear devices, the opto-isolator is not used in transmitting analog signals: the signals would be distorted because of the nonlinear diode *i-v* characteristic. However, opto-isolators find a very important application when on/off signals need to be transmitted from high-power machinery to delicate computer control circuitry. The optical interface ensures that potentially damaging large currents cannot reach delicate instrumentation and computer circuits.

EXAMPLE 8.13 Analysis of Light-Emitting Diode

Problem

For the circuit of Figure 8.64, determine (1) the LED power consumption, (2) the resistance R_s, and (3) the power required by the voltage source.

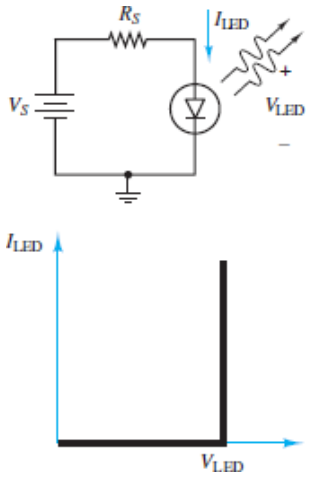


Figure 8.64 LED drive circuit and *i-v* characteristic based on offset model

Solution

Known Quantities: Diode operating point: $V_{\text{LED}} = 1.7$ V; $I_{\text{LED}} = 40$ mA; $V_S = 5$ V.

Find: P_{LED}; R_S; P_S.

Assumptions: Use the offset diode model.

Analysis:

1. The power consumption of the LED is determined directly from the specification of the operating point:

$$P_{\text{LED}} = V_{\text{LED}} \times I_{\text{LED}} = 68 \text{ mW}$$

2. To determine the required value of R_S to achieve the desired operating point, we apply KVL around the circuit of Figure 8.64:

$$V_S = I_{\text{LED}} R_S + V_{\text{LED}}$$

$$R_S = \frac{V_S - V_{\text{LED}}}{I_{\text{LED}}} = \frac{5 - 1.7}{40 \times 10^{-3}} = 82.5 \ \Omega$$

n

Page 560

3. To satisfy the power requirement of the circuit, the battery must be able to supply 40 mA to the diode. Thus,

$$P_S = V_S \times I_{\text{LED}} = 200 \text{ mW}$$

Comments: A more practical LED biasing (drive) circuit is found in Example 9.3.

CHECK YOUR UNDERSTANDING

Determine the source resistance required to bias the LED of Example 8.13 if the required LED current is 24 mA.

Answer: 137.5 Ω

Conclusion

This chapter introduces the topic of electronic devices by presenting the semiconductor diode.

Upon completing this chapter, you should have mastered the following learning objectives:

1. *Understand the basic principles underlying the physics of semiconductor devices in general and of the pn junction in particular. Become familiar with the diode equation and i-v characteristic.* Semiconductors have conductive properties that fall between those of conductors and insulators. These properties make the materials useful in the construction of many electronic devices that exhibit nonlinear *i-v* characteristics. Of these devices, the diode is one of the most commonly employed.
2. *Use various circuit models of the semiconductor diode in simple circuits. These are divided into two classes: the large-signal models, useful to study rectifier circuits, and the small-signal models.* The semiconductor diode acts as a one-way current valve, permitting the flow of current only when it is forward-biased. The behavior of the diode is described by an exponential equation, but it is possible to approximate the operation of the diode by means of simple circuit models. The simplest (ideal) model treats the diode either as a short-circuit (when it is forward-biased) or as an open-circuit (when it is reverse-biased). The ideal model can be extended to include an offset voltage, which represents the contact potential at the diode *pn* junction. A further model, useful for small-signal circuits, includes a resistance that models the forward resistance of the diode. With the aid of these models it is possible to analyze diode circuits by using the DC and AC circuit analysis methods of earlier chapters.
3. *Study practical full-wave rectifier circuits, and learn to analyze and determine the practical specifications of a rectifier by using large-signal diode models.* One of the most important properties of the diode is its ability to rectify AC voltages and currents. Diode rectifiers can be of the half-wave and full-wave types. Full-wave rectifiers can be constructed in a two-diode configuration or in a four-diode bridge configuration. Diode rectification is an essential element of DC power supplies. Another important part of a DC power supply is the filtering, or smoothing, that is usually accomplished by using capacitors.
4. *Understand the basic operation of Zener diodes as voltage references, and use simple circuit models to analyze elementary voltage regulators.* In addition to rectification and filtering, the power supply requires output voltage regulation. Zener diodes can be used to provide a voltage reference that is useful in voltage regulators.

5. *Understand the basic principle of operation of photodiodes, including solar cells, photosensors, and light-emitting diodes.* Semiconductor material properties can also be affected by light intensity. Certain types of diodes, known as *photodiodes,* find applications in light detectors, solar cells, or light-emitting diodes.

Page 561

HOMEWORK PROBLEMS

Section 8.1: Electrical Conduction in Semiconductor Devices

8.1 In a semiconductor material, the net charge is such that the density of positive charges equals the density of negative charges. Charge carriers (free electrons and holes) and ionized dopant atoms have a charge equal to the magnitude of one electronic charge. Therefore the charge neutrality equation (CNE) is

$$p_o + N_d^+ - n_o - N_a^- = 0$$

where

n_o = equilibrium negative carrier density

p_o = equilibrium positive carrier density

N_a^- = ionized acceptor density

N_d^+ = ionized donor density

The carrier product equation (CPE) states that as a semiconductor is doped, the product of the charge carrier densities remains constant:

$$n_o p_o = \text{const} \qquad \text{(Equation 8.2)}$$

For intrinsic silicon at $T = 300$ K:

$$\begin{aligned}\text{Const} &= n_{io}\, p_{io} = n_{io}^2 = p_{io}^2 \\ &= \left(1.5 \times 10^{16} \frac{1}{\text{m}^3}\right)^2 = 2.25 \times 10^{32} \frac{1}{\text{m}^2}\end{aligned}$$

The semiconductor material is *n*- or *p*-type depending on whether donor or acceptor doping is greater. Almost all dopant atoms are ionized at room temperature. Assume intrinsic silicon is doped such that

$$N_A \approx N_a^- = 10^{17}\frac{1}{\text{m}^3} \qquad N_d = 0$$

Determine:

a. If this is an *n*- or *p*-type extrinsic semiconductor.

b. Which are the major and which are the minority charge carriers.

c. The density of majority and minority carriers.

8.2 Assume intrinsic silicon is doped such that

$$N_a \approx N_a^- = 10^{17}\frac{1}{\text{m}^3} \qquad N_d \approx N_d^+ = 5\times 10^{18}\frac{1}{\text{m}^3}$$

Determine:

a. If this is an *n*- or *p*-type extrinsic semiconductor.

b. Which are the majority and which are the minority charge carriers.

c. The density of majority and minority carriers.

8.3 Describe the microscopic structure of semiconductor materials. What are the three most commonly used semiconductor materials?

8.4 Describe the thermal production of charge carriers in a semiconductor and how this process limits the operation of a semiconductor device.

8.5 Describe the properties of donor and acceptor dopant atoms and how they affect the densities of charge carriers in a semiconductor material.

Section 8.2: The *pn* Junction and the Semiconductor Diode

8.6 Describe the behavior of the charge carriers and ionized dopant atoms in the vicinity of a semiconductor *pn* junction that creates the potential (energy) barrier that inhibits charge carriers from crossing the junction.

Section 8.3: Large-signal Models for the Semiconductor Diode

8.7 Consider the circuit of Figure P8.7. Determine whether the diode is conducting. Assume V_A = 12 V, V_B = 10 V, and that the diode is ideal.

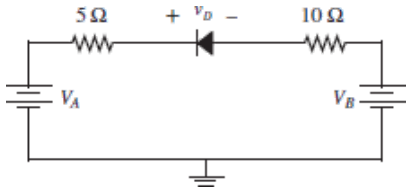


Figure P8.7

8.8 Repeat Problem 8.7 for $V_A = 12$ V and $V_B = 15$ V.

8.9 Consider the circuit of Figure P8.9. Determine whether the diode is conducting. Assume $V_A = 12$ V, $V_B = 10$ V, $V_C = 5$ V and that the diode is ideal.

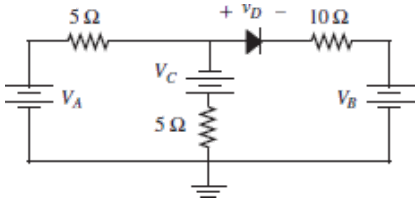


Figure P8.9

8.10 Repeat Problem 8.9 for $V_B = 15$ V.

8.11 Repeat Problem 8.9 for $V_C = 15$ V.

8.12 Repeat Problem 8.9 for $V_B = 15$ V and $V_C = 10$ V.

Page 562

8.13 For the circuit of Figure P8.13, sketch $i_D(t)$ using:

a. The ideal diode model.

b. The offset diode model with $V_\gamma = 0.6$ V.

c. The piecewise linear diode model (see Section 8.4) with $r_D = 1$ kΩ and $V_\gamma = 0.6$ V.

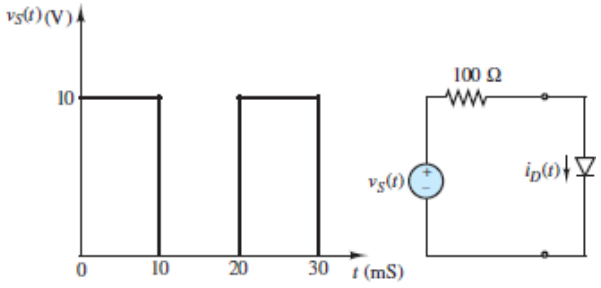


Figure P8.13

8.14 For the ideal diode circuit of Figure P8.14, find the range of V_{in} for which D_1 is forward-biased.

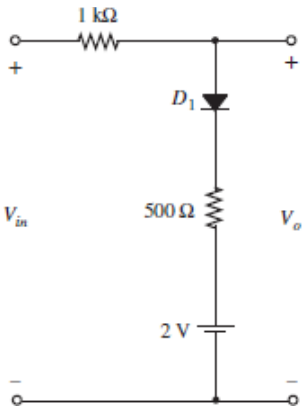


Figure P8.14

8.15 One of the more interesting applications of a diode, based on the diode equation, is an electronic thermometer. The concept is based on the empirical observation that if the current through a diode is nearly constant, the diode voltage is nearly a linear function of the diode temperature, as shown in Figure P8.15(a).

a. Show that i_D in the circuit of Figure P8.15(b) is nearly constant in the face of variations in the diode voltage v_D. To do so, compute the percent change in i_D for a given percent change in v_D. Assume that v_D changes by 5 percent from 0.6 V to 0.63 V.

b. On the basis of the graph of Figure P8.15(a), write an equation for v_D (T°) of the form

$$v_D = \alpha T^\circ + \beta$$

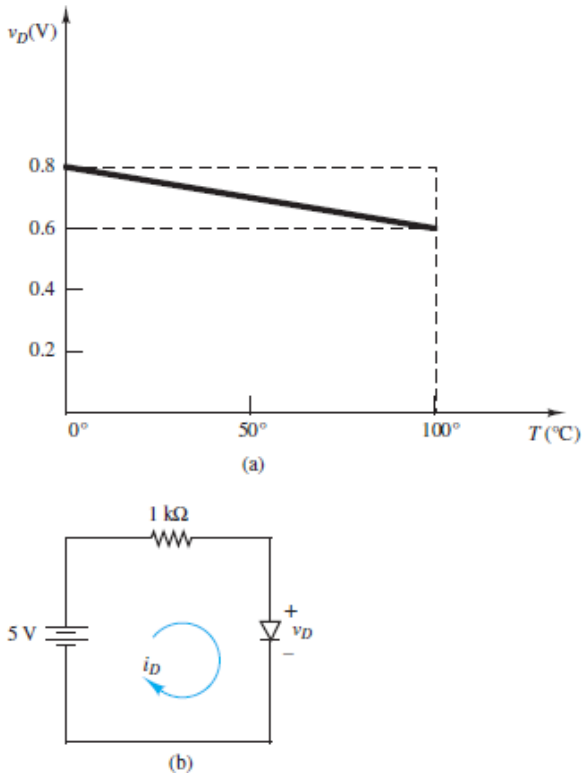


Figure P8.15

8.16 Find expressions for the voltage v_o in Figure P8.16, assuming D is an ideal diode, for positive and negative values of v_S. Sketch a plot of v_o versus v_S.

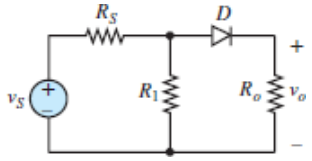


Figure P8.16

8.17 Repeat Problem 8.16, using the offset diode model.

8.18 Find the power dissipated in diode D, and the power dissipated in R in Figure P8.18. Use the exponential diode equation and assume $R = 2$ kΩ, $V_S = 5$V, $V_D = 900$ mV, $q/KT = \frac{1}{52}$ mV, and $I_0 = 15$ nA.

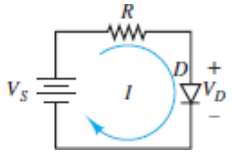


Figure P8.18

8.19 Determine the Thévenin equivalent network seen by the ideal diode D in Figure P8.19, and use it to Page 563determine the diode current i_D. Also, solve for the currents i_1 and i_2. Assume $R_1 = 5$ kΩ, $R_2 = 3$ kΩ, $V_{cc} = 10$ V, and $V_{dd} = 15$ V.

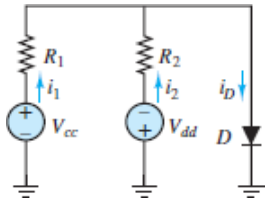


Figure P8.19

8.20 In Figure P8.20, assume a sinusoidal source $V_S = 50$ V rms, $R = 170$ Ω, and $V_\gamma = 0.6$ V. Use the offset diode model for a silicon diode to determine:

a. The maximum forward current.

b. The peak reverse voltage across the diode.

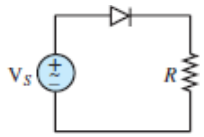


Figure P8.20

8.21 Determine voltages V_2 assuming the diodes are ideal in each of the configurations shown in Figure P8.21.

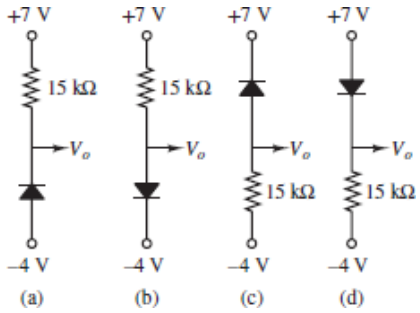


Figure P8.21

8.22 In the circuit of Figure P8.22, find the range of V_{in} for which D_1 is forward-biased. Assume ideal diodes.

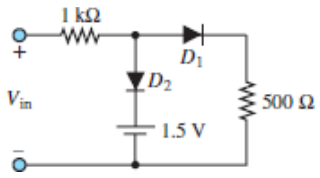


Figure P8.22

8.23 Determine which diodes are forward-biased and which are reverse-biased in the configurations shown in Figure P8.23. Assuming a 0.7-V drop across each forward-biased diode, determine v_{out}.

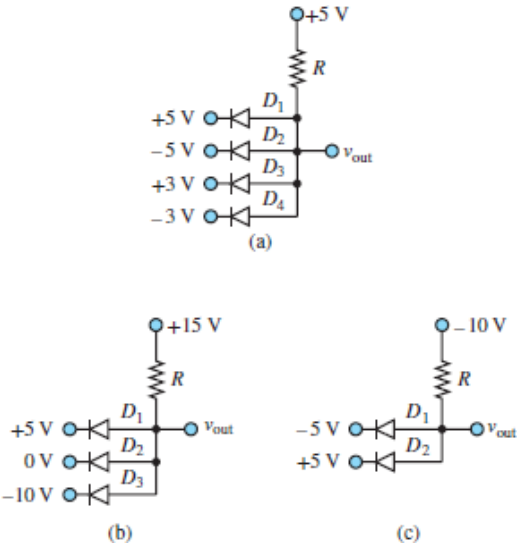


Figure P8.23

8.24 Sketch the output waveform and the voltage transfer characteristic for the circuit of Figure P8.24. Assume an ideal diode, $v_s(t) = 8 \sin(\pi t)$, $V_1 = 3$ V, $R_1 = 8\ \Omega$, and $R_2 = 5\ \Omega$.

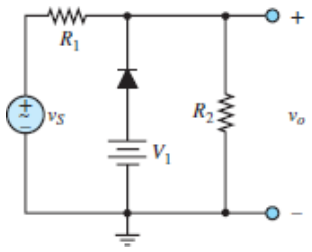


Figure P8.24

8.25 Repeat Problem 8.24, using the offset diode model with $V_\gamma = 0.55$ V.

Section 8.4: Small-signal Models for the Semiconductor Diode

8.26 Repeat Problem 8.24 for $v_S(t) = 1.5 \sin(2{,}000\,\pi t)$, $V_1 = 1$ V, and $R_1 = R_2 = 1$ kΩ. Use the piecewise linear model with $r_D = 200\ \Omega$.

8.27 The silicon diode shown in Figure P8.27 is described by:

$$i_D = I_o(e^{v_D/V_T} - 1)$$

where at $T = 300$ K

$$I_o = 250 \times 10^{-12}\,\text{A} \qquad V_T = \frac{kT}{q} \approx 26\ \text{mV}$$
$$v_S = 4.2\ \text{V} + 110\cos(\omega t)\ \text{mV}$$
$$\omega = 377\ \text{rad/s} \qquad R = 7\ \text{k}\Omega$$

Page 564

Determine the current i_D at the operating point Q:

a. Using the diode offset model.

b. By graphically solving the circuit characteristic (i.e., the DC load-line equation) and the device characteristic (i.e., the diode equation).

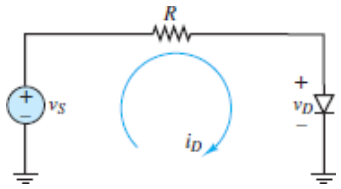


Figure P8.27

8.28 Repeat Problem 8.27 using the following data:

$$i_D = I_o(e^{v_D/V_T} - 1)$$

where at $T = 300$ K

$$I_o = 2.030 \times 10^{-15}\,\text{A} \qquad V_T = \frac{kT}{q} \approx 26\ \text{mV}$$
$$v_S = 5.3\ \text{V} + 7\cos(\omega t)\ \text{mV}$$
$$\omega = 377\ \text{rad/s} \qquad R = 4.6\ \text{k}\Omega$$

8.29 A diode with the *i-v* characteristic shown in Figure 8.8 is connected in series with a 5-V DC voltage source (in the forward-bias direction) and a load resistance of 200 Ω, similar to that shown in Figure P8.34. Determine:

a. The load current and voltage.

b. The power dissipated by the diode.

c. The load current and voltage if the load is changed to 100 Ω and 500 Ω.

8.30 A diode with the *i-v* characteristic shown in Figure 8.28 is connected in series with a 2-V DC source (in the forward-bias direction) and a 200-Ω load resistance, similar to that shown in Figure P8.34. Determine:

a. The load current and voltage.

b. The power dissipated by the diode.

c. The load current and voltage if the load is changed to 100 Ω and 300 Ω.

8.31 The silicon diode shown in Figure P8.32 is described by:

$$i_D = I_o(e^{v_D/V_T} - 1)$$

where at $T = 300$ K

$$I_o = 250 \times 10^{-12}\,\text{A} \qquad V_T = \frac{kT}{q} \approx 26\,\text{mV}$$
$$v_S = V_S + v_s = 4.2\,\text{V} + 110\cos(\omega t)\,\text{mV}$$
$$\omega = 377\,\text{rad/s} \qquad R = 7\,\text{k}\Omega$$

The DC operating (quiescent) point Q and the AC small-signal equivalent resistance at Q are

$$I_{DQ} = 0.548\,\text{mA} \qquad V_{DQ} = 0.365\,\text{V} \qquad r_d = 47.45\,\Omega$$

Determine the AC voltage across the diode and the AC current through it.

8.32 The silicon diode shown in Figure P8.32 is in series with two voltage sources and a resistor, where:

$$R = 2.2\,\text{k}\Omega \qquad V_{S2} = 3\,\text{V} \qquad V_r = 0.7\,\text{V}$$

Determine the minimum value of V_{S1} at which the diode will be forward-biased and conduct charge.

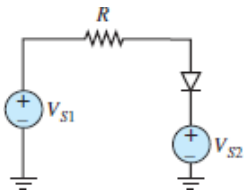


Figure P8.32

8.33 Find the average value of the output voltage v_o shown in Figure P8.33. Assume $v_{in} = 10 \sin(\omega t)$ V, $C = 80$ nF, and $V_\gamma = 0.5$ V. (See the Focus on Measurements box "Diode Peak Detector. . .".)

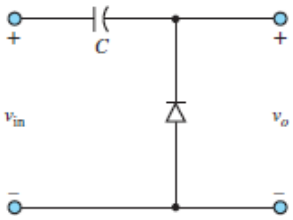


Figure P8.33

8.34 The circuit of Figure P8.34 is driven by a sinusoidal source $v_S(t) = 6 \sin(314t)$ V. Determine the average and peak diode currents, using:

a. The ideal diode model.

b. The offset diode model.

c. The piecewise linear model with resistance r_D.

Assume $R_o = 200\ \Omega$, $r_D = 25\ \Omega$, and $V_\gamma = 0.8$ V.

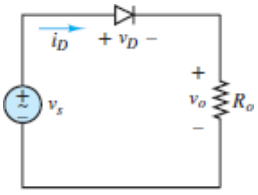


Figure P8.34

Section 8.5: Rectifier Circuits

8.35 A half-wave rectifier produces an average voltage of 50 V at its output.

a. Draw a schematic diagram of the circuit.

b. Sketch the output voltage waveform.

c. Determine the peak value of the output voltage.

d. Sketch the input voltage waveform.

e. What is the rms voltage at the input?

8.36 A half-wave rectifier is used to provide a DC supply to a 80-Ω load. If the AC source voltage is 32 V rms, find the peak and average current in the load. Assume an ideal diode.

8.37 The bridge rectifier in Figure P.8.37 is driven by a sinusoidal voltage source $v_s(t) = 6 \sin(314t)$ V. Redraw this figure to show that it is functionally identical to Figure 8.38. Determine the average and peak forward current through each diode when $R_o = 200\ \Omega$. Assume ideal diodes.

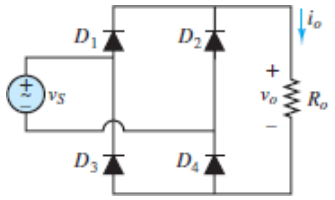


Figure 8.37 Full-wave rectifier current and voltage waveforms ($R = 1\ \Omega$)

8.38 In the full-wave power supply shown in Figure P8.38 the silicon diodes are 1N4001 with a rated peak reverse voltage of 25 V.

$$n = 0.05883$$
$$C = 80\ \mu\text{F} \qquad R_o = 1\ \text{k}\Omega$$
$$v_{\text{line}} = 170 \cos(377t)\ \text{V}$$

a. Determine the actual peak reverse voltage across each diode.

b. Explain why these diodes are or are not suitable for the specifications given.

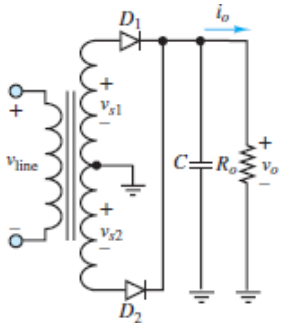


Figure P8.38

8.39 In the full-wave power supply shown in Figure P8.38,

$$n = 0.1$$
$$C = 80\ \mu\text{F} \qquad R_o = 1\ \text{k}\Omega$$
$$v_{\text{line}} = 170\cos(377t)\ \text{V}$$

The silicon diodes are 1N914 switching diodes (but used here for AC-DC conversion) with the following performance ratings:

$$P_{\text{max}} = 500\ \text{mW} \qquad \text{at } T = 25°\text{C}$$
$$V_{\text{pk-rev}} = 30\ \text{V}$$

The derating factor is 3 mW/°C for 25°C < $T \leq$ 125°C and 4 mW/°C for 125°C < $T \leq$ 175°C.

a. Determine the actual peak reverse voltage across each diode.

b. Are these diodes suitable for the specifications given? Explain.

8.40 Refer to Problem 8.38 and assume a load voltage waveform as shown in Figure P8.40. Also assume:

$$|i_o|_{avg} = 60\ \text{mA} \qquad |v_o|_{avg} = 5\ \text{V} \qquad |V_{\text{ripple}}| = 5\%$$
$$v_{\text{line}} = 170\cos(\omega t)\ \text{V} \qquad \omega = 377\ \text{rad/s}$$

Determine:

a. The turns ratio n.

b. The capacitor C.

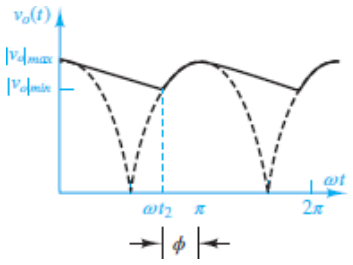


Figure P8.40

8.41 Refer to Problem 8.38. Assume:

$|i_o|_{avg} = 600$ mA $\quad |v_o|_{avg} = 50$ V

$V_r = 8\% = 4$ V

$v_{line} = 170\cos(\omega t)$ V $\quad \omega = 377$ rad/s

Determine:

a. The turns ratio n.

b. The capacitor C.

Page 566

8.42 Repeat Problem 8.37, using the diode offset model with $V_\gamma = 0.8$ V.

8.43 You have been asked to design a bridge rectifier for a power supply. A step-down transformer has already been chosen. It will supply 12 V rms to your rectifier. The bridge rectifier is shown in Figure P8.43.

a. If the diodes have an offset voltage of 0.6 V, sketch the input source voltage $v_S(t)$ and the output voltage $v_o(t)$, and state which diodes are on and which are off in the appropriate cycles of $v_{S2}(t)$. The frequency of the source is 60 Hz.

b. If R_o = 1,000 Ω and a filtering capacitor has a value of 8 μF, sketch the output voltage $v_o(t)$.

c. Repeat part b, with the capacitance equal to 100 μF.

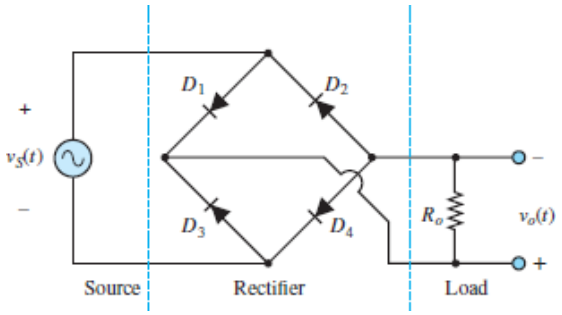


Figure P8.43

8.44 In the bridge rectifier of the power supply shown in Figure P8.44 the silicon diodes are 1N4001 with a rated peak reverse voltage of 50 V.

$v_{line} = 170\cos(377t)$ V

$n = 0.2941$

$C = 700\ \mu$F $\quad R_o = 2.5$ kΩ

a. Determine the actual peak reverse voltage across each diode.

b. Are these diodes suitable for the specifications given? Explain.

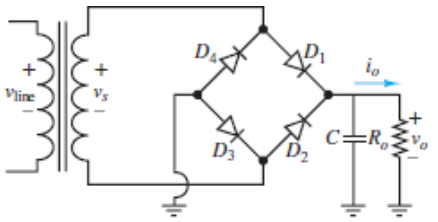


Figure P8.44

8.45 Refer to Problem 8.44. Assume the diodes have a rated peak reverse voltage of 10 V and:

$$v_{line} = 156\cos(377t)\text{ V}$$
$$n = 0.04231 \qquad V_r = 0.2\text{ V}$$
$$|i_o|_{avg} = 2.5\text{ mA} \qquad |v_o|_{avg} = 5.1\text{ V}$$

a. Determine the actual peak reverse voltage across the diodes.

b. Explain why these diodes are or are not suitable for the specifications given.

8.46 Refer to Problem 8.44. Assume:

$$|i_o|_{avg} = 650\text{ mA} \qquad |v_o|_{avg} = 10\text{ V}$$
$$V_r = 1\text{ V} \qquad \omega = 377\text{ rad/s}$$
$$v_{line} = 170\cos(\omega t)\text{ V} \qquad \phi = 23.66°$$

Determine the value of the average and peak current through each diode.

8.47 Repeat Problem 8.37, using the piecewise linear diode model with $V_\gamma = 0.8$ V and resistance $R_D = 25\ \Omega$.

8.48 Refer to Problem 8.44. Assume:

$$|i_o|_{avg} = 250\text{ mA} \qquad |v_o|_{avg} = 10\text{ V}$$
$$V_r = 2.4\text{ V} \qquad \omega = 377\text{ rad/s}$$
$$v_{line} = 156\cos(\omega t)\text{ V}$$

Determine:

a. The turns ratio n.

b. The capacitor C.

Section 8.6: Zener Diodes and Voltage Regulation

8.49 The diode shown in Figure P8.49 has a piecewise linear characteristic that passes through the points (−10 V, −5 μA), (0, 0), (0.5 V, 5 mA), and (1 V, 50 mA). Determine the piecewise linear model, and using that model, solve for i and v.

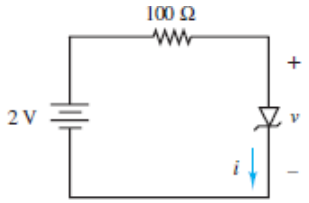


Figure P8.49

8.50 In the circuit shown in Figure P8.50, R must maintain the Zener diode within its specified limits. Page 567If $V_{batt} = 15 \pm 3$ V, $R_o = 1{,}000\ \Omega$, $V_z = 5$ V, 4 mA $\leq I_z \leq$ 18 mA, determine the minimum and maximum values of R that can be used.

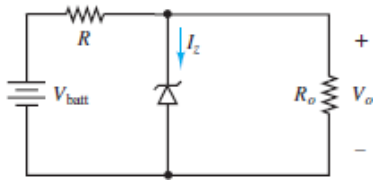


Figure P8.50

8.51 Determine the minimum value and the maximum value that the series resistor may have in a regulator circuit whose output voltage is to be 25 V, whose input voltage varies from 35 to 40 V, and whose maximum load current is 75 mA. The Zener diode used in this circuit has a maximum current rating of 250 mA.

8.52 The i-v characteristic of a semiconductor diode designed to operate in the Zener breakdown region is shown in Figure P8.52. The Zener or breakdown region extends from the knee of the curve, located here at $v_D = -3$ V and $i_D = -10$ mA, to a maximum rated current equal to −80 mA. The test point is $v_D = -5$ V and $i_D = -32$ mA. Determine the Zener resistance and Zener voltage of the diode.

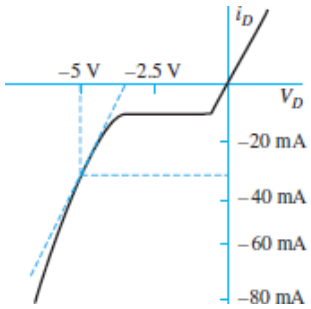


Figure P8.52

8.53 The Zener diode in the simple voltage regulator shown in Figure P8.53 is a 1N5231B. The source voltage is obtained from a DC power supply. It has DC and ripple components:

$$v_S = V_S + v_r$$

where:

$$V_S = 20\text{ V} \quad |v_r| = 250\text{ mV}$$
$$R = 220\ \Omega \quad |i_o|_{avg} = 65\text{ mA} \quad |v_o|_{avg} = 5.1\text{ V}$$
$$V_z = 5.1\text{ V} \quad r_z = 17\ \Omega \quad P_{rated} = 0.5\text{ W}$$
$$|i_z|_{min} = 10\text{ mA}$$

Determine the maximum rated current the diode can handle without exceeding its power limitation.

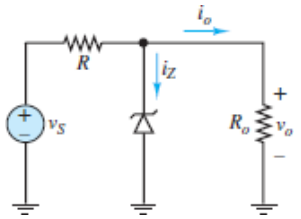


Figure P8.53

8.54 Repeat Problem 8.53 for the following specifications:

$$V_z = 12\text{ V} \quad r_z = 11.5\ \Omega \quad P_{rated} = 400\text{ mW}$$

At the knee of the reverse-biased Zener diode curve (see Figure P8.52):

$$I_{zk} = 12\text{ mA} \quad r_{zk} = 700\ \Omega$$

8.55 In the simple voltage regulator shown in Figure P8.53, R must maintain the Zener diode current within its specified limits for all values of the source voltage, load current, and Zener diode voltage. Determine the minimum and maximum values of R that can be used.

$$V_z = 5\text{ V} \pm 10\% \qquad r_z = 15\,\Omega$$
$$|i_z|_{\min} = 3.5\text{ mA} \qquad |i_z|_{\max} = 65\text{ mA}$$
$$|v_S| = 12 \pm 3\text{ V} \qquad |i_o| = 70 \pm 20\text{ mA}$$

8.56 Repeat Problem 8.55 for the following specifications:

$$V_z = 12\text{ V} \pm 10\% \qquad r_z = 9\,\Omega$$
$$|i_z|_{\min} = 3.25\text{ mA} \qquad |i_z|_{\max} = 80\text{ mA}$$
$$v_S = 25 \pm 1.5\text{ V}$$
$$|i_o| = 31.5 \pm 21.5\text{ mA}$$

8.57 In the circuit shown in Figure P8.57, compute the diode currents. Let V_{cc} = 24 V, I_o = 5 mA, R_1 = 1 kΩ, V_{dd} = 6 V, $V_{z1} = V_{z2}$ = 5 V, R_2 = 3 kΩ.

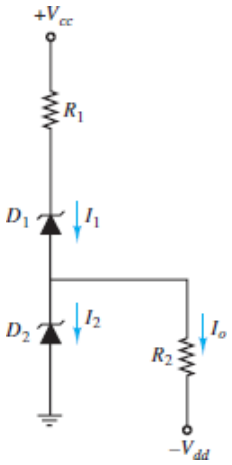


Figure P8.57

Page 568

8.58 In the circuit shown in Figure P8.58, compute the currents I_1 and I_2. Let V_{cc} = 18 V, V_{dd} = 24 V, V_{z1} = 7.5 V, V_{z2} = 5 V, R_1 = 5 kΩ, R_2 = 2 kΩ.

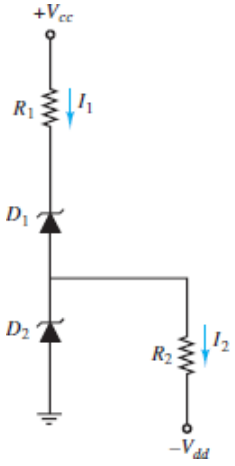


Figure P8.58

8.59 The Zener regulator shown in Figure P8.59 holds the load voltage at V_o = 14 V. Find the range of load resistance R_o for which regulation can be obtained if the Zener diode is rated at 14 V, 5 W.

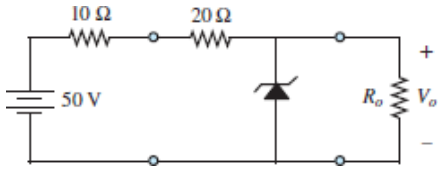


Figure P8.59

8.60 A Zener diode ideal *i-v* characteristic is shown in Figure P8.60(a). Given a Zener voltage, V_Z of 7.7 V, find the output voltage V_o for the circuit of Figure P8.60(b) if V_S is

a. 12 V

b. 20 V

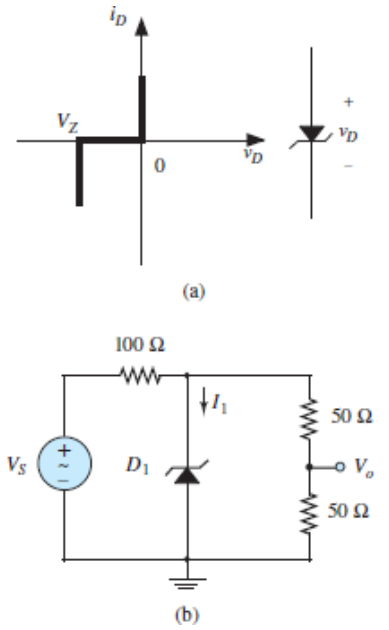


Figure P8.60

Section 8.7: Photodiodes

8.61 For the LED circuit of Example 8.13, determine the LED power consumption if the LED consumes 20 mA at the same voltage. How much power is required of the source?

8.62 For the LED circuit of Example 8.13, determine the LED power consumption if the LED consumes 30 mA and the diode voltage is 1.5 V. How much power is required of the source?

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]Semiconductors can also be made of more than one element, in which case the elements are not necessarily from group IV.

[2]Another reported relation [A.B. Sproul and M.A. Green, *J. Appl. Phys.* 70, 846 (1991)] is $n_i \propto T^2 e^{-E_G/2kT}$ with a value at 300 K of approximately 1.0×10^{10} carriers/cm^3.

[3]The group number system used here is somewhat antiquated. Under the current international system, groups III, IV, and V were renumbered as groups 13, 14, and

15. However, the older number system has value in this context since it represents the number of valence electrons.

[4]The positive terminal of a battery is referred to as the cathode because internally it is the source of negative ions traveling toward the negative terminal.

C H A P T E R

9

BIPOLAR JUNCTION TRANSISTORS: OPERATION, CIRCUIT MODELS AND APPLICATIONS

Over the last half-century, transistor technology has revolutionized the manner in which power and information are transmitted and utilized within our society. The impact of this technology is difficult to overstate, and examples of it are ubiquitous. Moreover, the technology and the products that depend upon it continue to develop at an exponential rate. It is astounding to consider that the first Macintosh personal computer was introduced by Apple Computer Co. in January 1984 with 64 kB of ROM, 128 kB of RAM, a motherboard running at 8 MHz, a display with 384 × 256 *pixel* resolution, all for the modest price of $2,495, which is roughly equivalent to $6,038 in 2018. In the same year, IBM released its second-generation AT (advanced technology) personal computer, which featured the 16-bit, 6-MHz Intel 80286 microprocessor, a 20-MB hard

drive. Just 35 years later, the minimum recommended specifications of a modest desktop computer for college students typically include a 64-bit, 3.2-GHz quad core Page 570processor, a 1.3-GHz data bus with 8 GB of RAM, 500 GB of internal drive storage and a monitor resolution of 1600 × 900.

Of course, advances in analog and digital technology have not been limited to personal computers. In general, communication systems of all kinds have been revolutionized. Until 1983, interpersonal telecommunications were limited to land-line phone calls. The only asynchronous form of telecommunication was provided by analog telephone tape recorders and by letter and package carriers such as the U.S. Postal Service, UPS, and FedEx. While these services continue to play an important role in our society, new forms of communication, particularly real-time asynchronous communications, have exploded. Today we transmit, exchange, and broadcast digital images, video, text, and voice using handheld and wearable mobile devices. It is not unreasonable to describe these smartphones as pocket-sized supercomputers. According to the Pew Research Center's Internet & American Life Project, as of May 2011, roughly 35 percent of American adults owned a smartphone of one type or another. Today that number has risen to 81 percent.

Fundamentally, all this progress has relied on advances in transistor technology. Given the broad impact of this technology, it would seem essential that engineers of all stripes possess a basic understanding of transistors and how they are used to produce the two building blocks of all communication and power devices. These two building blocks are the **switch** and the **amplifier**. Chapters 9 and 10 are dedicated to revealing how transistors are utilized to produce various types of switches and amplifiers. Chapter 9 focuses on a family of transistors known as **bipolar junction transistors** (BJTs). The underlying physics is discussed in sufficient detail to provide a comfortable basis for understanding the three modes of BJT operation. Practical examples are provided to illustrate important BJT circuits and their analysis using linear circuit models.

Learning Objectives

Students will learn to...

1. Understand the basic principles of amplification and switching. *Section 9.1*.
2. Understand the physical operation of bipolar junction transistors determine the operating point of a bipolar transistor circuit. *Section 9.2*.
3. Understand the large-signal model of the bipolar junction transistor and apply it to simple amplifier circuits. *Section 9.3*.
4. Select the operating point of a bipolar junction transistor circuit; understand the principle of small-signal amplifiers. *Section 9.4*.
5. Understand the operation of a bipolar junction transistor as a switch and analyze basic analog and digital gate circuits. *Section 9.5*.

9.1 AMPLIFIERS AND SWITCHES

A transistor is a three-terminal semiconductor device that can perform two functions that are fundamental to the design of electronic circuits: **amplification** and **switching**. Amplification consists of using an external power source to produce a scaled reproduction of a signal. Switching consists of using a relatively small input current or voltage to control a larger output current or voltage.

Page 571

Four different linear amplifier models are shown in Figure 9.1. Controlled voltage and current sources generate an output proportional to an input current or voltage; each proportionality constant A_i, A_v, G_m and R_m is called the internal *gain* of the transistor. (The internal gain G_m is a transconductance with units of A/V; it is not a dimensionless gain G as defined in previous chapters.) Bipolar junction transistor (BJTs) are well-modeled as current-controlled devices.[1]

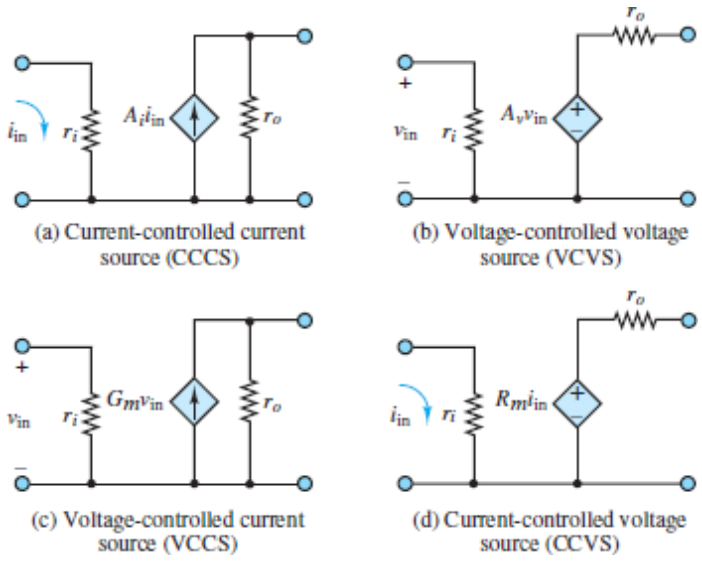


Figure 9.1 Controlled-source models of linear amplifiers

Transistors are also operated in a nonlinear mode, as voltage- or current-controlled switches. Figure 9.2 depicts the idealized operation of the transistor as a switch, suggesting that the switch is closed (on) whenever a control voltage or current is greater than zero and is open (off) otherwise. More realistic conditions on transistors acting in a switch mode are discussed later in this chapter and Chapter 10.

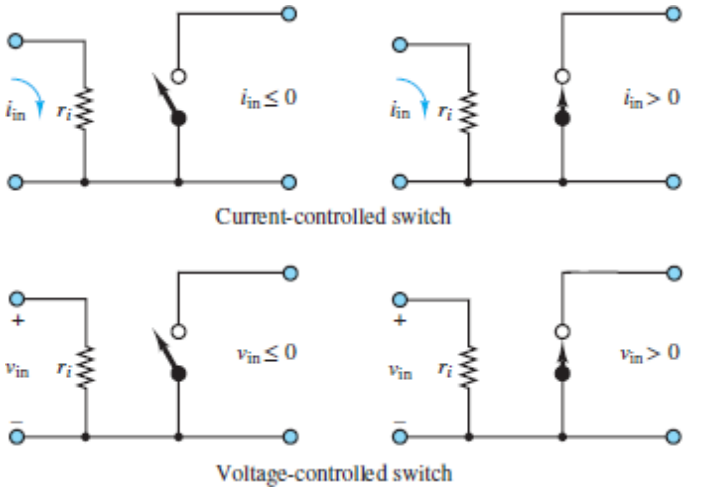


Figure 9.2 Models of ideal transistor switches

EXAMPLE 9.1 Model of Linear Amplifier

Problem

Determine the voltage gain of the amplifier circuit model shown in Figure 9.3.

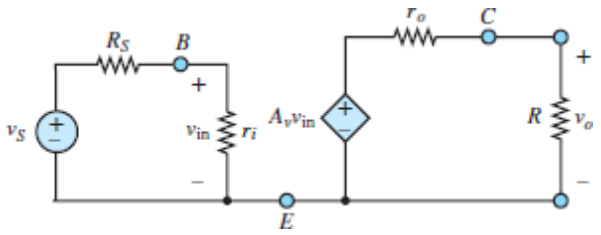


Figure 9.3

Solution

Known Quantities: Amplifier internal input and output resistances r_i and r_o; amplifier internal gain A_v; source and load resistances R_S and R.

Find: $G = \frac{v_o}{v_S}$

Analysis: Apply voltage division to determine v_{in}:

$$v_{in} = \frac{r_i}{r_i + R_S} v_S$$

Then, the output of the controlled voltage source is

$$A_v v_{in} = A_v \frac{r_i}{r_i + R_S} v_S$$

and the output voltage can also be found using voltage division:

$$v_o = A_v \frac{r_i}{r_i + R_S} v_S \times \frac{R}{r_o + R}$$

Finally, the amplifier voltage gain can be computed:

$$G = \frac{v_o}{v_S} = A_v \frac{r_i}{r_i + R_S} \times \frac{R}{r_o + R}$$

Comments: Note that the voltage gain computed above is always less than the transistor internal voltage gain A_v . One can easily show that if $r_i \gg R_S$ and $r_o \ll R$, then $G \approx A_v$. In general, the amplifier gain always depends on the ratio of the source R_S to input r_i resistances and the ratio of output r_o to load R resistances.

CHECK YOUR UNDERSTANDING

Repeat Example 9.1 for a current-controlled voltage source (CCVS) as shown in Figure 9.1(d). What is the amplifier voltage gain? Under what conditions would $G = R_m/R_S$?

Page 573

Repeat Example 9.1 for the current-controlled current source (CCCS) of Figure 9.1(a). What is the amplifier voltage gain?

Repeat Example 9.1 for the voltage-controlled current source (VCCS) of Figure 9.1(c). What is the amplifier voltage gain?

Answers: $G = R_m \frac{1}{r_i + R_S} \frac{R}{r_o + R}$ $r_i \to 0,\ r_o \to 0;\ G = \frac{1}{r_i + R_S} \frac{r_o R}{r_o + R} A_i$; $G = G_m \frac{r_i}{r_i + R_S} \frac{r_o R}{r_o + R}$

9.2 THE BIPOLAR JUNCTION TRANSISTOR (BJT)

A BJT is formed by joining three sections of alternating *p*- and *n*-type material. An *npn* transistor is a BJT with a thin, lightly doped *p*-type **base** region sandwiched between a heavily doped *n*-type **emitter** region and a large, lightly doped *n*-type **collector** region. The BJT counterpart to the *npn* is the *pnp* transistor, which utilizes the same doping scheme except that the *n* and *p* regions are swapped with respect to the *npn*. In both of these BJT types, the heavily doped emitter region is often labeled n^+ or p^+ to distinguish it from the lightly doped collector. Figure 9.4 illustrates the construction, symbols, and nomenclature for the two types of BJTs. Notice that there are the two *pn* junctions in a BJT: the **emitter-base junction** (EBJ) and the **collector-base junction** (CBJ). The operating mode of a BJT depends upon whether these junctions are reverse- or forward-biased, as indicated in Table 9.1.

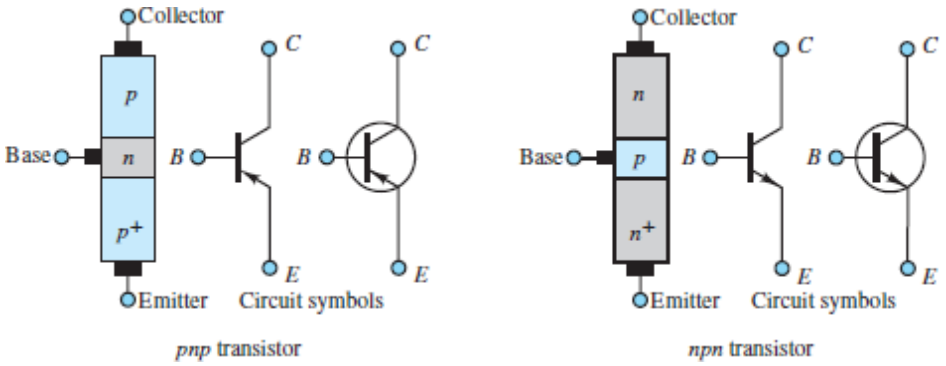


Figure 9.4 Bipolar junction transistors

Table 9.1 BJT operating modes

Mode	EBJ	CBJ	Application
Cutoff	Reverse-biased	Reverse-biased	Open switch
Active	Forward-biased	Reverse-biased	Amplifier
Saturation	Forward-biased	Forward-biased	Closed switch

Page 574

Although the construction of a BJT results in two opposing *pn* junctions, it is important to avoid modeling a BJT as two identical but opposing diodes. The EBJ always behaves as a true diode; however, because of the thin base region and the lightly doped collector region, the CBJ does not.

Figure 9.5 depicts the basic geometry of a cross section of a BJT. The base region is shown much thicker (compared to the emitter and collector) for the sake of clarity. There are two key points to note from the figure: (1) the base is a very thin envelope around the emitter, and (2) the collector is much larger than the emitter and the base because it envelopes both and is itself relatively thick compared to the emitter. The result of this geometry is that the collector can receive large numbers of mobile charge carriers without any significant impact upon its density of charge carriers.

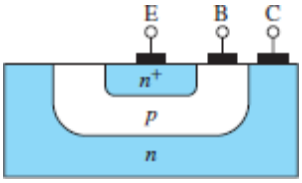


Figure 9.5 Cross section of an *npn* transistor. Notice that the collector is much larger and much more lightly doped than the emitter. However, the base is, in fact, very thin compared to the emitter and collector.

Cutoff Mode (EBJ Reverse-Biased; CBJ Reverse-Biased)

When both *pn* junctions are reverse-biased, no current is present across either junction and the path from collector to emitter can be approximated as an open-circuit. In fact, small reverse currents due to minority carriers are present across the junctions, but for most practical applications these reverse currents are negligible. In silicon-based BJTs, the offset voltage for the EBJ is the same V_γ presented in Chapter 8 for single silicon diodes, where $V_\gamma \approx 0.6$ V. Thus, in cutoff mode, when $v_{BE} < V_\gamma$, the transistor acts as a switch in its off (open-circuit) condition.

Active Mode (EBJ Forward-Biased; CBJ Reverse-Biased)

Figure 9.6 shows a Norton source connected across the base and emitter terminals of an *npn* transistor and the resulting *i*-*v* characteristic of its EBJ.

Notice that $i_B \approx 0$ when $v_{BE} \leq V_\gamma$, which is cutoff mode. However, when the EBJ is forward-biased such that $v_{BE} \leq V_\gamma$, current is conducted as in a typical diode. Majority carriers in the emitter and base *drift* across the EBJ under the influence of the forward-bias voltage in excess of the potential barrier of the depletion region. However, since the emitter is heavily doped while the base is lightly doped, the current I_E through the EBJ (see Figure 9.7) is dominated by the majority carriers from the emitter.

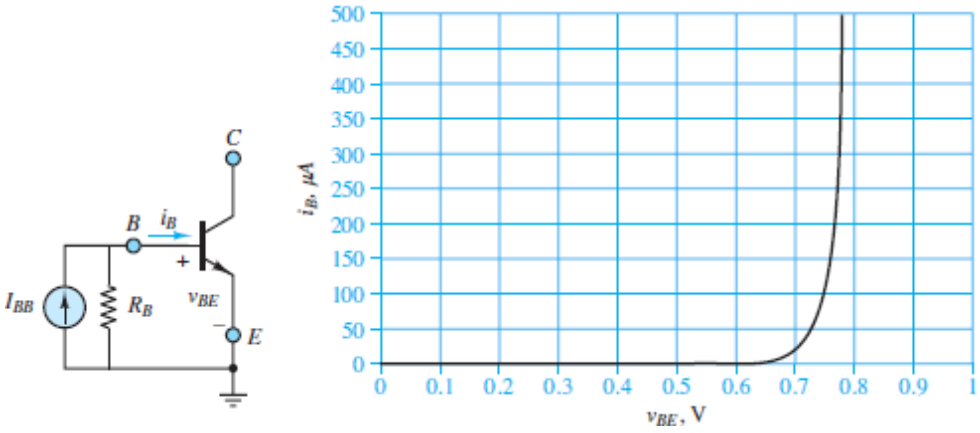


Figure 9.6 The *i-v* characteristic of the emitter-base junction of a typical *npn* transistor

Page 575

The *i-v* characteristics of the EBJ for *npn* and *pnp* transistors are identical except that the abscissas are v_{BE} and v_{EB}, respectively. The discussion below is based upon the behavior of an *npn* transistor; however, the behavior of a *pnp* transistor is completely analogous to that of an *npn* transistor, except that positive and negative charge carriers are interchanged and the EBJ is forward-biased from emitter to base rather than from base to emitter.

The behavior of a *pnp* transistor is completely analogous to that of an *npn* transistor, except that positive and negative charge carriers are

interchanged and the EBJ is forward-biased from emitter to base rather than from base to emitter.

For an *npn* BJT, the majority carriers in the emitter are free electrons while the majority carriers in the base are holes, as indicated in Figure 9.7. Some of these free electrons recombine with holes in the base; however, since the base is lightly doped, most of these electrons remain mobile *minority carriers* in the *p*-type base. As these mobile electrons cross the EBJ, their growing concentration in the base causes them to *diffuse* toward the CBJ. The equilibrium concentration of these mobile electrons throughout the base region is a maximum at the EBJ and is given by:

$$(n_p)_{\max} = (n_p)_o(e^{v_{BE}/V_T} - 1) \tag{9.1}$$

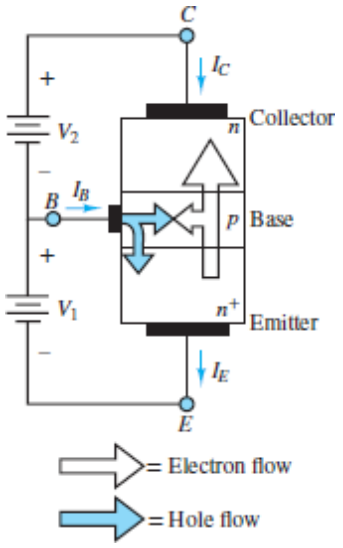


Figure 9.7 Flow of emitter electrons into the collector in an *npn* transistor

where v_{BE} is the forward-bias voltage from base to emitter and $(n_p)_o$ is the thermal equilibrium concentration of electrons in the base. Since the base is very thin, the equilibrium concentration *gradient* across the base is nearly linear, as depicted in Figure 9.8, such that the electron diffusion *rate* from the EBJ to the CBJ can be approximated as:

$$\frac{A q_e D_n (n_p)_{\max}}{W} \tag{9.2}$$

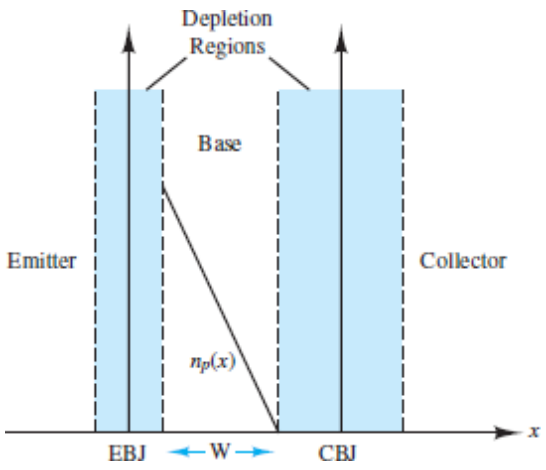


Figure 9.8 Equilibrium concentration gradient of free electrons in the *p*-type base of a forward-biased EBJ of an *npn* transistor

where A is the cross-sectional area of the EBJ, W is the width of the base (not including the width of the two bounding depletion regions), and D_n is the diffusivity of electrons in the base. It is important to note that this electron diffusion rate is Page 576temperature dependent and that it represents a diffusion *current* directed from the CBJ to the EBJ because of the convention that the direction of positive current is the direction of flow of positive charge carriers. Once these diffusing electrons reach the CBJ they are swept into the collector by the reverse-bias voltage across the CBJ. Thus, the **collector current** i_C is

$$\begin{aligned} i_C &= \frac{A q_e D_n (n_p)_o}{W} (e^{v_{BE}/V_T} - 1) \\ &= \frac{A q_e D_n n_i^2}{W N_A} (e^{v_{BE}/V_T} - 1) \\ &= I_S (e^{v_{BE}/V_T} - 1) \end{aligned} \tag{9.3}$$

where N_A is the doping concentration of holes in the base and I_S is known as the **scale current** because it scales with the cross-sectional area A of the EBJ. Typical values of I_S range from 10^{-12} A to 10^{-15} A.

The **base current** i_B (from base to emitter) is comprised of those majority carriers in the base (e.g., holes for an *npn* transistor) that traverse the EBJ. Some of these carriers recombine with the majority carriers in the emitter (e.g., electrons for an *npn* transistor); however, those majority carriers lost to recombination are replaced by additional majority carriers supplied by V_1. Because the concentration of these majority carriers is proportional to $i_B = \frac{i_C}{\beta} = \frac{i_C}{h_{FE}}$, the base current is proportional to the collector current i_C such that:

$$\begin{aligned} i_E &= I_{ES}(e^{v_{BE}/V_T} - 1) \\ &= i_C + i_B = \frac{\beta + 1}{\beta} i_C = \frac{i_C}{\alpha} \end{aligned} \tag{9.4}$$

where β is known as the forward **common-emitter current gain** with typical values ranging from 20 to 200. Although β can vary significantly from one transistor to another, most practical electronic devices only require that $\beta \gg 1$. Figure 9.7 depicts the flow of charge carriers from emitter to base to collector and from base to emitter, as discussed above, for an *npn* transistor. A BJT is a *bipolar* device because its current is comprised of both electrons and holes.[2]

The parameter β is not often found in a data sheet. Instead, the forward **DC** value of β is listed as h_{FE}, which is the **large-signal current gain**. A related parameter, h_{fe}, is the **small-signal current gain**.

Finally, to satisfy KCL, the **emitter current** i_E must be the sum of the collector and base currents and, therefore, must also be proportional to e^{v_{BE}/V_T}. Thus:

$$\begin{aligned} i_E &= I_{ES}(e^{v_{BE}/V_T} - 1) \\ &= i_C + i_B = \frac{\beta + 1}{\beta} i_C = \frac{i_C}{\alpha} \end{aligned} \tag{9.5}$$

where I_{ES} is the reverse **saturation current** and α is the **common-base current gain** with a typical value close to, but not exceeding, 1.

Saturation Mode (EBJ Forward-Biased; CBJ Forward-Biased)

A BJT remains in active mode as long as the CBJ is reverse-biased; that is, as long as $V_2 > 0$, electrons diffusing across the base will be swept away into the collector Page 577once they reach the CBJ. However, when the CBJ is forward-biased ($V_2 < 0$), these diffusing electrons are no longer swept away into the collector but instead accumulate at the CBJ such that the concentration of minority carrier electrons there is no longer zero. The magnitude of this concentration increases as V_2 decreases, such that the concentration gradient across the base decreases. The result is that the diffusion of minority carrier electrons across the base decreases; in other words, the collector current i_C decreases as the forward bias of the CBJ increases.

It is important to realize that as the concentration gradient across the base decreases and the rate of diffusion across the base decreases, the rate of increase of the concentration near the CBJ slows and the concentration gradient across the base approaches zero asymptotically. This asymptotic process expresses itself as an upper limit on the forward-bias voltage across the CBJ. Figure 9.9 defines three voltages across the terminals of an *npn* transistor. In saturation, the action of the transistor limits v_{CB} such that v_{CE} is always positive, although small. In fact, deep saturation mode is often best recognized as when the value of v_{CE} is approximately 0.2 V for a silicon-based BJT.

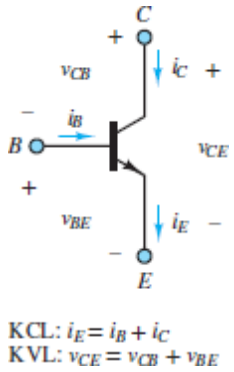


Figure 9.9 Definition of BJT voltages and currents

In saturation, the collector current is no longer proportional to the base current and the collector-emitter voltage v_{CE} for a silicon-based BJT is small (< 0.4 V). An increasing base current drives a BJT further into saturation, and v_{CE} approaches the saturation limit of $V_{CE\,\text{sat}} \approx 0.2$ V.

$$V_{CE\,\text{sat}} \approx 0.2 \text{ V} \qquad \text{Saturation limit} \tag{9.6}$$

Key BJT Characteristics

The voltages and currents shown in Figure 9.9 for an *npn* transistor are related by KCL and KVL.

$$v_{CE} = v_{CB} + v_{BE} \qquad \text{KVL} \tag{9.7}$$

$$i_E = i_C + i_B \qquad \text{KCL} \tag{9.8}$$

The BJT currents are temperature dependent because they are proportional to both n_i^2 and e^{v_{BE}/V_T}. These currents are also proportional to the cross-sectional area A of the EBJ and inversely proportional to the effective width W of the base.

The relationships between these voltages and currents are commonly represented by a graph of i_C versus v_{CE}, with i_B treated as a parameter. A typical example of such a graph is shown in Figure 9.10. The operating mode of a BJT is completely specified by these three variables. The three modes of operation are indicated in the figure. Cutoff and saturation modes occur when i_C and v_{CE} are very small, respectively.

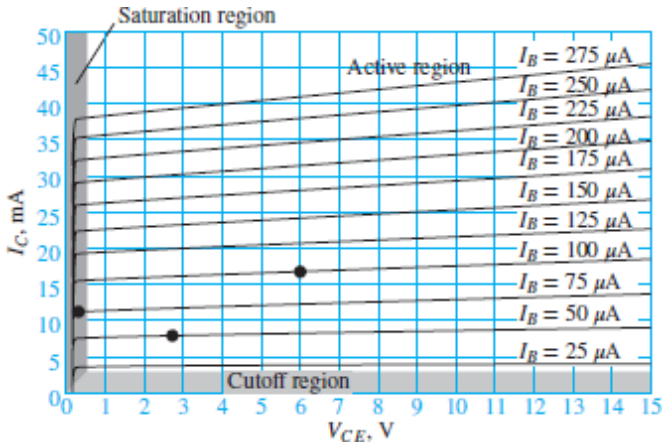


Figure 9.10 Typical characteristic lines of a BJT

For any fixed value of i_B, the slope of the transistor characteristic is very small in active mode. In the ideal case, this slope would be zero; however, the effective width of the base decreases with v_{CE} such that the concentration gradient of charge carriers in the base increases and, thus, the collector current increases as well. This increase in i_C with v_{CE} is known as the **Early effect** or **base-width modulation**.

Page 578

It is important to realize that the operating values of i_B, i_C, and v_{CE}, and the operating mode itself, are determined by the external circuitry attached to the BJT. An important objective of this chapter is to provide a method to design the external circuitry so as to dictate and control the operating mode of a BJT. To understand the development of such a method, it is essential to keep in mind the key characteristics of the cutoff, active, and saturation modes, which are the same for both *npn* and *pnp* transistors, and which are summarized in the box below.

Cutoff mode: Both the EBJ and CBJ are reverse-biased such that all three currents i_C, i_B, and i_E are approximately zero. In cutoff mode, a BJT

acts as an open switch between the collector and emitter.

Active mode: The EBJ is forward-biased while the CBJ is reverse-biased.

The BJT currents are related by:

$$i_C = \beta i_B \qquad i_C = \alpha i_E$$

In active mode, these currents are largely independent of v_{CB} and the BJT acts as a linear amplifier.

Saturation mode: Both the EBJ and CBJ are forward-biased such that $v_{BE} \approx 0.7$ V and $v_{CE} \approx 0.2$ V. The collector current i_C is highly sensitive to small changes in v_{CE}, and, since v_{CE} is small, i_C is largely determined by external circuitry attached to the collector terminal. In saturation mode, the BJT approximates a closed switch between the collector and emitter.

Determining the Operating Mode of a BJT

A few simple voltage measurements permit a quick determination of the state of a transistor. Consider, for example, an *npn* transistor placed in the circuit of Figure 9.11, where:

$$R_B = 40\ \text{k}\Omega \qquad R_C = 1\ \text{k}\Omega \qquad R_E = 161\ \Omega$$

Page 579

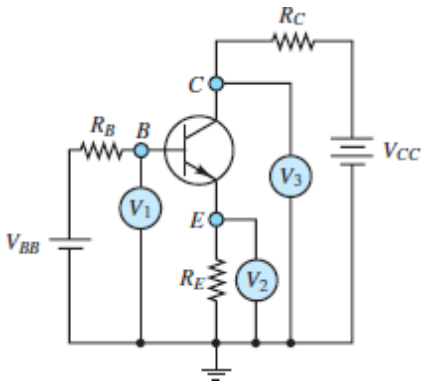


Figure 9.11 Determination of the operating mode of a BJT

and

$$V_{BB} = 4\,\text{V} \qquad V_{CC} = 12\,\text{V}$$

Assume that for these external parameter values the measured collector, emitter, and base terminal voltages are

$$V_B = V_1 = 2.0\,\text{V} \qquad V_E = V_2 = 1.3\,\text{V} \qquad V_C = V_3 = 4.0\,\text{V}$$

The method used in determining the state of a transistor is to assume an operating mode and then test the assumption against the known data. It is usually best to first assume cutoff mode and check whether the EBJ is reverse-biased. The voltage across the EBJ is

$$V_{BE} = V_B - V_E = 0.7\,\text{V}$$

Thus, the EBJ is forward-biased, not reverse-biased, and the transistor is not in cutoff mode.

One can next assume either active or saturation mode and test the assumption. For this example, assume saturation mode and test whether the CBJ is forward-biased. The voltage across the CBJ is

$$V_{BC} = V_B - V_C = -2.0\,\text{V}$$

Thus, the CBJ is reverse-biased and the transistor is in active mode. The same determination could be made by evaluating the voltage across the collector-emitter terminals.

$$V_{CE} = V_C - V_E = 2.7\,\text{V}$$

The requirement for saturation mode is $V_{CE} < 0.4$ V, which is clearly not satisfied.

Since the transistor is in active mode, it is possible to calculate the common- emitter current gain β. The base current is

$$I_B = \frac{V_{BB} - V_B}{R_B} = \frac{4 - 2}{40{,}000} = 50\ \mu\text{A}$$

The collector current is

$$I_C = \frac{V_{CC} - V_C}{R_C} = \frac{12 - 4}{1{,}000} = 8 \text{ mA}$$

Thus, the current amplification factor is

$$\frac{I_C}{I_B} = \beta = 160$$

Page 580

The operating point of the transistor in the given circuit can be located on a characteristic plot such as that in Figure 9.10. It is important to note that the operating mode of the transistor is determined by the attached circuitry. In this example, the values of V_B, V_C, and V_E were measured. However, for analytic problems, these values can be calculated using KCL, KVL, Ohm's law, and the known characteristics of the three possible modes of operation.

EXAMPLE 9.2 Determining the Operating Mode of a BJT

Problem

Determine the operating mode of the BJT in the circuit of Figure 9.11.

Solution

Known Quantities: Base, collector, and emitter voltages with respect to ground.

Find: Operating mode of the transistor.

Schematics, Diagrams, Circuits, and Given Data: $V_1 = V_B = 1.0$ V; $V_2 = V_E = 0.3$ V; $V_3 = V_C = 0.6$ V; $R_B = 40$ kΩ; $R_C = 1$ kΩ; $R_E \approx 26$ Ω.

Analysis: Compute V_{BE} and V_{BC} to determine the bias conditions of the EBJ and CBJ, which determine the mode of operation of the transistor.

$$V_{BE} = V_B - V_E = 0.7\,\text{V}$$
$$V_{BC} = V_B - V_C = 0.4\,\text{V}$$

Since both junctions are forward-biased, the transistor is in saturation mode. Also, notice that $V_{CE} = V_C - V_E = 0.3$ V is less than 0.4 V, which indicates that the BJT is operating near or in saturation.

The operating point of this transistor can be located in Figure 9.10 by calculating:

$$I_C = \frac{V_{CC} - V_C}{R_C} = \frac{12 - 0.6}{1{,}000} = 11.4\ \text{mA}$$

and

$$I_B = \frac{V_{BB} - V_B}{R_B} = \frac{4 - 1.0}{40{,}000} = 75.0\ \mu\text{A}$$

Notice that the operating point in Figure 9.10 is near the elbow in the $I_B = 75.0\ \mu$A curve at $V_{CE} = 0.3$ V.

Comments: KCL requires $I_E = I_C + I_B$. The latter sum is 11.475 mA, whereas I_E is 0.3 V/$R_E \approx 11.5$ mA. The difference between these two currents is due entirely to the approximate value of R_E. In fact, KCL is—as it must be—satisfied exactly.

It is important to notice that by only changing R_E from 26 to 161 Ω (as in the analysis described on the previous page), the operating mode of the transistor is changed from saturation mode to active mode.

CHECK YOUR UNDERSTANDING

Describe the operation of a *pnp* transistor in active mode by analogy with that of the *npn* transistor.

CHECK YOUR UNDERSTANDING

For the circuit of Figure 9.11, the voltmeter readings are $V_1 = 3.1$ V, $V_2 = 2.4$ V, and $V_3 = 2.7$ V. Determine the operating mode of the transistor. What is the value of R_E required to satisfy KCL? Assume $R_B = 40$ kΩ and $R_C = 1$ kΩ.

Answer: Saturation, $R_E \approx 257\ \Omega$

9.3 BJT LARGE-SIGNAL MODEL

The *i-v* characteristics of a BJT indicate that it acts as a current-controlled current source (CCCS) in the cutoff and active operating modes. In those two modes, the base current dictates the behavior of the BJT. These characteristics form part of a **large-signal model** for the BJT that describes its behavior in terms of the amplitudes of the base and collector currents. Like all models, the large-signal model does not account for every characteristic of a BJT. In particular, it does not account for the Early effect nor temperature effects. However, this model does provide a useful and simple starting point for the analysis of transistor circuits.

Large-Signal Model of the *npn* BJT

In cutoff mode, the *BE* junction is reverse-biased, the base and collector currents are approximately zero, and therefore the transistor acts as a *virtual* open-circuit. In practice, there is always a leakage current, denoted by I_{CEO}, through the collector, even when $V_{BE} = 0$ and $I_B = 0$.

In active mode, the *BE* junction is forward-biased, and the collector current is proportional to the base current, where the constant of proportionality is β.

$$I_C = \beta I_B \tag{9.9}$$

Since $\beta \gg 1$, this relationship indicates that the collector current is controlled by a relatively small base current.

Finally, in saturation mode, the base current is sufficiently large that the collector-emitter voltage V_{CE} reaches its saturation limit, and the collector current is no longer proportional to the base current. In fact, the collector-emitter pathway acts like a *virtual* short-circuit, except for the small potential drop $V_{CE\,\text{sat}} \approx 0.2\text{V}$.

All three of these operating modes are described by the simple circuit models shown in Figure 9.12. Each of these individual models approximates one of the three operating modes indicated in Figure 9.10. Notice that the large-signal model treats the forward-biased *BE* junction as an offset diode.

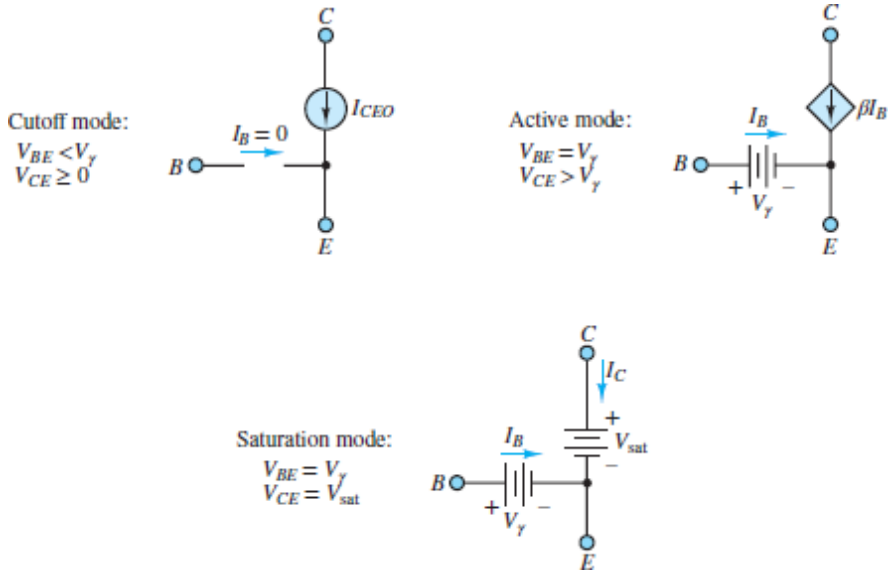


Figure 9.12 An *npn* BJT large-signal model

Selecting an Operating Point for a BJT

The family of curves shown for the collector *i-v* characteristic in Figure 9.10 reflects the dependence of the collector current on the base current. For each

value of the base current i_B there exists a corresponding i_C-v_{CE} curve. Thus, by selecting the base current and collector current (or collector-emitter voltage), an operating point Q for the transistor is determined. Q is defined in terms of the **quiescent** (or **idle**) **currents** and **voltages** that are present at the terminals of the device under DC conditions. The circuit of Figure 9.13 illustrates an ideal (not practical) **DC bias circuit**, used to set Page 582the operating point Q such that $V_{CE} \approx V_{CC}/2$. (A practical bias circuit is discussed later in this chapter.) In this illustration, Norton and Thévenin sources represent the linear networks seen by the base and collector terminals, respectively. The underlying principle is to pick R_C and R_B such that under quiescent DC conditions the BJT is maintained in active mode for all anticipated variations in I_B, I_C, and V_{CE} under operating (nonquiescent) conditions.

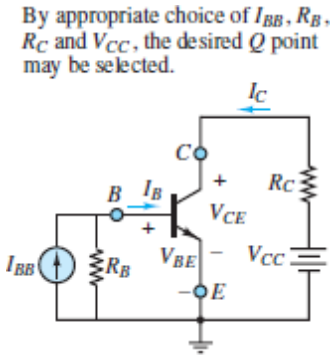


Figure 9.13 A simple ideal bias circuit for a BJT amplifier

KVL can be applied to yield the following equations:

$$I_B = I_{BB} - \frac{V_{BE}}{R_B} \tag{9.10}$$

and

$$V_{CE} = V_{CC} - I_C R_C \tag{9.11}$$

or

$$I_C = \frac{V_{CC} - V_{CE}}{R_C} \tag{9.12}$$

Note that equation 9.11 represents a *load line* for the source network of V_{CC} in series with R_C. When $V_{CE} = 0$, the collector current is $I_C = V_{CC}/R_C$; when $I_C = 0$, the collector-emitter voltage is $V_{CE} = V_{CC}$. These two conditions represent the virtual short- and open-circuit cases for the collector-emitter pathway, that is, the saturation and cutoff modes of the BJT, respectively. The load line can be superimposed upon the plot of BJT characteristics as shown in Figure 9.14. The slope of the load line is $-1/R_C$. The operating point Q is the intersection of the load line with one of the BJT characteristic lines. The particular characteristic line is determined by the base current I_B, as given by equation 9.10. The particular load line shown in Figure 9.14 assumes $V_{CC} = 15$ V, $V_{CC}/R_C = 40$ mA, which are the open-circuit voltage and short-circuit current, respectively, of the Thévenin source seen by the collector.

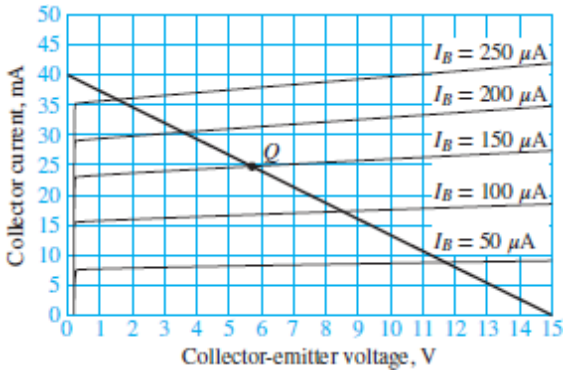


Figure 9.14 Load-line analysis of a simplified BJT amplifier
2

Once the operating point is established, the BJT is considered **biased** and prepared to operate as a linear amplifier. It is important to note that in circuit diagrams transistors are usually designated Q_1, Q_2, etc. The use of Q to denote transistors is related to the use of Q to denote an operating point, but the two uses serve two different purposes.

FOCUS ON MEASUREMENTS

Large-Signal Amplifier Design for a Diode Thermometer

Problem:

A diode can be used as the temperature transducer in an electronic thermometer (see the Focus on Measurements box, “Diode Thermometer” in Chapter 8). In this example, a diode element acts as a temperature transducer within the transistor amplifier circuit shown in Figure 9.15.

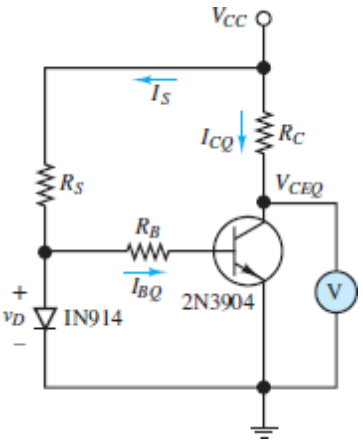


Figure 9.15 Large-signal amplifier for diode thermometer

Solution:

Known Quantities—Diode and transistor amplifier bias circuits; diode voltage versus temperature response.

Find—Resistor values and transistor output voltage versus temperature.

Schematics, Diagrams, Circuits, and Given Data—Figures 8.58, 9.10 and 9.15.

Assumptions—See discussion.

Analysis—The objective of this exercise is to design an amplifier that will output a linearly scaled reproduction of the diode voltage as it varies with temperature. The first design specification is that the amplifier be linear; that is, that the BJT remain in active Page 584mode throughout the temperature range $0 < T < 100°C$. A common approach to satisfying this specification is to establish the transistor nominal operating point so that $V_{CEQ} \approx V_{CC}/2$, where V_{CEQ} is the collector-emitter voltage v_{CE} at the midrange temperature $T_0 = 50°C$. This choice provides headroom and legroom above and below V_{CEQ} to allow v_{CE} to vary as the diode temperature varies without leaving active mode. Thus, for this example, an initial design specification is

$$v_{CE} \approx 6 \text{ V} \quad \text{at} \quad T_0 = 50°C$$

The second design specification is that the diode voltage remain a linear function of temperature during operation. Recall from Figure 8.58 that the diode voltage is a nearly linear function of temperature when the diode current is held constant, which can be accomplished by choosing $v_{CC} \gg v_D$ such that the voltage drop across R_S is relatively constant. It is assumed here that $i_S \gg i_B$, which can be accomplished by selecting $R_S \approx R_B$ since the voltage drop across R_B is the small difference between the two diode voltages v_D and v_{BE}. Thus, two more design specifications could be

$$V_{CC} \gg v_D \qquad \text{and} \qquad R_S \approx R_B$$

In addition, to maintain active mode it is essential to maintain $v_D > v_{BE}$ so that the base current is directed into the base of the NPN transistor. This result will also hold as long as $i_S \gg i_B$; that is, as long as these two specifications are met. Also, the assumption that $i_S \gg i_B$ implies that $i_D \approx i_S$.

Assuming the BJT is maintained in active mode, the base-emitter voltage v_{BE} can be assumed to be significantly larger than V_γ . Since $V_\gamma \approx$ 0.65 V a minimum design specification could be $v_D > V_\gamma + 0.5$ V at the high end of the temperature range.

$$v_D > V_\gamma + 0.5 \text{ V} \qquad \text{at} \qquad T = 100°\text{C}$$

Figure 8.58 shows the diode voltage versus temperature for a diode current $i_D \approx 1.5$ mA. At $T = 100°$C the diode voltage is approximately 0.6 V, which does not satisfy the minimum specification on v_D. However, as with all diodes, the forward-biased diode voltage v_D increases as the diode current i_D increases. Figure 9.16(a) shows the diode voltage as a function of temperature when the diode current is approximately 100 mA. Notice that the diode voltage at $T = 100°$C is roughly 0.73 V, which satisfies $v_D > V_\gamma +$ 0.5 V. Thus, another design specification could be

$$i_D \approx 100 \text{ mA}$$

Consider the specifications $V_{CC} \gg v_D$ and $i_D \approx 100$ mA. The implication that $i_D \approx i_S$ also implies $i_S \approx 100$ mA. If $V_{CC} = 12$ V and $v_D \approx 0.8$ V, then $R_S \approx (12 - 0.8)$ V/100 mA = 112 Ω. The nearest standard resistor values are 100 Ω and 120 Ω. A slightly larger value for i_S is preferable to a slightly lower value, so choose $R_S = 100$. Consequently, $R_B = 100$ Ω also to satisfy the $R_S \approx R_B$ specification.

$$R_S = R_B = 100 \ \Omega$$

To produce a design specification for R_C it is necessary to estimate the collector current i_C when $v_{CE} \approx 6$ V at $T_0 = 50°C$. Ohm's law could then be applied to R_C to yield:

$$R_C = \frac{V_{CC} - v_{CE}}{i_C}$$

Page 585

The *i-v* characteristic curves for a typical BJT shown in Figure 9.14 can be used to estimate i_C for various values of i_B given that $v_{CE} = 6$ V. Ohm's law applied to R_B is

$$i_B = \frac{v_D - v_{BE}}{R_B}$$

Figure 9.16(a) indicates that $v_D \approx 0.78$ V when $i_D \approx 100$ mA at $T_0 = 50°C$. But what about v_{BE} at the same operating point? No additional information is available for this particular 2N3904 NPN transistor. However, a reasonable assumption is that v_{BE} increases and decreases with v_D and that the difference between them is quite small. Thus, assume that $v_D - v_{BE} \approx 0.01$ V to make an initial estimate of i_B, i_C and finally R_C. Simulations and/or experiments can be performed to evaluate and iterate on these initial estimates to improve the performance of the amplifier. Such iterations are a common aspect of any design process.

Assuming $v_D - v_{BE} \approx 0.01$ V and $R_B = 100\ \Omega$ the initial estimate is $i_B \approx 100\ \mu$A. Using this value and Figure 9.14, the initial estimate of the collector current is $i_C \approx 17$ mA and thus $R_C \approx 353\ \Omega$. The nearest standard resistor values are 330 Ω and 390 Ω. A larger R_C results in a smaller i_C and, thus, a smaller i_B in active mode, which supports the implication that $i_D \approx i_S$. Thus, choose $R_C = 390\ \Omega$ for the initial amplifier design.

$$R_C = 390\ \Omega$$

The amplifier output v_{CE} for these parameter values over the temperature range $0 < T < 100°C$ is shown in Figure 9.16(b).

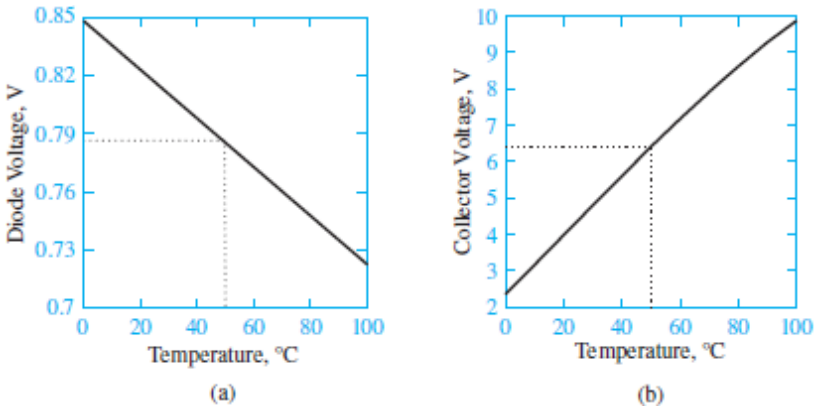


Figure 9.16 (a) Diode voltage temperature dependence; (b) Amplifier output

Notice the greatly increased sensitivity of the output voltage to changes in temperature when compared to that of the diode voltage. Also notice that the slope of $v_{CE}(T)$ is positive. This result is due to the fact that a *common-emitter amplifier* inverts the slope of the collector-emitter voltage relative to the slope of the base voltage.

The design parameters chosen here are not the only viable options and are unlikely optimal for this particular application. Other combinations of R_S, R_B and R_C yield useful results. The efficacy of the choices depends upon the objectives for any particular project. The process outlined here is intended to illustrate some of the issues that should be considered when designing a BJT amplifier. Typically, an actual design would involve iterations on an initial design to improve performance. A practical amplifier Page 586design would also include coupling and bypass capacitors to address other common amplifier design issues. It is also true that other amplifier configurations can be used to accomplish the same task.

It is a worthwhile exercise to compare the results of the final design to the assumptions that led to that design. For the values of R_S, R_B and R_C selected here, the collector current is 14.4 mA at 50°C, which is somewhat lower than the 17 mA used to determine R_C. Also, the actual base current is approximately 75 μA at 50°C, which is lower than the 100 μA design estimate. However, the important question is whether the design performance satisfies the objectives. It does.

Comments: —In active mode, $i_C = \beta i_B$. However, in practice, the value of β is not used as a design specification because its value can vary greatly even within the same type of transistor. Instead, amplifier designs rely on $\beta \gg 1$, which is a feature of all BJTs.

The results shown in Figure 9.16(a)–(b) were generated using the TINACloud circuit simulator. Experimental results may vary somewhat. To reproduce the simulation results it is important that only the diode temperature be permitted to vary. If the BJT and diode temperatures are allowed to vary together the amplifier output voltage is greatly compromised.

EXAMPLE 9.3 LED Driver

Problem

Design a transistor switch to control an LED as shown in Figure 9.17. The LED is required to turn on and off with the on/off signal from a digital output port of a microcontroller that is in series with the transistor base.

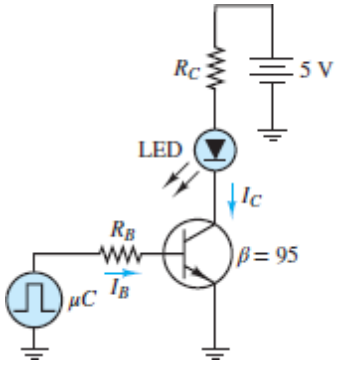


Figure 9.17 LED driver circuit

Solution

Known Quantities: Microcontroller output resistance and output signal voltage and current levels; LED offset voltage, required current, and power rating; BJT current gain and base-emitter junction offset voltage.

Find: (a) Collector resistance R_C such that the transistor is in saturation when the microcontroller outputs 5 V; (b) power dissipated by LED.

Schematics, Diagrams, Circuits, and Given Data:

Microcontroller: output resistance = R_B = 1 kΩ; V_{ON} = 5 V; V_{OFF} = 0 V.

Transistor: VCC = 5V; V_γ = 0.7V; β = 95; $V_{CE\ \text{sat}}$ = 0.2V.

LED: V_{LED} = 1.4 V; I_{LED} = 30 mA; P_{max} = 100 mW.

Assumptions: Use the large-signal models for cutoff and saturation as shown in Figure 9.12. In saturation, $V_{CE} \approx V_{CE\ \text{sat}}$ = 0.2 V.

Analysis: When the microcontroller output voltage is zero, the BJT is in cutoff mode since the base current is zero. When the microcontroller output voltage is V_{ON} = 5 V, the transistor should be in saturation mode so that the LED sees a virtual short-circuit from collector to emitter. Figure 9.18(a) depicts the equivalent base-emitter circuit when the microcontroller output voltage is V_{ON} = 5 V. Figure 9.18(b) depicts the collector circuit, and Figure 9.18(c), Page 587the same collector circuit with the large-signal model for the transistor in place of the BJT. Apply KVL to obtain:

$$V_{CC} = R_C I_C + V_{\text{LED}} + V_{CE\,\text{sat}}$$

or

$$R_C = \frac{V_{CC} - V_{\text{LED}} - V_{CE\,\text{sat}}}{I_C} = \frac{3.4}{I_C}$$

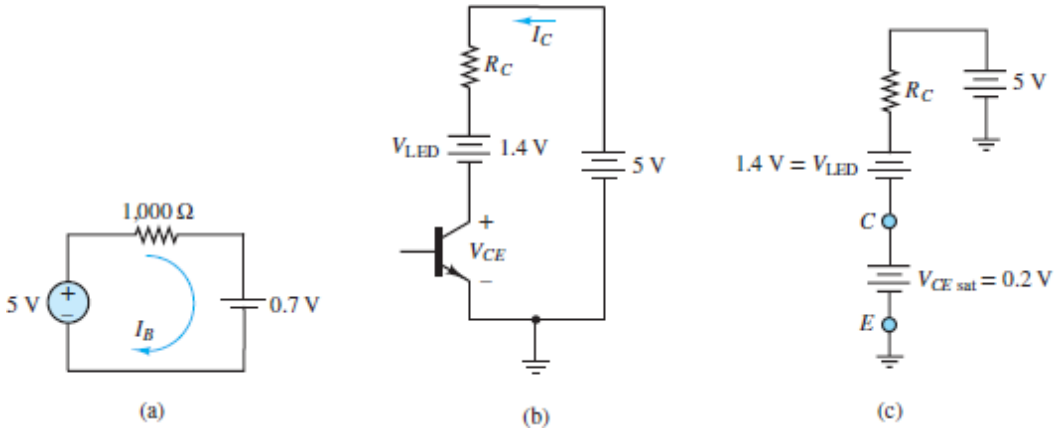


Figure 9.18 (a) *BE* circuit for LED driver; (b) equivalent collector circuit of LED driver, assuming that the BJT is in the linear active mode; (c) LED driver equivalent collector circuit, assuming that the BJT is in saturation mode

A typical LED requires at least 15 mA to be on. In this example, the LED current is specified as 30 mA to ensure that the LED is reasonably bright when on. The collector resistance R_C needed to provide this current is ≈ 113 Ω.

To confirm that the transistor is in saturation when the microcontroller voltage is 5 V, the ratio I_C/I_B should be less than β. For the given specifications, the base current is

$$I_B = \frac{V_{ON} - V_\gamma}{R_B} = \frac{4.3}{1{,}000} = 4.3 \text{ mA}$$

Thus:

$$\frac{I_C}{I_B} \approx 7$$

In active mode, the ratio $I_C/I_B = \beta = 95$. For sufficiently large values of the base current, the transistor leaves active mode and enters saturation. In saturation, the ratio I_C/I_B is no longer constant and is always less than β. Clearly, this condition is satisfied when the microcontroller output is on. For any particular transistor, the value of β can be significantly different from its typical value given in a generic data sheet. Thus, in practice, it is a good idea

to make sure that $I_C/I_B \ll \beta_{typ}$. In this example, $7 \ll 95$ such that it is reasonably certain that the transistor will be in saturation for the design specification of $R_C \approx 113\ \Omega$.

The power dissipated by the LED is

$$P_{LED} = V_{LED} I_C = 1.4 \times 0.3 = 42\ \text{mW} < 100\ \text{mW}$$

Since the power rating of the LED has not been exceeded, the design is complete.

Comments: The large-signal model of the BJT is easy to apply because the *BE* and *CE* junctions are approximated as short-circuits in series with an independent voltage source. Remember to first check the operating mode of the transistor by assuming a mode and then verifying that the assumption is not contradicted by the resulting voltages across the EBJ, CBJ and CEJ.

Page 588

EXAMPLE 9.4 Simple BJT Battery Charger (Current Source)

Problem

Design a constant-current battery charging circuit by selecting values of V_{CC}, R_1, R_2 (a potentiometer) that will cause the transistor Q_1 to act as a current-controlled current source (CCCS) with a selectable range $10\ \text{mA} \leq i_C \leq 100\ \text{mA}$.

Solution

Known Quantities: Transistor large-signal parameters; Li-ion battery nominal voltage.

Find: V_{CC}, R_1, R_2.

Schematics, Diagrams, Circuits, and Given Data: Figure 9.19. $V_{CC} = 12$ V; $V_\gamma = 0.6$ V; $\beta = 100$.

Assumptions: Assume that the transistor can be represented by the large-signal model.

Analysis: To determine the operating mode of the transistor, assume one of the three possible modes and check for any contradictions. First, if cutoff mode is assumed, $i_B = 0$ and $i_C = 0$. Clearly, this mode is not useful for charging. Moreover, KVL requires $V_{BE} + i_B(R_1 + R_2) = V_{CC}$, or since $i_B = 0$, $V_{BE} = V_{CC} = 12$ V. Thus, the EBJ would be forward-biased if $i_B = 0$. This result is a contradiction of the cutoff mode assumption, which therefore must be incorrect.

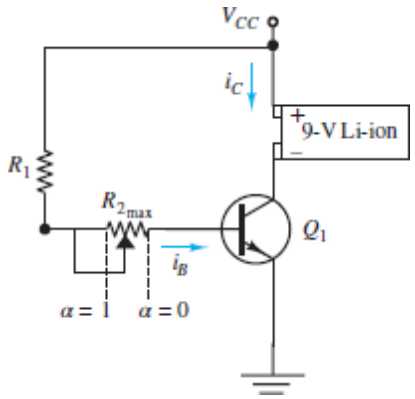


Figure 9.19 Simple battery charging circuit

Second, if the saturation mode is assumed, $V_{CE} \approx V_{CE\ sat} = 0.2$ V. However, KVL requires $V_{CE} + 9\text{ V} = V_{CC} = 12$ V, or $V_{CE} = 3$ V, which is a contradiction of the assumed saturation mode.

Thus, the transistor must be in active mode. The base and collector currents, i_B and i_C, are given by Ohm's law and $i_C = \beta i_B$, respectively.

$$i_B = \frac{V_{CC} - V_\gamma}{R_1 + R_2} \quad \text{and} \quad i_C = \beta \frac{V_{CC} - V_\gamma}{R_1 + R_2}$$

The bounds on the collector current i_C, which charges the battery, are

$$10\text{ mA} \le i_C \le 100\text{ mA}$$

The potentiometer wiper can be set to any value in the range $0 \leq \alpha \leq 1$ such that the resistance seen by the base is $R_1 + \alpha R_{2\max}$. The maximum collector current is obtained when the wiper is set to the far right position $\alpha = 0$. Thus, select R_1 by setting $i_C = i_{C_{max}} = 100$ mA when $\alpha = 0$.

$$100 \text{ mA} = \beta\left(\frac{V_{CC} - V_\gamma}{R_1}\right)$$

Page 589

or

$$R_1 = (V_{CC} - V_\gamma)\frac{\beta}{10^{-1}} = (12 - 0.6)\frac{100}{10^{-1}} = 11{,}400\ \Omega$$

If the value of R_1 is restricted to the E12 series of standard resistor values, the closest standard value is $R_1 = 12$ kΩ, which will result in a slightly lower maximum collector current. The rated value of the potentiometer $R_{2_{max}}$ is found by requiring that $i_C = i_{C_{min}} = 10$ mA when the wiper is set to the far left position $\alpha = 1$. Thus:

$$i_{C_{min}} = 10 \text{ mA} = \beta\frac{V_{CC} - V_\gamma}{R_1 + R_{2_{max}}}$$

or

$$R_{2_{max}} = \frac{\beta}{10 \text{ mA}}(V_{CC} - V_\gamma) - R_1 = 102{,}600\ \Omega$$

Again, if the value of $R_{2_{max}}$ is restricted to the E12 series of standard resistor values, the closest standard value is $R_{2_{max}} = 100$ kΩ, which results in a slightly higher minimum collector current.

Comments: A practical note on Li-ion batteries: a standard 9-V Li-ion battery is made up of two 3.6-V cells. Thus, the actual nominal battery voltage is 7.2 V. Further, as the battery becomes fully charged, each cell may rise as high as 4.2 V, leading to a fully charged voltage of 8.4 V.

EXAMPLE 9.5 Simple BJT Motor Drive Circuit

Problem

The aim of this example is to design a BJT driver for the Lego® 9V DC XL motor, model 8882. Figure 9.20 shows the driver circuit. The motor has a maximum (stall) current of 2,020 mA. The minimum current needed to start motor rotation is 110 mA. The aim of the circuit is to control the current to the motor (and therefore the motor torque, which is proportional to the current) through the potentiometer $R_{2_{max}}$.

Solution

Known Quantities: Transistor large-signal parameters; component values.

Find: Values of R_1 and $R_{2_{max}}$.

Schematics, Diagrams, Circuits, and Given Data: Figure 9.20. Maximum (stall) of 2,020 mA; minimum (start) current of 110 mA; $V_\gamma = 0.6$ V; $\beta = 40$; $V_{CC} = 12$ V.

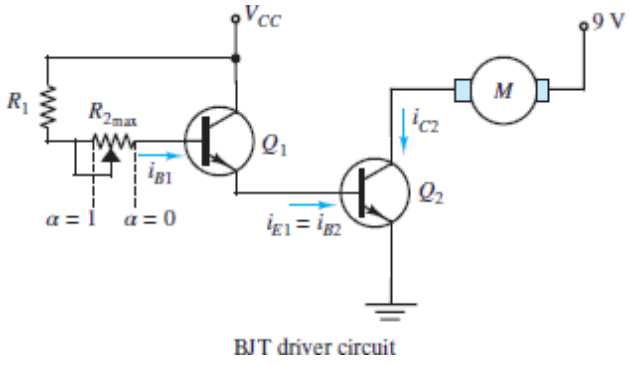


Figure 9.20 BJT motor driver circuit

Page 590

Assumptions: Use the large-signal model with $\beta = 40$ for each transistor.

Analysis: This circuit is a good example of how to stage transistors to accomplish a task that is difficult or impossible to accomplish with one transistor alone. Assume that both transistors are in active mode such that $i_C = \beta i_B$ for each transistor. Once a solution is found, the voltages across the EBJ and CBJ can be checked for compatibility with the active mode assumption.

It is important to recognize that the emitter current from Q_1 is the base current for Q_2. Since $i_{E1} = i_{C1} + i_{B1} = (\beta + 1)i_{B1}$, $i_{B2} = i_{E1}$, and $i_{C2} = \beta i_{B2}$, the collector current i_{C2} of Q_2 is related to the base current i_{B1} of Q_1 by:

$$i_{C2} = \beta i_{B2} = \beta i_{E1} = \beta(\beta + 1) i_{B1}$$

The base current i_{B1} of Q_1 is given by Ohm's law.

$$i_{B1} = \frac{V_{CC} - 2V_\gamma}{R_1 + R_{2_{max}}}$$

Therefore, the range of the motor current is

$$i_{C2_{min}} \leq \beta(\beta + 1)\left(\frac{V_{CC} - 2V_\gamma}{R_1 + R_{2_{max}}}\right) \leq i_{C2_{max}}$$

The potentiometer wiper can be set to any value in the range $0 \leq \alpha \leq 1$ such that the resistance seen by the base of Q_1 is $R_1 + \alpha R_{2_{max}}$. The maximum (stall) current for the motor is obtained when the wiper is set to the far right position $\alpha = 0$. Thus, select R_1 by setting $i_{C2} = i_{C2_{max}} = 2{,}020$ mA when $\alpha = 0$.

$$i_{C2_{max}} = 2.02 \text{ A} = \beta(\beta + 1)\left(\frac{V_{CC} - 2V_\gamma}{R_1}\right)$$

or

$$R_1 = \frac{\beta(\beta + 1)}{0.34}(V_{CC} - 2V_\gamma) = 8{,}768\ \Omega$$

If the value of R_1 is restricted to the E24 series of standard resistor values, the closest standard value is $R_1 = 9.1$ kΩ, which will result in a somewhat lower maximum motor current. The rated value of the potentiometer $R_{2_{max}}$ is

found by requiring that $i_{C2} = i_{C2_{min}} = 110$ mA when the wiper is set to the far left position $\alpha = 1$. Thus:

$$R_{2_{max}} = \frac{\beta(\beta+1)}{0.11}(V_{CC} - 2V_\gamma) - R_1 \approx 152 \text{ k}\Omega$$

Again, if the value of $R_{2_{max}}$ is restricted to the E24 series of standard resistor values, the closest standard value that is still greater than 152 kΩ is $R_{2_{max}} = 180$ kΩ, which results in a slightly lower minimum motor current. The lower minimum motor current will allow the motor to be turned off by adjusting the potentiometer. Great!

Comments: This design is simple and permits manual control of the motor current (and torque). If the motor is to be controlled by a microcontroller, the circuit should be redesigned to accept an external voltage input.

Page 591

CHECK YOUR UNDERSTANDING

Repeat the analysis of Example 9.3 for $R_C = 400$ Ω. In which mode is the transistor operating? What is the collector current?

What is the power dissipated by the LED in Example 9.3 if $R_C = 30$ Ω?

Answer: Saturation; 8.5 mA; 159 mW

CHECK YOUR UNDERSTANDING

In Example 9.4, what is V_{CE} when the battery is fully charged (8.4 V)? Is this value consistent with the assumption that the transistor is in active mode?

Answer: $V_{CE} \approx 3.6\text{V} \gg V_{CEsat} = 0.2\text{V}$; Yes

CHECK YOUR UNDERSTANDING

Compute the maximum and minimum possible motor currents for the circuit in Example 9.5 using the selected standard resistor values for R_1 and $R_{2_{max}}$.

Answer: $i_{C2_{max}} = 1{,}946.4\text{ mA}$; $i_{C2\,min} = 93.66\text{ mA}$

9.4 A BRIEF INTRODUCTION TO SMALL-SIGNAL AMPLIFICATION

The purpose of a DC operating point Q for a BJT circuit is to *bias* the BJT so that it is prepared to act as a linear amplifier for a *relatively small* time-varying input signal.

Typically, a time-varying voltage signal ΔV_B is superimposed upon a much larger DC voltage V_{BB}, as shown in Figure 9.21, such that the base current is also a time-varying function $I_B + \Delta I_B$. The primary objective of the DC biasing is to prevent the variation in the base current ΔI_B from driving the BJT out of active mode. This objective will be achieved if the maximum variation in the base current $\Delta I_{B_{max}}$ is *small compared to the DC bias current* I_B and if I_B is picked such that the operating point of the BJT is located in active mode, far from cutoff and saturation. An example of such an operating point Q is shown in Figure 9.14. In that figure, notice that I_B would have to change by at least ±100 μA from the 150-μA bias current for the BJT to leave active mode and enter either cutoff or saturation. As the base current changes, the location of Q simply moves along the load line,

either up and to the left as I_B increases, or down and to the right as I_B decreases.

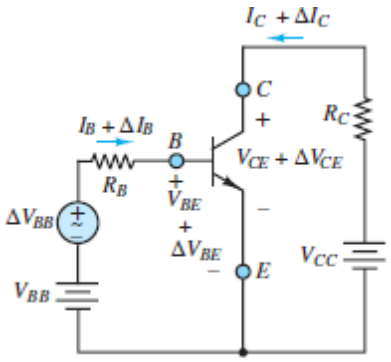


Figure 9.21 Circuit illustrating the amplification effect in a BJT

The phrase *small-signal model* refers to the fact that the maximum variation in the amplified signal must be small compared to the DC bias conditions.

Page 592

As long as the BJT remains in active mode the collector current will be roughly proportional to the base current, such that:

$$I_C + \Delta I_C = \beta(I_B + \Delta I_B) \tag{9.13}$$

Further, as seen in Figure 9.21, KVL around the collector source network yields:

$$V_{CE} + \Delta V_{CE} = V_{CC} - (I_C + \Delta I_C)R_C = V_{CC} - \beta(I_B + \Delta I_B)R_C \tag{9.14}$$

In the quiescent state (no time-varying input signal), this equation becomes:

$$V_{CE} = V_{CC} - I_C R_C = V_{CC} - \beta I_B R_C \tag{9.15}$$

Subtract equation 9.15 from equation 9.14 to obtain:

$$\Delta V_{CE} = \Delta I_C R_C = \beta \Delta I_B R_C \tag{9.16}$$

Notice that the variation in the collector-emitter voltage ΔV_{CE} is proportional to the variation in the base current, where the constant of proportionality is βR_C.

Further analysis applying KVL around the base source network yields:

$$V_{BB} + \Delta V_{BB} = (I_B + \Delta I_B) R_B + V_{BE} + \Delta V_{BE} \tag{9.17}$$

In the quiescent state, this equation becomes:

$$V_{BB} = I_B R_B + V_{BE} \tag{9.18}$$

Subtract equation 9.18 from equation 9.17 to obtain:

$$\Delta V_{BB} = \Delta I_B R_B + \Delta V_{BE} \tag{9.19}$$

or

$$\Delta I_B = \frac{\Delta V_{BB} - \Delta V_{BE}}{R_B} \tag{9.20}$$

Use this result to substitute for ΔI_B in equation 9.16 to obtain:

$$\Delta V_{CE} = \beta \frac{R_C}{R_B} (\Delta V_{BB} - \Delta V_{BE}) \tag{9.21}$$

This equation shows that the time-varying component ΔV_{BB} of the *input voltage* is amplified by a factor of $\beta R_C / R_B$ to produce a time-varying component ΔV_{CE} of the *output voltage*. Notice that the *output* of the BJT circuit in Figure 9.21 is considered to be the collector-emitter voltage.

It is important to mention that equation 9.21 shows that the expression for ΔV_{CE} is proportional to ΔV_{BB} only if ΔV_{BE} is negligibly small compared to ΔV_{BB}. Keep in mind that when the BJT is in active mode the EBJ is forward-biased such that the operating point for the EBJ diode is located

along the steep portion of the curve shown in Figure 9.6. As a result, ΔV_{BE} tends to be quite small for changes in I_B. Whether ΔV_{BE} is negligible requires more analysis than is appropriate here. Besides, there are other nonideal behaviors of a BJT that prevent the amplifier from being completely linear. The key point is that if the BJT is properly biased, these nonideal effects can be kept small.

An example of the amplification process described above is illustrated in Figure 9.22, where a time-varying sinusoidal collector current $I_C + \Delta I_C$ is shown Page 593to the right of the horizontal time axis and the resulting time-varying sinusoidal collector- emitter voltage $V_{CE} + \Delta V_{CE}$ is shown below the V_{CE} axis. Notice that the base current oscillates between 110 and 190 μA, causing the collector current to correspondingly fluctuate between 15.3 and 28.6 mA. Thus, the BJT acts as a *current amplifier.*

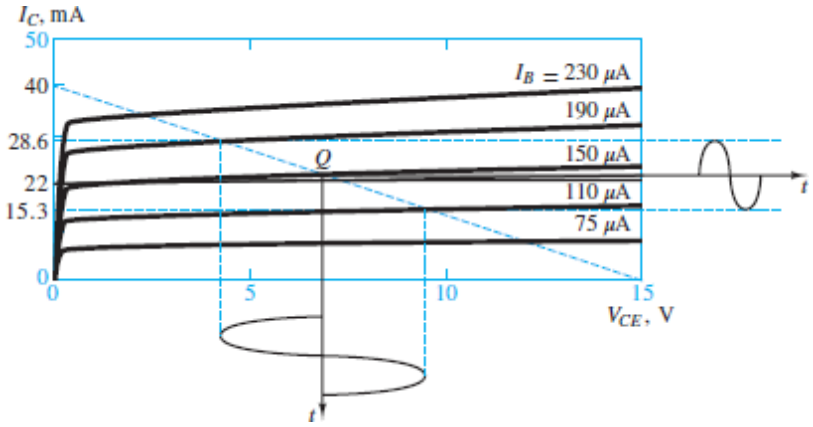


Figure 9.22 Amplification of sinusoidal oscillations in a BJT

A Practical Self-Biasing BJT Circuit

In practice, the circuit shown in Figure 9.21 can be used to bias a BJT; however, it has some weaknesses that can create serious problems in applications. In particular, variations in temperature can cause the operating point Q to shift significantly, and perhaps result in *thermal runaway.* Even if temperature effects are compensated for by other means, the Q points for two apparently identical reproductions of this circuit can be significantly different if the β values for the two BJTs are significantly different, as is often the case even in BJTs of the same type and lot.

A much better *self-biasing* circuit that automatically compensates for such parameter variations is shown in Figure 9.23. This circuit also has the added advantage of needing only one common power supply V_{CC}. Notice that V_{CC} appears across both (R_1, R_2) and (R_C, R_E) such that the circuit can be redrawn as shown in Figure 9.24(a). The Thévenin equivalent network seen by the base is shown in Figure 9.24(b), where:

$$V_{BB} = \frac{R_2}{R_1 + R_2} V_{CC} \tag{9.22}$$

and

$$R_B = R_1 \| R_2 = \frac{R_1 R_2}{R_1 + R_2} \tag{9.23}$$

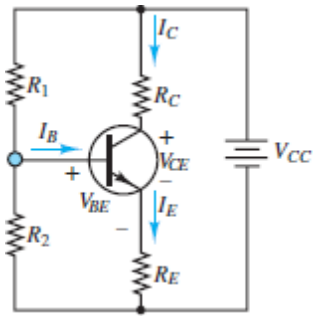


Figure 9.23 Practical single power supply BJT self-bias DC circuit

Notice that the circuit in Figure 9.24(b) closely resembles the circuit in Figure 9.21. The important difference is the presence of R_E between the emitter and the node along the bottom portion of the diagram.

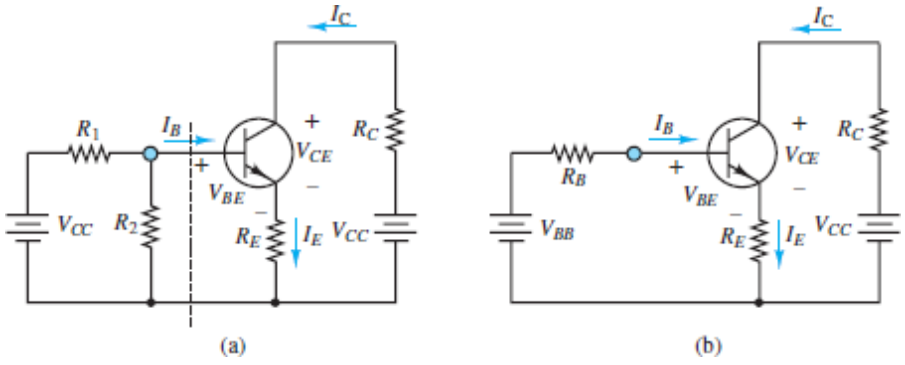


Figure 9.24 DC self-bias circuit represented in equivalent-circuit form

KVL can be applied around the base and collector networks to yield:

$$V_{BB} = I_B R_B + V_{BE} + I_E R_E = [R_B + (\beta + 1)R_E] I_B + V_{BE} \tag{9.24}$$

and

$$V_{CC} = I_C R_C + V_{CE} + I_E R_E = I_C\left(R_C + \frac{\beta + 1}{\beta} R_E\right) + V_{CE} \tag{9.25}$$

where

$$I_E = I_C + I_B = (\beta + 1)I_B = \frac{\beta + 1}{\beta} I_C \tag{9.26}$$

These two equations can be solved to obtain:

$$I_B = \frac{V_{BB} - V_{BE}}{R_B + (\beta + 1)R_E} \qquad I_C = \beta I_B \tag{9.27}$$

and

$$V_{CE} = V_{CC} - I_C\left(R_C + \frac{\beta + 1}{\beta} R_E\right) \tag{9.28}$$

The latter equation is the load line for the bias circuit. Notice that the effective load resistance seen by the collector circuit is now:

$$R_C + \frac{\beta + 1}{\beta} R_E \approx R_C + R_E \qquad \beta \gg 1$$

rather than simply R_C.

The role of R_E is to provide *negative feedback* to a change in the operating point Q due to, for example, a change in temperature that, in turn, changes β of the transistor. Refer to Figure 9.24(b) for the case of a change

$\Delta\beta$. The most immediate effect is a change in the collector current $\Delta I_C = \Delta\beta I_B$. In turn, this change results in a change in the emitter current $\Delta I_E = \Delta I_C + \Delta I_B$. It is here that R_E plays its part. The change in the emitter current results in a change in the voltage across R_E of $\Delta I_E R_E$, which then brings about a change in the voltage V_{BE} across the EBJ. Finally, this change in V_{BE} brings about a change in the base current due to the fact that the EBJ is a diode. At this point, it is important to realize that the change in base current *always* tends to offset the original change in the collector current because $\Delta I_C = \beta\Delta I_B$. In other words, if $\Delta\beta$ is positive, then ΔI_B will be negative, and vice versa. Thus, while a change in β tends to move the operating point Q, the effect of R_E is to restrain Q from moving.

Page 595

EXAMPLE 9.6 A BJT Small-Signal Amplifier

Problem

With reference to the BJT amplifier of Figure 9.25 and to the collector characteristic curves of Figure 9.22, determine (1) the DC operating point of the BJT, (2) the nominal current gain β at the operating point, and (3) the AC voltage gain $G = \Delta V_o/\Delta V_B$.

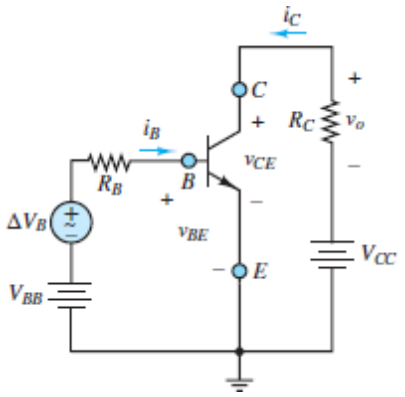


Figure 9.25

Solution

Known Quantities: Base, collector, and emitter resistances; base and collector supply voltages; collector characteristic curves; *BE* junction offset voltage.

Find: (1) DC (quiescent) base and collector currents I_{BQ} and I_{CQ} and collector-emitter voltage V_{CEQ}, (2) $\beta = \Delta I_C/\Delta I_B$, and (3) $G = \Delta V_o/\Delta V_B$.

Schematics, Diagrams, Circuits, and Given Data: $R_B = 10$ kΩ; $R_C = 375\ \Omega$; $V_{BB} = 2.1$ V; $V_{CC} = 15$ V; $V_\gamma = 0.6$ V. Collector characteristic curves such as those shown in Figure 9.27.

Assumptions: Assume that the *BE* junction resistance is negligible compared to the base resistance. Assume that each voltage and current can be represented by the superposition of a DC (quiescent) value and an AC component, for example, $v_0 = V_{0Q} + \Delta V_0$.

Analysis:

1. *DC operating point.* If the resistance of the *BE* junction is assumed to be much smaller than R_B, any change in the voltage across the EBJ is negligible such that $v_{BE} = V_{BEQ} = V_\gamma$. Figure 9.26 shows the resulting DC equivalent base circuit. KVL yields:

$$V_{BB} = R_B I_{BQ} + V_{BEQ}$$

The quiescent base current can be computed as:

$$I_{BQ} = \frac{V_{BB} - V_{BEQ}}{R_B} = \frac{V_{BB} - V_\gamma}{R_B} = \frac{2.1 - 0.6}{10{,}000} = 150\ \mu\text{A}$$

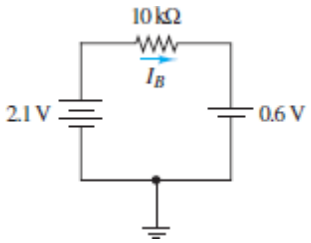


Figure 9.26

The load-line equation for the collector circuit is given by KVL as:

$$V_{CE} = V_{CC} - R_C I_C = 15 - 375 I_C$$

The load line and its intersection Q with the $I_B = 150\ \mu A$ line is shown in Figure 9.27. At the operating or quiescent point Q, $V_{CEQ} = 6.75$ V, $I_{CQ} = 22$ mA, and $I_{BQ} = 150\ \mu$A.

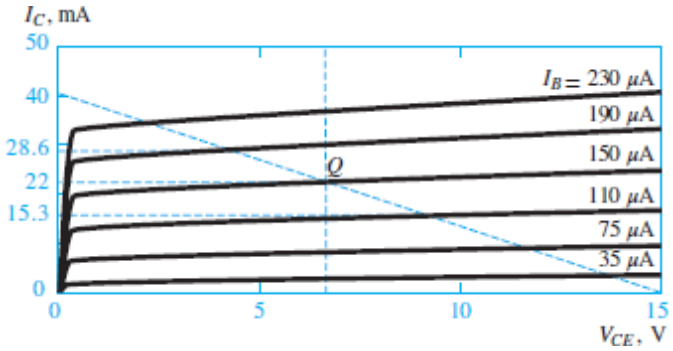


Figure 9.27 Operating point on the characteristic curve

2. *AC gain.* The current gain is determined from the characteristic curves of Figure 9.27. The collector current values corresponding to base currents of 190 and 110 μA are 28.6 Page 596and 15.3 mA, respectively. These collector current excursions ΔI_C from the Q point correspond to the effects of an oscillation ΔI_B in the base current. Thus, the current gain of the BJT amplifier can be computed as:

$$\beta = \frac{\Delta I_C}{\Delta I_B} = \frac{28.6 \times 10^{-3} - 15.3 \times 10^{-3}}{190 \times 10^{-6} - 110 \times 10^{-6}} = 166.25$$

which is the nominal current gain of the transistor.

3. *AC voltage gain.* To determine the AC voltage gain $G = \Delta V_o / \Delta V_B$, express ΔV_o as a function of ΔV_B. Observe that $v_o = -R_C i_C = -R_C I_{CQ} - R_C\ \Delta I_C$. Thus:

$$\Delta V_o = -R_C\ \Delta I_C = -R_C \beta\ \Delta I_B$$

The principle of superposition allows ΔI_B to be computed from the KVL equation for the base circuit.

$$\Delta V_B = R_B \, \Delta I_B + \Delta V_{BE}$$

However, due to the assumed small EBJ resistance, ΔV_{BE} is negligible. Thus:

$$\Delta I_B = \frac{\Delta V_B}{R_B}$$

Substitute this result into the expression for ΔV_o to find

$$\Delta V_o = -R_C \beta \, \Delta I_B = -\frac{R_C \beta \Delta V_B}{R_B}$$

or

$$\frac{\Delta V_o}{\Delta V_B} = G = -\frac{R_C}{R_B}\beta = -6.23$$

Comments: The circuit examined in this example is not self-biasing, but it demonstrates most of the essential features of BJT amplifiers, which are summarized below.

- Transistor amplifier analysis is greatly simplified by applying the principle of superposition to consider the DC bias circuit and the AC equivalent circuits separately.
- Once the bias point Q has been determined, the current gain can also be determined. Its value is somewhat dependent on the location of Q.
- The AC voltage gain of the amplifier is strongly dependent on R_B and R_C. Note that the AC voltage gain ΔV_o is negative! This inversion corresponds to a 180° phase shift for a sinusoidal AC input. This result is typical of all *common-emitter* amplifiers.

It is important to master this example when studying this section.

EXAMPLE 9.7 Practical BJT Bias Circuit

Problem

Determine the DC bias point of the transistor in the circuit of Figure 9.23.

Solution

Known Quantities: Base, collector, and emitter resistances; collector supply voltage; nominal transistor current gain; *BE* junction offset voltage.

Find: DC (quiescent) base and collector currents I_{BQ} and I_{CQ} and collector-emitter voltage V_{CEQ}.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 100$ kΩ; $R_2 = 50$ kΩ; $R_C = 5$ kΩ; $R_E = 3$ kΩ; $V_{CC} = 15$ V; $V_\gamma = 0.7$ V, $\beta = 100$.

Page 597

Analysis: Compute the equivalent base voltage from equation 9.22,

$$V_{BB} = \frac{R_2}{R_1 + R_2} V_{CC} = \frac{50}{100 + 50} 15 = 5 \text{ V}$$

and the equivalent base resistance from equation 9.23.

$$R_B = R_1 \| R_2 = 33.3 \text{ k}\Omega$$

Compute the base current from equation 9.27.

$$I_B = \frac{V_{BB} - V_{BE}}{R_B + (\beta + 1) R_E} = \frac{V_{BB} - V_\gamma}{R_B + (\beta + 1) R_E} = \frac{5 - 0.7}{33{,}000 + 101 \times 3{,}000} = 12.8 \ \mu\text{A}$$

Knowing the current gain of the transistor β, compute the collector current:

$$I_C = \beta I_B = 1.28 \text{ mA}$$

Finally, the collector-emitter junction voltage can be computed with reference to equation 9.28:

$$V_{CE} = V_{CC} - I_C\left(R_C + \frac{\beta+1}{\beta}R_E\right)$$
$$= 15 - 1.28\times 10^{-3}\left(5\times 10^3 + \frac{101}{100}\times 3\times 10^3\right) = 4.78 \text{ V}$$

Thus, the Q point of the transistor is given by:

$$V_{CEQ} = 4.73 \text{ V} \qquad I_{CQ} = 1.28 \text{ mA} \qquad I_{BQ} = 12.8\ \mu\text{A}$$

Comment In practice, the value of β is not used in calculations because its value can vary greatly even within the same type of transistor. Instead, amplifier designs rely on $\beta \gg 1$, which is a feature of all BJTs.

CHECK YOUR UNDERSTANDING

In Example 9.6, find the new Q point if R_C is increased to 680 Ω.

Answer: Since V_{BB} and R_B are unchanged and the change in V_{BEQ} is negligible, I_{BQ} will remain approximately equal to 150 μA. By observation, $V_{CEQ} \approx 0.5$ V is much smaller and the BJT is close to saturation. The new collector current $I_{CQ} \approx 20$ mA.

CHECK YOUR UNDERSTANDING

In the circuit of Figure 9.24, find the value of V_{BB} that yields a collector current I_C = 6.3 mA. What is the corresponding collector-emitter voltage? Assume that V_{BE} = 0.6 V, R_B = 50 kΩ, R_E = 200 Ω, R_C = 1 kΩ, β = 100, and V_{CC} = 14 V.

What percentage change in collector current would result if β were changed to 150 in Example 9.8? Why does the collector current increase less than 50 percent?

Answer: $V_{BB} = 5$ V, $V_{CE} = 6.43$ V; 3.74 percent. Because R_E provides negative feedback action that will keep I_C and I_E nearly constant

Page 598

9.5 GATES AND SWITCHES

In describing the properties of transistors, it was suggested that, in addition to serving as amplifiers, three-terminal devices can be used as electronic switches in which one terminal controls the current between the other two. It had also been hinted in Chapter 8 that diodes can act as on/off devices as well. In this section, we discuss the operation of diodes and transistors as electronic switches, illustrating the use of these electronic devices as the switching circuits that are at the heart of **analog** and **digital gates**. Transistor switching circuits form the basis of digital logic circuits, which are discussed in greater detail in Chapter 11. The objective of this section is to discuss the internal operation of these circuits and to provide the reader interested in the internal workings of digital circuits with an adequate understanding of the basic principles.

An **electronic gate** is a device that, on the basis of one or more input signals, produces one of two or more prescribed outputs; as will be seen shortly, one can construct both digital and analog gates. A word of explanation is required, first, regarding the meaning of the words *analog* and *digital.* An analog voltage or current—or, more generally, an analog signal—is one that varies in a continuous fashion over time, in *analogy* (hence the expression *analog*) with a physical quantity. An example of an analog signal is a sensor voltage corresponding to ambient temperature on any given day, which may fluctuate between, say, 30 and 50°F. A digital signal, on the other

hand, is a signal that can take only a finite number of values; in particular, a commonly encountered class of digital signals consists of **binary signals**, which can take only one of two values (for example, 1 and 0). A typical example of a binary signal would be the control signal for the furnace in a home heating system controlled by a conventional thermostat, where one can think of this signal as being on (or 1) if the temperature of the house has dropped below the thermostat setting (desired value), or off (or 0) if the house temperature is greater than or equal to the set temperature (say, 68°F). Figure 9.28 illustrates the appearance of the analog and digital signals in this furnace example.

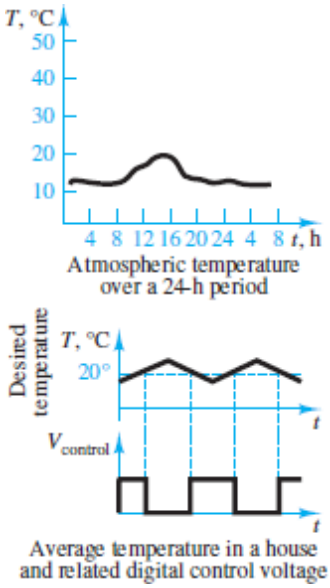


Figure 9.28 Illustration of analog and digital signals

The discussion of digital signals will be continued and expanded in Chapters 11 and 12. Digital circuits are an especially important topic because a large part of today's industrial and consumer electronics is realized in digital form.

Diode Gates

Recall that a diode conducts current when it is forward-biased and otherwise acts very much as an open-circuit. Thus, the diode can serve as a switch if properly employed. The circuit of Figure 9.29 is called an **OR gate**; it operates as follows. Let voltage levels greater than, say, 2 V correspond to a "logic 1" and voltages less than 2 V represent a "logic 0." Suppose, then,

that input voltages v_A and v_B can be equal to either 0 V or 5 V. If v_A = 5 V, diode D_A will conduct; if v_A = 0 V, D_A will act as an open-circuit. The same argument holds for D_B. It should be apparent, then, that the voltage across the resistor R will be 0 V, or logic 0 if both v_A and v_B are 0. If either v_A or v_B is equal to 5 V, though, the corresponding diode will conduct, and—assuming an offset model for the diode with V_γ = 0.6 V—we find that v_o = 4.4 V, or logic 1. Similar analysis yields an equivalent result if both v_A and v_B are equal to 5 V.

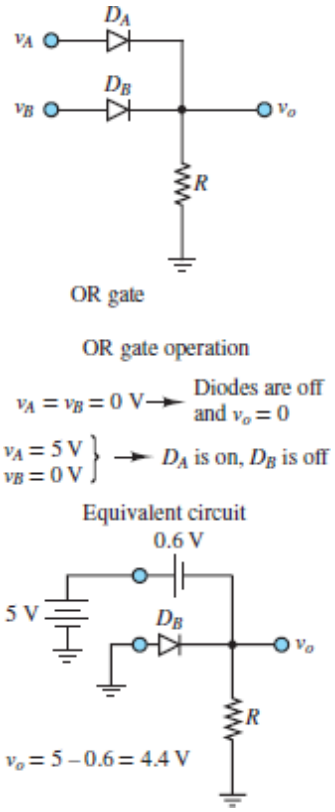


Figure 9.29 Diode OR gate

This type of gate is called an OR gate because v_o is equal to logic 1 (or "high") if either v_A *or* v_B is on while it is logic 0 (or "low") if neither v_A *nor* v_B is on. Other functions can also be implemented; however, the discussion of diode gates will be limited to this simple introduction because diode gate circuits, such Page 599as the one of Figure 9.29, are rarely, if ever,

employed in practice. Most modern digital circuits employ transistors to implement switching and gate functions.

BJT Gates

The large-signal models and the *i-v* characteristics of BJTs include a *cutoff* mode, where the collector current is virtually zero. On the other hand, when sufficient current is injected into the base, a BJT will reach *saturation,* and the collector current is independent of the base current. This behavior is quite well suited to the design of electronic gates and switches and can be visualized by superimposing a load line on the collector characteristic, as shown in Figure 9.30.

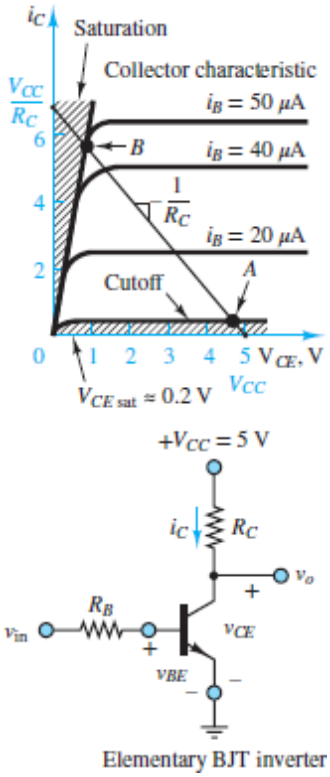


Figure 9.30 BJT switching characteristic and simple logic inverter circuit

The operation of the simple **BJT switch** is illustrated in Figure 9.30. The load-line equation of the collector circuit is

$$v_{CE} = V_{CC} - i_C R_C \tag{9.29}$$

and

$$v_o = v_{CE} \tag{9.30}$$

Thus, when the input voltage v_{in} is low (say, 0 V), the transistor is in cutoff mode and the collector current is very small. Then:

$$v_o = v_{CE} \approx V_{CC} \tag{9.31}$$

such that the output is "logic high." This result is represented by point *A* in Figure 9.30.

When v_{in} is large enough to drive the transistor into saturation, the collector-emitter voltage saturates at $V_{CE\ \text{sat}}$, which is typically on the order of 0.2 V. This result corresponds to point *B* in Figure 9.30. For the input voltage v_{in} to drive the BJT into saturation, a base current of approximately 50 μA is required. Assuming $v_{\text{in}} = 5$ V and $R_B = 82$ kΩ, then $i_B = (v_{\text{in}} - V_\gamma)/R_B = (5 - 0.6)/82{,}000 \approx 54\ \mu$A such that the BJT is in saturation, and $v_o = V_{CE\ \text{sat}} \approx 0.2$ V.

Thus, whenever v_{in} corresponds to a logic high (or logic 1), v_o takes a value close to 0 V, or logic low (or 0); conversely, v_{in} = "0" (logic "low") leads to v_o = "1." The values of 5 and 0 V for the two logic levels 1 and 0 are quite common in practice and are the standard values used in a family of logic circuits denoted by the acronym **TTL**, which stands for **transistor-transistor logic**.[3] One of the more common TTL blocks is the **inverter** shown in Figure 9.30, so called because it "inverts" the input by providing a low output for a high input, and vice versa. This type of inverting, or "negative," logic behavior is typical of BJT *common-emitter* amplifier circuits.

EXAMPLE 9.8 TTL NAND Gate

Problem

Refer to Figure 9.31 and complete the table below to determine the logic gate operation of a TTL NAND gate, which acts as an inverted AND gate (thus the prefix N in NAND, which stands for NOT).

Page 600

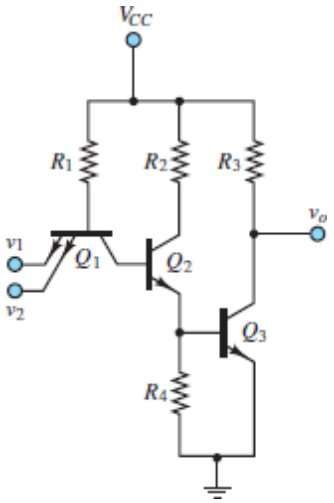


Figure 9.31 TTL NAND gate

v_1 (V)	v_2 (V)	State of Q_1	State of Q_2	v_o
0	0			
0	5			
5	0			
5	5			

Solution

Known Quantities: Resistor values; $V_{BE\ \text{on}}$ and $V_{CE\ \text{sat}}$ for each transistor.

Find: v_o for each of the four combinations of v_1 and v_2.

Schematics, Diagrams, Circuits, and Given Data: $R_1 = 5.7$ kΩ; $R_2 = 2.2$ kΩ; $R_3 = 2.2$ kΩ; $R_4 = 1.8$ kΩ; $V_{CC} = 5$ V; V_{BE} on $= V_\gamma = 0.7$ V; $V_{CE\,\text{sat}} = 0.2$ V.

Assumptions: Model the *BE* junctions of Q_1 as offset diodes. Assume that the transistors are in saturation when conducting.

Analysis: The inputs to the TTL gate, v_1 and v_2, are applied to the emitter of transistor Q_1. The transistor is designed so as to have two emitter circuits in parallel. Transistor Q_1 is modeled by the offset diode model, as shown in Figure 9.32. Consider each of the four cases.

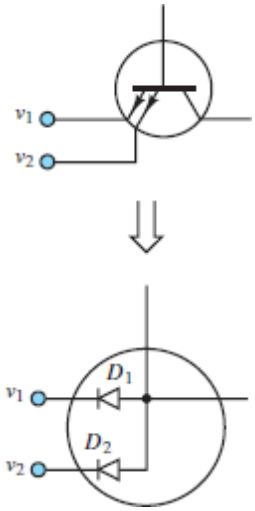


Figure 9.32

1. $v_1 = v_2 = 0$ V. With the emitters of Q_1 connected to ground and $V_{CC} = 5$ V, the *BE* junction will clearly be forward-biased and Q_1 is on. This result means that the base current of Q_2 (equal to the collector current of Q_1) is negative, and therefore Q_2 must be off. If Q_2 is off, its emitter current must be zero, and therefore the base current into Q_3 is zero. With Q_3 off, the current through R_3 is zero, and therefore $v_o = 5 - v_{R3} = 5$ V.
2. $v_1 = 5$ V; $v_2 = 0$ V. With reference to Figure 9.32, diode D_2 is still forward-biased, but D_1 is now reverse-biased because of the 5-V

potential at v_1. Thus, the EBJ conducts current and Q_1 is on. The remainder of the analysis is the same as in case 1, and Q_2 and Q_3 are both off, leading to $v_o = 5$ V.

3. $v_1 = 0$ V; $v_2 = 5$ V. By symmetry with case 2, one emitter branch is conducting, Q_1 is on, Q_2 and Q_3 are off, and $v_o = 5$ V.
4. $v_1 = 5$ V; $v_2 = 5$ V. Here, diodes D_1 and D_2 are both reverse-biased, there is no emitter current, and Q_1 is off. Note, however, that although D_1 and D_2 are reverse-biased, the BCJ of Q_1 is forward-biased, and a base current exists for Q_2; thus, Q_2 is on and its emitter current turns on Q_3. To determine the output voltage, assume that Q_3 is operating in saturation such that:

$$v_o \approx V_{CE\,\text{sat}}$$

Ohm's law can be applied to R_3 to find:

$$I_{C3} = \frac{V_{CC} - v_o}{R_3} = \frac{V_{CC} - V_{CE\,\text{sat}}}{R_3} = \frac{5 - 0.2}{2{,}200} \approx 2.2 \text{ mA}$$

A reasonable question is, Can Q_2 also be in saturation? If it is, then R_2 and R_4 are virtually in series and the base voltage of Q_3 can be computed by voltage division.

$$V_{B_3} \approx \frac{R_4}{R_2 + R_4}(V_{CC} - V_{CE\,\text{sat}}) = \frac{1.8}{2.2 + 1.8}4.8 = 2.16 \text{ V}$$

Since the emitter of Q_3 is tied directly to the reference node ($V = 0$), the voltage across the EBJ of Q_3 would also be 2.16 V. But this value is incompatible with the assumption that $V_\gamma \approx 0.7$ V for silicon-based transistors. Thus, Q_2 cannot be in saturation. But since it is on, it must be in active mode.

Page 601

The results for all four cases are summarized in the table below. The output values are consistent with TTL logic; the output voltage for case 4 is

sufficiently close to zero to be considered zero for logic purposes.

v_1 (V)	v_2 (V)	State of Q_2	State of Q_3	v_o (V)
0	0	Off	Off	5
0	5	Off	Off	5
5	0	Off	Off	5
5	5	On	On	0.2

Comments: While exact analysis of TTL logic gate circuits could be tedious and involved, the method demonstrated in this example—to determine whether transistors are on or off—leads to a very simple analysis. When working with logic devices, the primary interest is in logic levels rather than exact values; thus, approximations are appropriate.

CHECK YOUR UNDERSTANDING

Use the BJT switching characteristic of Figure 9.30 to find the value of R_B required to drive the transistor to saturation. Assume a base current of 50 μA when the minimum v_{in} to turn on the transistor is 2.5 V.

Answer: $R_B \leq 38$ kΩ

Conclusion

This chapter introduces the bipolar junction transistor, and by way of the simple circuit model demonstrates its operation as an amplifier and a switch. Upon completing this chapter, you should have mastered the following learning objectives:

1. *Understand the basic principles of amplification and switching.* Transistors are three-terminal electronic semiconductor devices that can serve as amplifiers and switches.

2. *Understand the physical operation of bipolar junction transistors; determine the operating point of a bipolar junction transistor circuit.* The transistor has four modes of operation. These can be readily identified by simple voltage measurements.
3. *Understand the large-signal model of the bipolar junction transistor and apply it to simple amplifier circuits.* The large-signal model of the BJT is very easy to use, requiring only a basic understanding of DC circuit analysis, and can be readily applied to many practical situations.
4. *Select the operating point of a bipolar junction transistor circuit.* Biasing a transistor consists of selecting the appropriate values for the DC supply voltage(s) and for the resistors that comprise a transistor amplifier circuit. When biased in the forward active mode, the transistor acts as a current-controlled current source and can amplify small currents injected into the base by as much as a factor of 200.
5. *Understand the principle of small-signal amplifiers*. WHEN MAINTAINED IN ACTIVE MODE, A BJT CAN BE USED TO PRODUCE A LINEAR AMPLIFIER, WHERE THE SHAPE OF A SMALL INPUT WAVEFORM IS REPRODUCED ON A MUCH LARGER SCALE AT AN OUTPUT TERMINAL.

Page 602

6. *Understand the operation of a bipolar junction transistor as a switch and analyze basic analog and digital gate circuits.* The operation of a BJT as a switch is very straightforward, and consists of designing a transistor circuit that will go from cutoff to saturation when an input voltage changes from a high to a low value, or vice versa. Transistor switches are commonly used to design digital logic gates.

HOMEWORK PROBLEMS

Section 9.2: The Bipolar Junction Transistor

9.1 For each transistor shown in Figure P9.1, determine whether the *BE* and *BC* junctions are forward- or reverse-biased, and determine the operating mode.

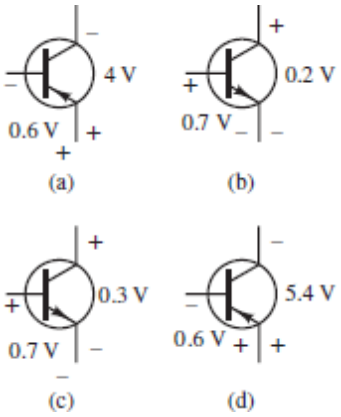


Figure P9.1

9.2 Determine the mode of operation for the following transistors:

a. *npn*, $V_{BE} = 0.8$ V, $V_{CE} = 0.4$ V

b. *npn*, $V_{CB} = 1.4$ V, $V_{CE} = 2.1$ V

c. *pnp*, $V_{EB} = 0.9$ V, $V_{EC} = 0.4$ V

d. *npn*, $V_{BE} = -1.2$ V, $V_{CB} = 0.6$ V

9.3 Given the circuit of Figure P9.3, estimate the operating point of the transistor. Assume $V_\gamma = 0.65$ V and $\beta = 150$.

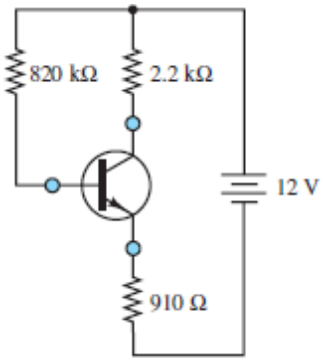


Figure P9.3

9.4 Refer to Figure 9.4 and assume that for a *pnp* transistor the emitter and base currents are $I_E = 5$ mA and $I_B = 0.2$ mA, respectively. The voltage

drops across the emitter-base and collector-base junctions are V_{EB} = 0.67 V and V_{CB} = 7.8 V. Find:

a. V_{CE}.

b. The collector current.

c. The total power dissipated in the transistor, defined here as $P = V_{CE}I_C + V_{BE}I_B$.

9.5 For the circuit shown in Figure P9.5, determine the emitter current I_E and the collector-base voltage V_{CB}, as defined in Figure 9.9. Assume V_γ = 0.62 V.

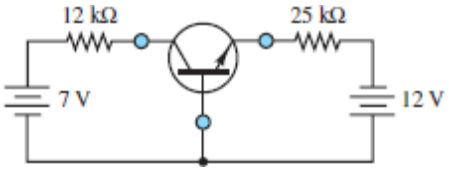


Figure P9.5

9.6 Given the circuit of Figure P9.6, determine V_{CE} and I_C. Assume β = 80, R_1 = 15 kΩ, R_2 = 25 kΩ, R_c = 2 kΩ, V_{BB} = 5 V, V_{CC} = 10 V, and V_{AA} = −4 V.

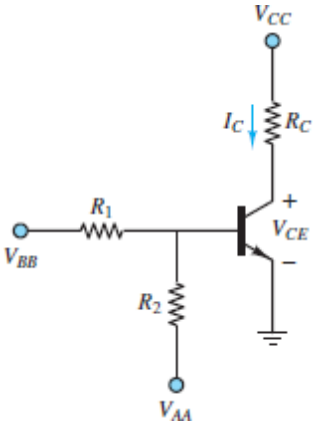


Figure P9.6

9.7 Given the circuit of Figure P9.7, determine the emitter current I_E and the collector-base voltage V_{CB}. Assume the offset voltage is $V_\gamma = 0.6$ V.

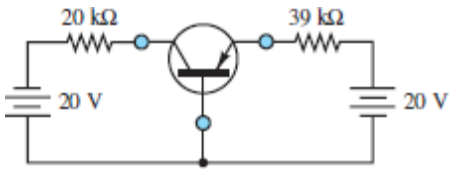


Figure P9.7

9.8 Given the circuit of Figure P9.8, estimate V_{CE} and I_C. Assume $R_1 = 50$ kΩ, $R_2 = 10$ kΩ, $R_C = 600\ \Omega$, $R_E = 400\ \Omega$, $V_{BE} = 0.7$ V, $I_B = 25\ \mu$A, $I_2 = 230\ \mu$A, and $V_{CC} = 15$ V. Assume a 2N2222A BJT.

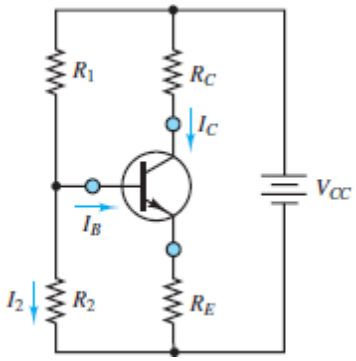


Figure P9.8

9.9 The collector characteristics for a certain transistor are shown in Figure P9.9.

a. Find the ratio I_C/I_B for $V_{CE} = 10$ V and $I_B = 100$, 200, and 600 μA.

b. If the maximum allowable collector power dissipation is $P = i_C v_{CE} = 0.5$ W for $I_B = 500\ \mu$A, find V_{CE}.

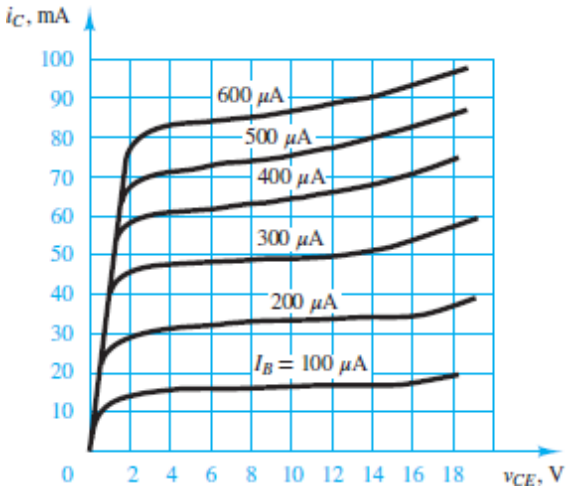


Figure P9.9

9.10 Given the circuit of Figure P9.10, determine the current I_R. Let $R_B = 30$ kΩ, $R_{C1} = 1$ kΩ, $R_{C2} = 3$ kΩ, $R = 7$ kΩ, $V_{BB1} = 4$ V, $V_{BB2} = 3$ V, $V_{CC} = 10$ V, $\beta_1 = 40$, and $\beta_2 = 60$.

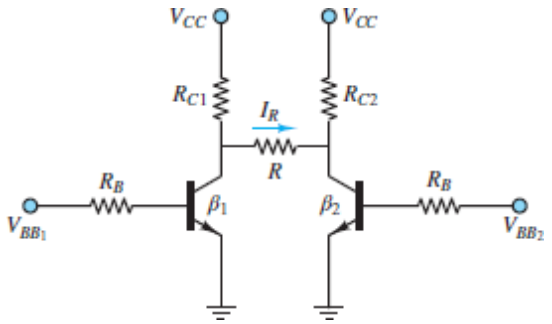


Figure P9.10

9.11 For the circuit shown in Figure P9.11, determine I_R. Let $R_B = 50$ kΩ, $R_C = 1$ kΩ, $R = 2$ kΩ, $V_{BB} = 2$ V, $V_{CC} = 12$ V, and $\beta = 120$.

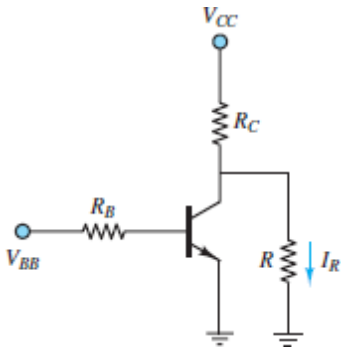


Figure P9.11

9.12 For the circuit shown in Figure P9.12, determine whether the transistor is in saturation. Let $R_B = 8\ \text{k}\Omega$, $R_E = 260\ \Omega$, $R_C = 1.1\ \text{k}\Omega$, $V_{CC} = 13$ V, $V_{BB} = 7$ V, and $\beta = 100$.

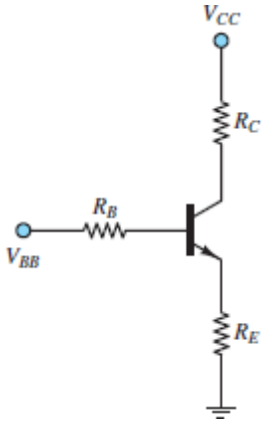


Figure P9.12

Page 604

9.13 For the circuit shown in Figure P9.8, $V_{CC} = 20$ V, $R_C = 5\ \text{k}\Omega$, and $R_E = 1\ \text{k}\Omega$. Determine the operating mode of the transistor if:

a. $I_C = 1$ mA, $I_B = 20\ \mu$A, $V_{BE} = 0.7$ V

b. $I_C = 3.2$ mA, $I_B = 0.3$ mA, $V_{BE} = 0.8$ V

c. $I_C = 3$ mA, $I_B = 1.5$ mA, $V_{BE} = 0.85$ V

9.14 For the circuit shown in Figure P9.14, find the minimum input voltage v_{in} required to saturate the transistor. Assume $V_{CC} = 5$ V, $R_C = 2$ kΩ, $R_B = 50$ kΩ, $V_{CE\ sat} = 0.2$ V, $V_{BE\ sat} = 0.8$ V, and $\beta = 50$.

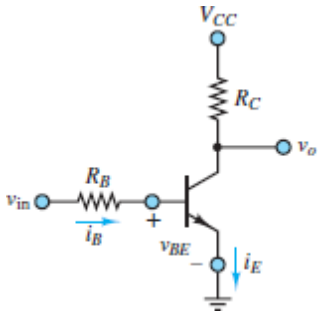
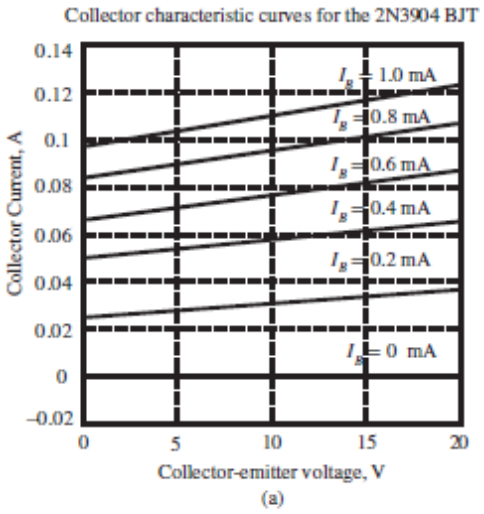
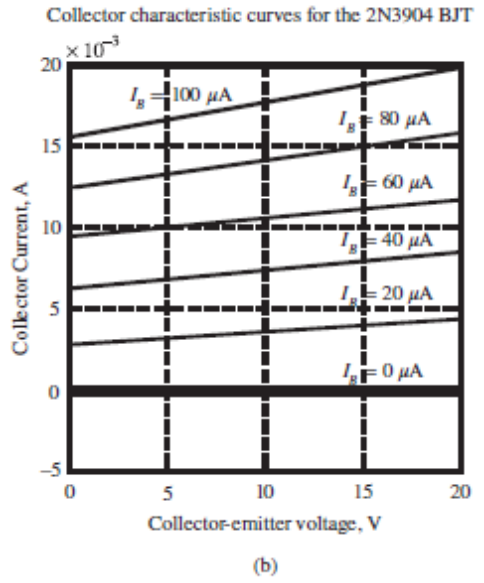


Figure P9.14

9.15 An *npn* transistor, such as that in Figure 9.9, is operated in active mode with $i_C = 60 i_B$ and with junction voltages of $V_{BE} = 0.6$ V and $V_{CB} = 7.2$ V. If $I_E = 4$ mA, find (a) I_B and (b) V_{CE}.

9.16 Use the collector characteristics of the 2N3904 *npn* transistor shown in Figure P9.16(a) and (b) to determine I_C and V_{CE} of the transistor in Figure P9.16(c). Is the transistor in the active mode? If so, determine its value of β.


Collector characteristic curves for the 2N3904 BJT
Collector Current, A
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
–0.02
I_B = 1.0 mA
I_B = 0.8 mA
I_B = 0.6 mA
I_B = 0.4 mA
I_B = 0.2 mA
I_B = 0 mA
0
5
10
15
20
Collector-emitter voltage, V
(a)


Collector characteristic curves for the 2N3904 BJT
× 10⁻³
Collector Current, A
20
15
10
5
0
–5
I_B = 100 μA
I_B = 80 μA
I_B = 60 μA
I_B = 40 μA
I_B = 20 μA
I_B = 0 μA
0
5
10
15
20
Collector-emitter voltage, V
(b)

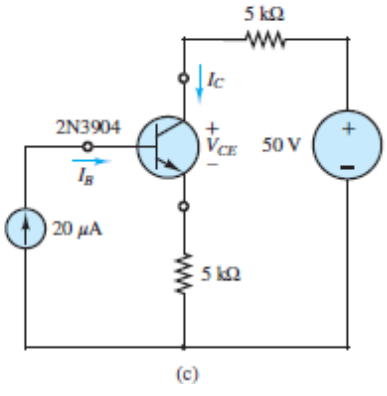

5 kΩ
I_C
2N3904
+
V_{CE}
–
50 V
I_B
20 μA
5 kΩ
(c)

Figure P9.16

Section 9.3: BJT Large-Signal Model

9.17 Refer to Example 9.3 and Figure 9.17. Assume that all given values are unchanged except that the application requires $I_{LED} \leq 10$ mA. Find the range of collector resistance R_C values that will permit the transistor to supply the required current.

9.18 Refer to the Focus on Measurements box, "Large-Signal Amplifier Design for a Diode Thermometer" and Figure 9.15. Assume $R_B = R_S = 100\ \Omega$. Estimate (by analysis or simulation) V_{CEQ} when $R_C = 330\ \Omega$ and when $R_C = 470\ \Omega$. Do the results make sense? How do these changes impact the performance of the diode thermometer?

9.19 Refer to Example 9.3 and Figure 9.17. Assume that all given values are unchanged except that $R_C = 340\ \Omega$, $I_{LED} \leq 10$ mA, and that the maximum base current supplied by the microprocessor is 5 mA. Find the range of values of R_B that satisfy these requirements.

Page 605

9.20 Use the same data given in Problem 9.19 but assume that $R_B = 10\ \text{k}\Omega$. Find the minimum value of β that satisfies the requirements.

9.21 Repeat Problem 9.20 for the case of a microprocessor operating on a 2.8-V supply (that is, $V_{ON} = 2.8$ V).

9.22 Consider the LED driver circuit of Figure 9.17. This circuit is now used to drive an automotive fuel injector (an electromechanical solenoid valve). The differences in the circuit are as follows: The collector resistor and the LED are replaced by the fuel injector, which can be modeled as a series *RL* circuit. The voltage supply for the fuel injector is 13 V (instead of 5 V). For the purposes of this problem, it is reasonable to assume $R = 12\ \Omega$ and $L \sim 0$. Assume that the maximum current that can be supplied by the microprocessor is 1 mA, that the current required to drive the fuel injector must be at least 1 A, and that the transistor saturation voltage is $V_{CE\ sat} = 1$ V. Find the minimum value of β required for the transistor.

9.23 Refer to Problem 9.22. Assume $\beta = 7{,}000$. Find the allowable range of R_B.

9.24 Given the circuit of Figure P9.8, find the minimum value of R_C such that transistor operates in active mode and dissipates less than 15 mW. Let $V_{CC} = 10$ V, $R_1 = R_2 = 40$ kΩ, $R_E = 1.5$ kΩ, $V_{BE} = 0.7$ V, $\beta = 70$, and $V_{CE\,\text{sat}} = 0.25$ V.

9.25 The circuit shown in Figure P9.25 is a 9-V battery charger. The purpose of the Zener diode is to provide a constant voltage across resistor R_2, such that the transistor will source a constant emitter (and therefore collector) current. Select the values of R_2, R_1, and V_{CC} such that the battery will be charged with a constant 40-mA current.

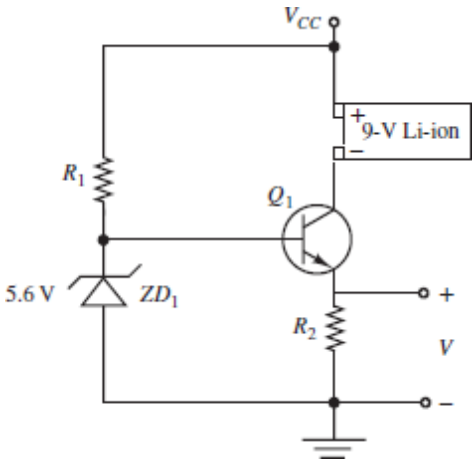


Figure P9.25

9.26 The circuit shown in Figure P9.26 is a variation of that shown in Figure P9.25. Analyze the operation of the circuit and explain how this circuit will provide a decreasing charging current (taper current cycle) until the Li-ion battery is fully charged (8.4 V—see Comments in Example 9.4). Choose appropriate values of V_{CC} and R_1 that would result in a practical design. Use standard resistor values.

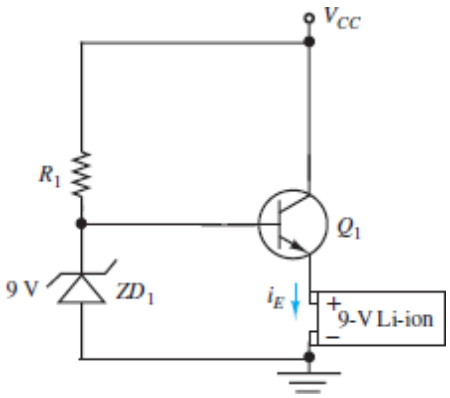


Figure P9.26

9.27 The circuit shown in Figure P9.27 is a variation of the motor driver circuit of Example 9.5. The external voltage v_{in} represents the analog output of a microcontroller and alternates between 0 and 5 V. Complete the design of the circuit by selecting the value of the base resistor R_b such that the motor will see the maximum design current when $v_{in} = 5$ V. Use the other specifications given in the example.

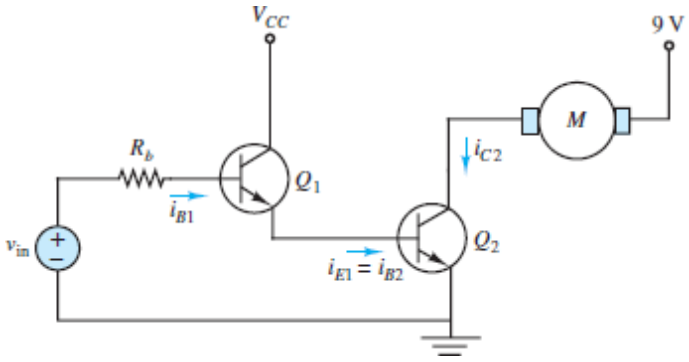


Figure P9.27

9.28 For the circuit shown in Figure P9.28, $R_C = 1$ kΩ, $V_{BB} = 5$ V, $\beta_{min} = 50$, and $V_{CC} = 10$ V. Find the range of R_B so that the transistor is in saturation.

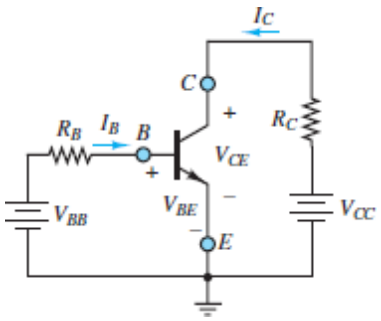


Figure P9.28

Page 606

9.29 For the circuit shown in Figure P9.28, $V_{CC} = 5$ V, $R_C = 1$ kΩ, $R_B = 10$ kΩ, and $\beta_{\text{min}} = 50$. Find the range of values of V_{BB} so that the transistor is in saturation.

9.30 For the circuit shown in Figure 9.13, $V_\gamma = 0.6$ V, $R_B = 100$ kΩ, $I_{BB} = 26$ μA, $R_C = 2$ kΩ, $V_{CC} = 10$ V, and $\beta = 100$. Find I_C, I_E, V_{CE}, and V_{CB}.

Section 9.4: A Brief Introduction to Small-Signal Amplification

9.31 The circuit shown in Figure P9.31 is a common-emitter amplifier stage. Determine the DC Thévenin equivalent of the network between the base node and the reference node. Use it to redraw the circuit.

$V_{CC} = 20$ V	$\beta = 130$
$R_1 = 1.8$ MΩ	$R_2 = 300$ kΩ
$R_C = 3$ kΩ	$R_E = 1$ kΩ
$R_o = 1$ kΩ	$R_S = 0.6$ kΩ
$v_S = 1\cos(6.28 \times 10^3 t)$ mV	

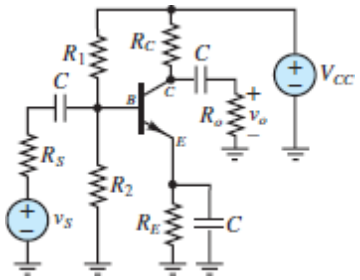


Figure P9.31

9.32 The circuit shown in Figure P9.32 is a common-collector (or emitter follower) amplifier stage implemented with an *npn* silicon transistor and a single DC supply V_{CC} = 12 V. Determine V_{CEQ} at the DC operating (Q) point.

$R_o = 16\ \Omega$ $\quad \beta = 130$
$R_1 = 82\ \text{k}\Omega$ $\quad R_2 = 22\ \text{k}\Omega$
$R_S = 0.7\ \text{k}\Omega$ $\quad R_E = 0.5\ \text{k}\Omega$

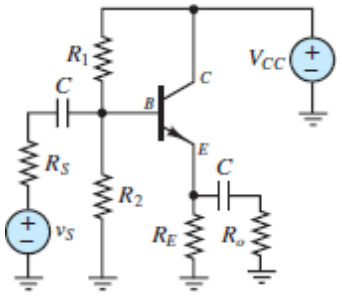


Figure P9.32

9.33 Shown in Figure P9.33 is a common-emitter amplifier stage implemented with an *npn* silicon transistor and two DC supplies V_{CC} = 12 V and V_{EE} = 4 V. Determine V_{CEQ} and the DC mode of operation.

$\beta = 100$ $\quad R_B = 100\ \text{k}\Omega$
$R_C = 3\ \text{k}\Omega$ $\quad R_E = 3\ \text{k}\Omega$
$R_o = 6\ \text{k}\Omega$ $\quad R_S = 0.6\ \text{k}\Omega$
$v_S = 1\cos(6.28 \times 10^3 t)\ \text{mV}$

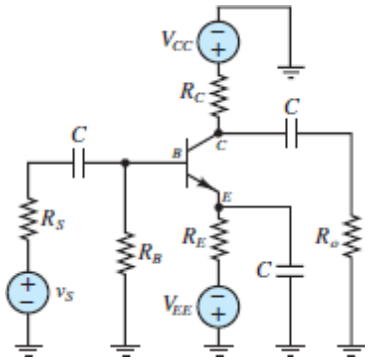


Figure P9.33

9.34 Shown in Figure P9.34 is a common-emitter amplifier stage implemented with an *npn* silicon transistor and a single DC supply V_{CC} = 12 V. Determine V_{CEQ} and the DC mode of operation.

$\beta = 130$ $\quad R_B = 325\ \text{k}\Omega$

$R_C = 1.9\ \text{k}\Omega$ $\quad R_E = 2.3\ \text{k}\Omega$

$R_o = 10\ \text{k}\Omega$ $\quad R_S = 0.5\ \text{k}\Omega$

$v_S = 1\cos(6.28 \times 10^3 t)\ \text{mV}$

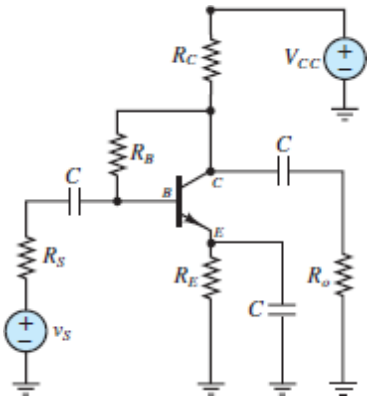


Figure P9.34

9.35 For the circuit shown in Figure P9.35 v_S is a small sine wave signal with average value of 3 V. If $\beta = 100$ and $R_B = 60\ \text{k}\Omega$,

a. Find the value of R_E so that I_E is 1 mA.

b. Find R_C so that V_C is 5 V.

c. For $R_o = 5\ \text{k}\Omega$, find the small-signal equivalent circuit of the amplifier.

d. Find the small-signal voltage gain.

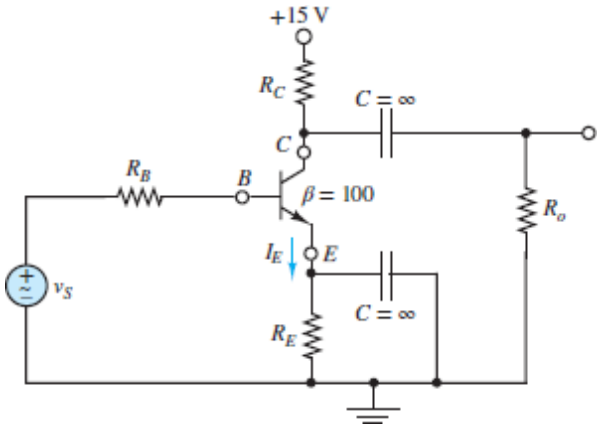


Figure P9.35

9.36 The circuit in Figure P9.36 is similar to a common collector when R_C is small. Assume $R_C = 200\ \Omega$. The AC coupling capacitor C_b blocks any DC component of v_S so assume v_S is a small sine wave signal as in Problem 9.35.

a. Find the operating point Q of the transistor.

b. Find the voltage gain v_o/v_{in}.

c. Find the current gain i_o/i_{in}.

d. Find the input resistance $r_i = v_{in}/i_{in}$.

e. Find the output resistance $r_o = v_o/i_o$.

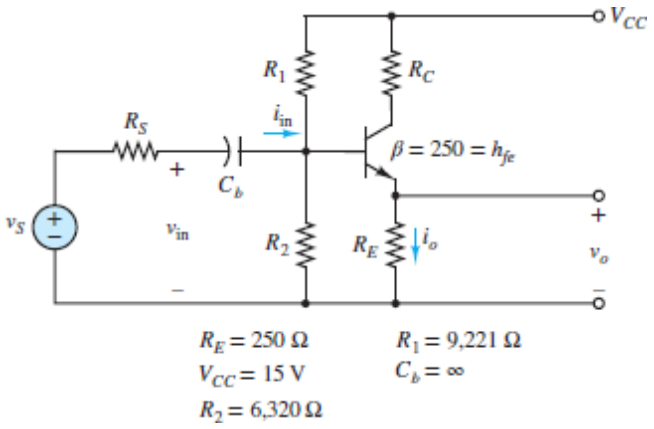


Figure P9.36

9.37 A Darlington pair of transistors is shown in Figure P9.37. The transistor parameters for large-signal operation are Q_1: $\beta = 130$; Q_2: $\beta = 70$. Calculate the overall small-signal current gain.

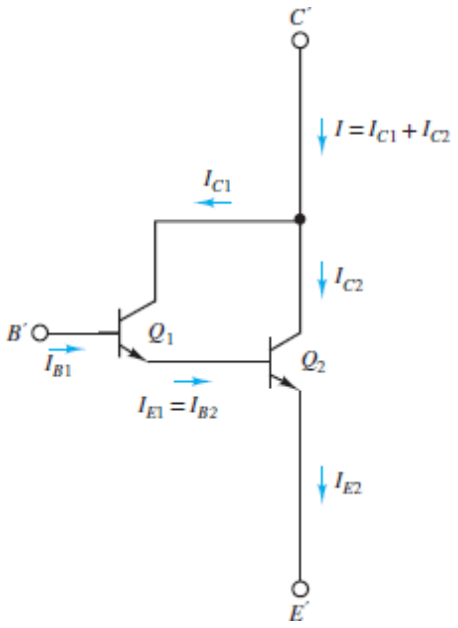


Figure P9.37

9.38 Assume the transistor shown in Figure P9.8 has $V_\gamma = 0.6$ V. Also assume $R_C = 1.5$ kΩ, $V_{CC} = 18$ V and $R_E = 1.0$ kΩ. Determine values for R_1 and R_2 such that:

a. The DC collector-emitter voltage V_{CEQ} is 5 V.

b. The DC collector current I_{CQ} will vary no more than 10 percent as β varies from 20 to 50.

c. Values of R_1 and R_2 that will permit maximum symmetrical swing in the collector current. Assume $\beta = 100$.

Section 9.5: BJT Switches and Gates

9.39 An automobile fuel injector system is depicted in Figure P9.39(a). The internal circuitry of the injector can be modeled as shown in Figure P9.39(b). The injector will inject gasoline into the intake manifold when $I_{inj} \leq 0.1$ A. A voltage pulse train v_{signal} is shown in Figure

P9.39(c). For a cold engine at start-up, the pulse width τ is determined by:

$$\tau = \text{BIT} \times K_C + \text{VCIT}$$

where

BIT = basic injection time = 1 ms
K_C = compensation constant of temperature of coolant (T_C)
VCIT = voltage-compensated injection time

The characteristics of VCIT and K_C are shown in Figure P9.39(d). Assume the transistor Q_1 saturates at Page 608V_{CE} = 0.3 V and V_{BE} = 0.9 V. Find the period of the fuel injector pulse if:

a. V_{batt} = 13 V, T_C = 100°C

b. Vbatt = 8.6 V, T_C = 20°C

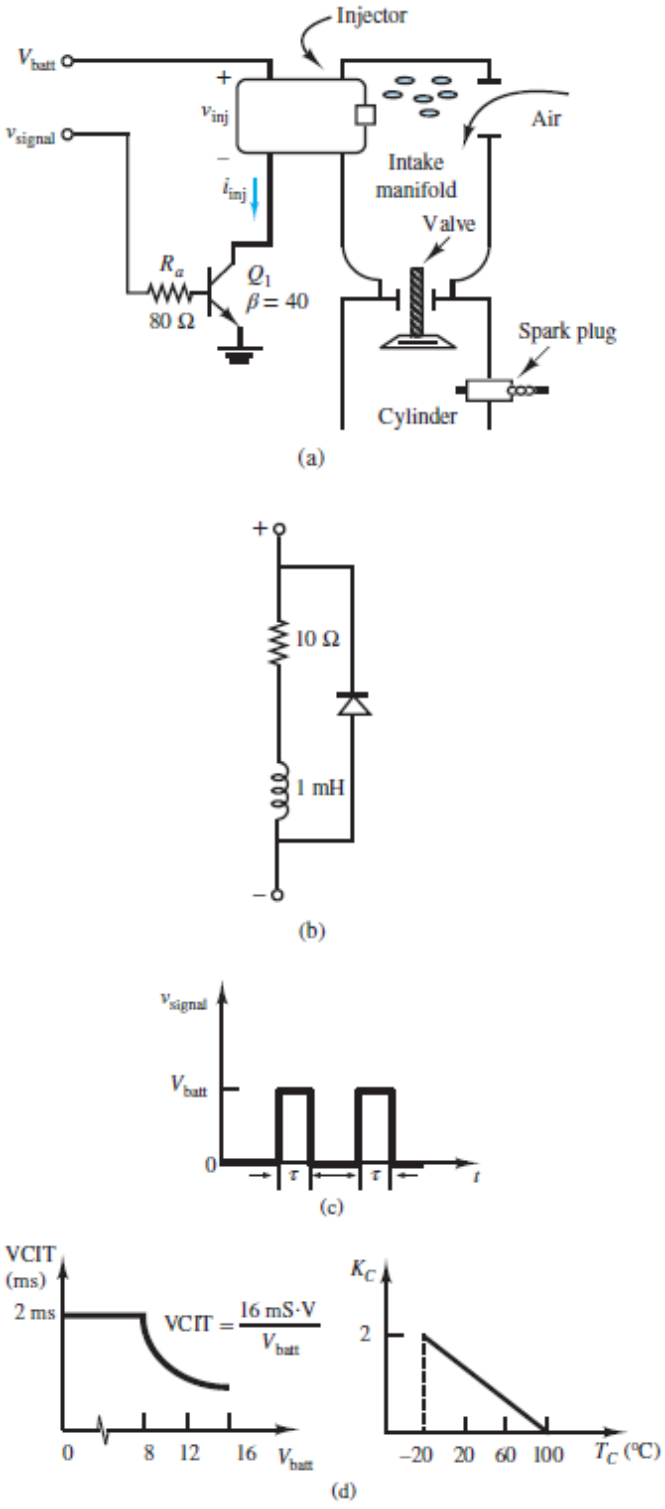


Figure P9.39

9.40 The circuit shown in Figure P9.40 is used to switch a relay under the control of a microcontroller. The relay dissipates 0.5 W at 5 VDC. It

switches on at 3 VDC and off at 1.0 VDC. What is the maximum frequency with which the relay can be switched? The inductance of the relay is 5 mH, and the transistor saturates at 0.2 V, $V_\gamma = 0.8$ V.

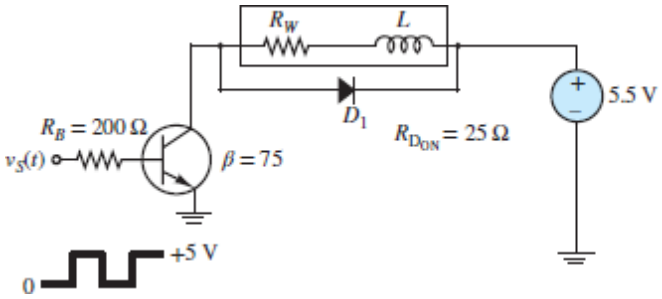


Figure P9.40

9.41 Show that the circuit of Figure P9.41 functions as an OR gate if the output is taken at v_{o1}.

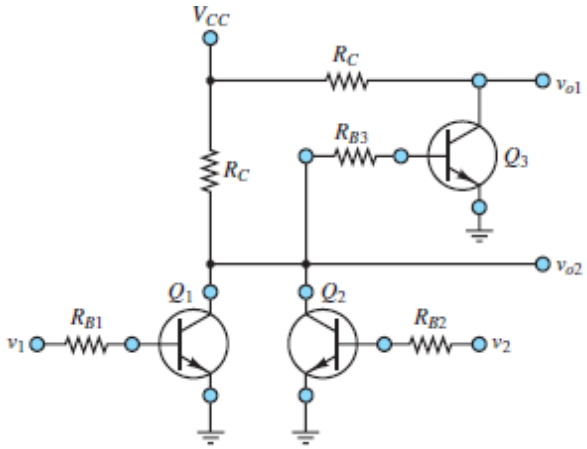


Figure P9.41

9.42 Show that the circuit of Figure P9.41 functions as a NOR gate if the output is taken at v_{o2}.

9.43 Show that the circuit of Figure P9.43 functions as an AND gate if the output is taken at v_{o1}.

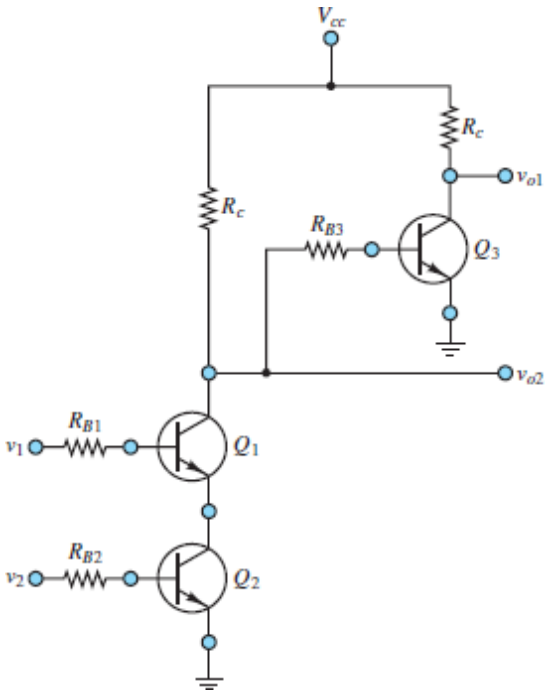


Figure P9.43

9.44 Refer to the circuit in Figure P9.14. The input voltage waveform is shown in Figure P9.44. Determine v_o assuming $\beta = 90$, $R_B = 40$ kΩ, $R_C = 2$ kΩ, and $V_{CC} = 4$ V.

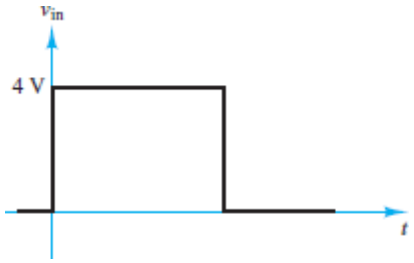


Figure P9.44

9.45 For the circuit shown in Figure P9.14, assume $\beta > 10$, and the minimum value of v_{in} for a high input is 2.0 V. Find the range for resistor R_B that guarantees the transistor is on.

9.46 Figure P9.46 shows a circuit with two transistor inverters connected in series, where

$$R_{1C} = R_{2C} = 10\ \text{k}\Omega \qquad \text{and} \qquad R_{1B} = R_{2B} = 27\ \text{k}\Omega$$

a. Find v_B, v_o, and the state of transistor Q_1 when v_{in} is low (0 V).

b. Find v_B, v_o, and the state of transistor Q_1 when v_{in} is high (5 V).

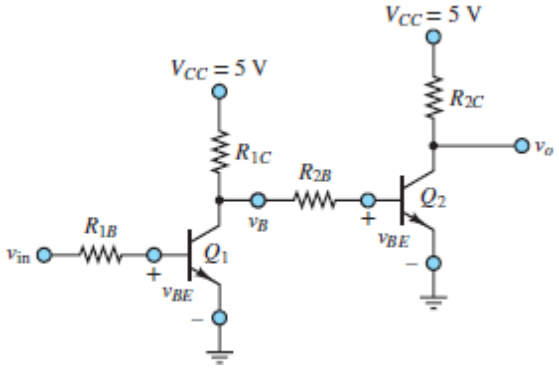


Figure P9.46

9.47 For the circuit shown in Figure P9.47, determine $v_o(t)$, where $v_{\text{in}}(t)$ is as shown in Figure P9.44. Let $\beta = 120$, $R_B = 10\ \text{k}\Omega$, $R_{C1} = R_{C2} = 1\ \text{k}\Omega$, and $V_{CC} = 4$ V.

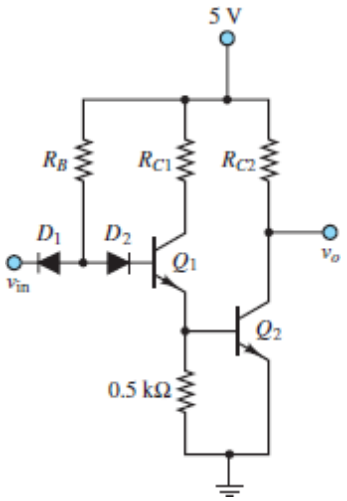


Figure P9.47

9.48 For the circuit shown in Figure P9.48, determine $v_o(t)$, where $v_{\text{in}}(t)$ is as shown in Figure P9.44. Let $\beta = 90$, $R_B = 3\ \text{k}\Omega$, $R_C = 5\ \text{k}\Omega$, and $V_{CC} = 6$

V.

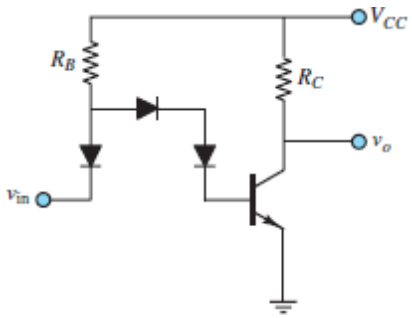


Figure P9.48

9.49 The basic circuit of a TTL gate is shown in Figure P9.49. Determine its logic function.

Page 610

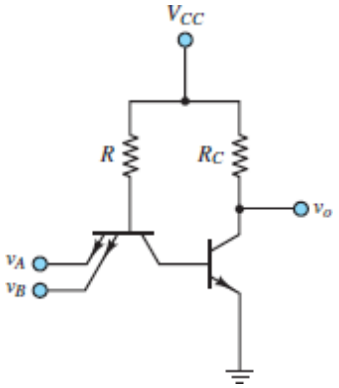


Figure P9.49

9.50 Figure P9.50 shows a three-input TTL NAND gate. Assuming that all the input voltages are high, find v_{B1}, v_{B2}, v_{B3}, v_{C2}, and v_o. Also indicate the operating mode of each transistor.

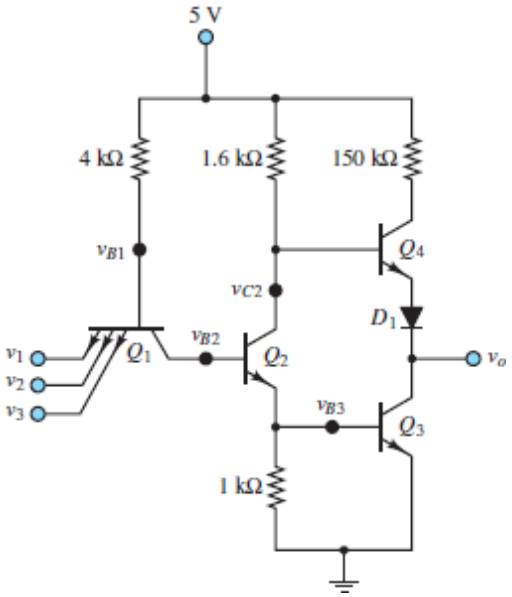


Figure P9.50

9.51 Show that two or more emitter-follower outputs connected to a common load, as shown in Figure P9.51, result in an OR operation; that is, $v_o = v_1 + v_2$. Here, the + sign represents a logical OR operation.

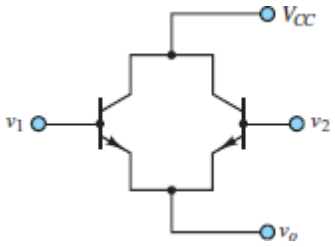


Figure P9.51

9.52 Verify that the circuit of Figure P9.52 is a NAND gate. Assume that a low state is 0.2 V, a high state is 5 V, and $\beta_{min} = 40$.

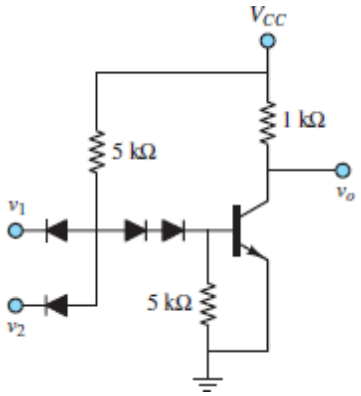


Figure P9.52

[1]Another family of transistors, the field-effect transistors (FETs), are well-modeled as voltage controlled devices. See Chapter 10.

[2]By contrast, a field-effect transistor (FET) is a *unipolar* device. See Chapter 10.

[3]TTL logic values are actually quite flexible, with v_{HIGH} as low as 2.4 V and v_{LOW} as high as 0.8 V.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

C H A P T E R

10

FIELD-EFFECT TRANSISTORS: OPERATION, CIRCUIT MODELS AND APPLICATIONS

Chapter 10 introduces the family of field-effect transistors, or FETs, in which an external electric field is used to control the conductivity of a *channel*, causing the FET to behave either as a voltage-controlled resistor or as a voltage-controlled current source. FETs are the dominant transistor family in today's integrated electronics, and although these transistors come in several different configurations, it is possible to understand the operation of the different devices by focusing principally on one type. Two large families of FETs are the JFETs (junction FETs) and the MOSFETs (metal-oxide semiconducting FETs). Both families can be further classified by a mode (enhancement or depletion) and a channel type (*n* or *p*). In this chapter, the focus is on the enhancement-mode MOSFET with either an *n*-type channel (NMOS) or *p*-type channel (PMOS). The very important CMOS technology, which combines both NMOS and PMOS, is also introduced.

Learning Objectives

Students will learn to...

1. Understand the classification of field-effect transistors. *Section 10.1*.
2. Learn the basic operation of enhancement-mode MOSFETs by understanding their *i-υ* curves and defining equations. *Section 10.2*.
3. Learn how enhancement-mode MOSFET circuits are biased. *Section 10.3*.
4. Understand the concept and operation of FET large-signal amplifiers. *Section 10.4*.
5. Understand the concept and operation of FET switches. *Section 10.5*.
6. Analyze FET switches and digital gates. *Section 10.5*.

10.1 FIELD-EFFECT TRANSISTOR CLASSES

There are three major classes of field-effect transistors:

1. **Enhancement-mode MOSFETs**
2. **Depletion-mode MOSFETs**
3. **Junction field-effect transistors**, or **JFETs**

Each of these classes is comprised of ***n*** and ***p*-channel** devices, where the *n* or *p* designation indicates the nature of the doping in the channel. The acronym MOSFET stands for **metal-oxide semiconductor field-effect transistor**, and although the specific materials and processes used in fabricating transistors has, of course, evolved over time, the acronym continues to be used to describe all enhancement-mode and depletion-mode FETs.

Figure 10.1 shows common circuit symbols for the *n*- and *p*-channel devices within each of the three transistor classes. These transistors have similar behaviors and applications; for the sake of brevity, only the enhancement-mode MOSFET is discussed in detail in this chapter. All the FETs are *unipolar* devices in that current is conducted by only one type of charge carrier, either holes or electrons, unlike BJTs that conduct current using both holes and electrons. Also, whereas both FETs and BJTs are three-terminal devices, the BJTs are asymmetric devices because the collector and emitter are not interchangeable. In concept, the analogous terminals of a FET, known as the drain and source, are symmetric and interchangeable; however, most commercially available FETs are constructed such that the drain and source are not interchangeable, as suggested by Figure 10.2.

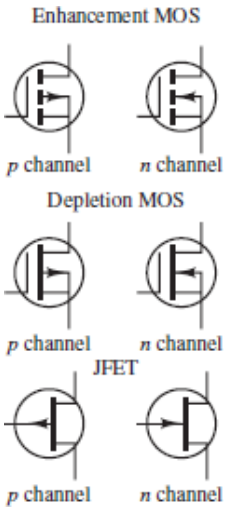


Figure 10.1 Classification of field-effect transistors

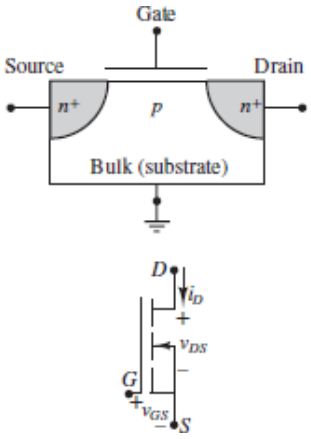


Figure 10.2 The *n*-channel enhancement MOSFET construction and circuit symbol

10.2 ENHANCEMENT-MODE MOSFETS

Figure 10.2 depicts the circuit symbol and the construction of a typical *n*-channel enhancement-mode MOSFET. The device has four regions: the **gate**, the **drain**, the **source**, and the **bulk**.[1] Each of these regions has its own conducting terminal. The bulk and source terminals are often electrically connected, in which case the bulk terminal is not shown in the circuit symbol. The gate consists of a conducting plate separated from the *p*-type bulk by a thin (10^{-9} m) insulating

layer, usually silicon dioxide SiO_2.[2] The drain and source regions are both composed of n^+ material.

Page 613

Consider the case when the gate and source terminals are connected to a reference node and the drain terminal is connected to a positive voltage supply V_{DD} , as shown in Figure 10.3(a). The bulk terminal is also connected to the reference node, by virtue of its connection to the source terminal, and so the pn^+ junction between the bulk and drain is reverse-biased. Obviously, the voltage across the pn^+ junction between the bulk and the source is zero, and thus that junction is also reverse-biased. Thus, a path between drain and source consists of two reverse-biased pn^+ junctions such that the current from drain to source is effectively zero. In this case, the resistance from drain to source is on the order of 10^{12} Ω.

When the voltage from gate to source is zero, the *n*-channel enhancement-mode MOSFET acts as an open-circuit. Thus, enhancement-mode devices are referred to as *normally off* and their channels as *normally open.*

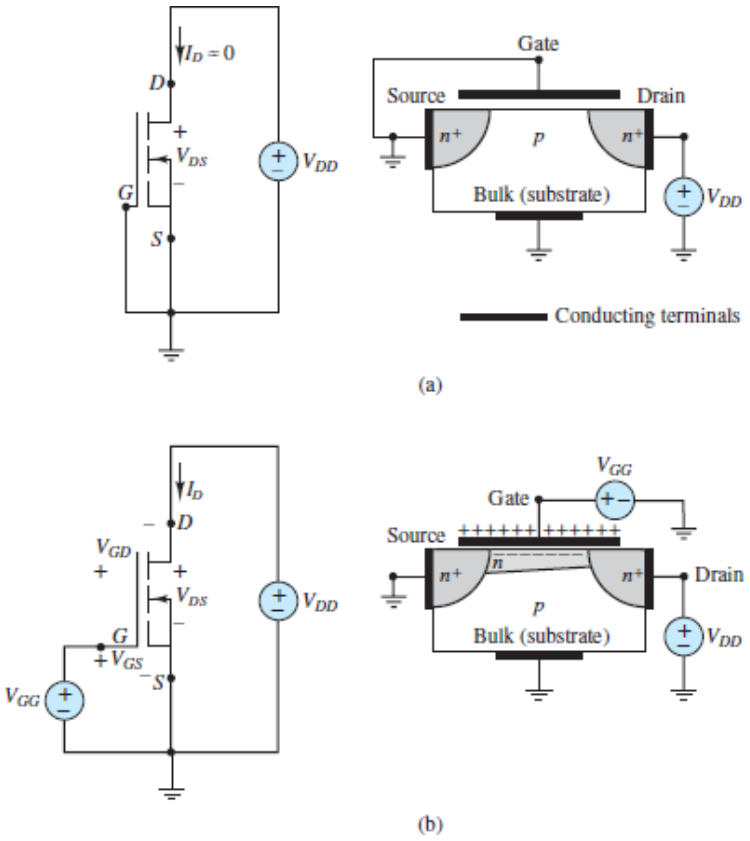


Figure 10.3 Channel formation in NMOS transistor: (a) With zero voltage from gate to source, the source-bulk and bulk-drain junctions are both reverse-biased, and the channel acts as an open-circuit; (b) when a positive gate-to-source voltage is applied, positive majority carriers in the bulk (i.e., holes) are repelled by the gate leaving behind negatively charged atoms. Also, negative majority carriers from the source and drain (i.e., electrons) are drawn toward the gate. The result is a conducting *n*-type channel between the source and drain regions.

Page 614

Suppose now that a positive DC voltage V_{GG} is applied to the gate as shown in Figure 10.3(b). Positive majority charge carriers in the bulk (i.e., holes) are repelled in the region nearest the gate. At the same time, negative majority charge carriers in the source and drain (i.e., electrons) are drawn to the same region. The result is a narrow *n*-type **channel** beneath the insulating layer that separates the gate from the bulk. For a given drain voltage, the higher the gate voltage, the higher the concentration of negative charge carriers in the channel, and the higher its conductivity. The term *enhancement mode* refers to the influence of the gate

voltage in enhancing the conductivity of the channel. The term *field effect* refers to the effect of the electric field from gate to bulk that is associated with the gate voltage.

Depletion-mode devices also exist, in which an externally applied field depletes the channel of charge carriers by reducing the effective channel width. Depletion-mode MOSFETs are normally on (i.e., the channel is conducting) and are turned off (i.e., the channel is not conducting) by an external gate voltage.

Both enhancement- and depletion-mode MOSFETs are available with either *n*- or *p*-type channels. Depending upon the mode and channel type, FETs can be *active high* or *active low* devices, where *high* and *low* refer to the voltage of the gate relative to a common reference. Table 10.1 summarizes these results. *n*- and *p*-channel MOSFETs are referred to as **NMOS** and **PMOS** transistors, respectively.

Table 10.1

Channel Type	Mode	
	Enhancement	Depletion
n	Active high	Active low
p	Active low	Active high

Operating Regions and the Threshold Voltage V_t

When the gate-to-bulk voltage of an NMOS transistor (Figure 10.4) is less than a threshold voltage V_t , a channel will not form between the source and drain. The result is that no current can be conducted from drain to source and the transistor is in the *cutoff region.* A typical value of V_t is between 0.3 and 1.0 V, although it can be significantly larger.

When the gate-to-bulk voltage is greater than the threshold voltage V_t a conducting *n*-type channel is formed. If, as usual, the source and bulk are both connected to a common reference, then the gate-to-bulk voltage is the same as the gate-to-source voltage v_{GS}. If the drain is also connected to the same common reference such that $v_{DS} = 0$, then a channel of uniform thickness and uniform resistance per unit length is formed from drain to source. In this state, known as the *ohmic region*, the channel effectively acts as a variable resistor whose resistance is dictated by the gate voltage. In other words, for a given value of v_{GS}, the channel current i_{DS} is proportional to v_{DS}. This linear relationship between i_D

and v_{DS} is valid for small values of v_{DS}. It is common to introduce the overdrive voltage $v_{OV} = v_{GS} - V_t$, which is the gate-to-source voltage Page 615in excess of what is necessary to create a channel. Note that $v_{OV} > 0$ is another way to write $v_{GS} > V_t$.

$$i_D \propto v_{DS} \quad \text{when} \quad v_{DS} \ll v_{OV} \quad \text{Ohmic region} \tag{10.1}$$

When $v_{GS} > V_t$ and the drain-to-source voltage v_{DS} is no longer small but held at a positive value V_{DD}, the channel is thinner near the drain than near the source, as depicted in Figure 10.3(b). In addition, if $v_{GD} > V_t$, which is equivalent to the requirement that $v_{DS} < v_{OV}$, the channel resistance per unit length is no longer uniform and the channel current i_D is proportional to v_{DS}^2. In this state, the transistor is in the *triode region.*

$$i_D \propto v_{DS}^2 \quad \text{when} \quad v_{DS} < v_{OV} \quad \text{Triode region} \tag{10.2}$$

It is important to realize that the ohmic region is simply one part of the triode region when $v_{DS} \ll v_{OV}$.

Eventually, if v_{DS} is increased to exceed v_{OV}, then $v_{GD} < V_t$ and the channel thickness at the drain goes to zero. However, due to the increase in v_{DS}, the depletion region of the bulk drain junction has expanded sufficiently to take the place of the channel. This condition is often called *channel pinch-off*. Although the channel thickness is now zero, current is still conducted in the channel because the voltage at the drain is large enough to drive mobile electrons in the channel across the depletion region. However, any increase in v_{DS} beyond v_{OV} is confined to the depletion region such that the voltage across the channel length remains constant. The result is that the channel current is independent of v_{DS} and depends only upon v_{OV}. In this state, the transistor is in the *saturation region*.

$$i_D \propto v_{OV}^2 \quad \text{when} \quad v_{DS} > v_{OV} \quad \text{Saturation region} \tag{10.3}$$

The boundary between triode and saturation is $v_{DS} = v_{OV}$. Since $v_{GD} = v_{GS} - V_t - v_{DS} + V_t = v_{OV} - v_{DS} + V_t$ another way to express that boundary is $v_{GD} = V_t$. Thus, the condition $v_{DS} < v_{OV}$ is equivalent to $v_{GD} > V_t$. Likewise, the condition $v_{DS} > v_{OV}$ is equivalent to $v_{GD} < V_t$. The three operating regions and their dependence upon v_{GD} and v_{GS} are depicted in Figure 10.4.

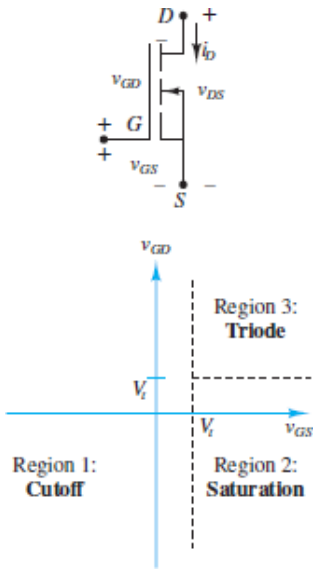


Figure 10.4 Regions of operation of NMOS transistor

Channel Current i_D and the Conductance Parameter K

The ability of the channel to conduct is dependent on various mechanisms, the effects of which are captured in a conductance parameter K, defined as:

$$K = \frac{W}{L}\frac{\mu C_{ox}}{2} \tag{10.4}$$

where W is the cross-sectional width of the channel, L is the channel length, μ is the mobility of the majority channel charge carrier (electrons in n-channel devices, holes in p-channel devices), and C_{ox} is the gate-channel capacitance due to the thin insulating oxide layer. The units of K are A/V^2.

With this definition of the conductance parameter, the relationship between i_D and v_{DS} can be expressed in the various operating regions as listed here. In the cutoff region, $v_{GS} < V_t$:

$$i_D = 0 \qquad \text{Cutoff region} \tag{10.5}$$

In the triode region, $v_{GS} > V_t$ and $v_{GD} > V_t$:

$$i_D = K(2v_{OV} - v_{DS})v_{DS} \qquad v_{GS} > V_t \qquad \text{Triode region} \tag{10.6}$$

When $v_{DS} \ll v_{OV}$ (or equivalently when $v_{GD} \approx v_{GS}$), this expression is approximated by

$$i_D \approx 2K v_{OV} v_{DS} \qquad \text{Ohmic region} \tag{10.7}$$

which is a linear relationship between i_D and v_{DS} such that the transistor acts as a resistor controlled by the overdrive voltage v_{OV}. This property allows transistors to act as resistors in integrated circuit (IC) designs. Other applications of a voltage-controlled resistor are found in tunable (variable-gain) amplifiers and in analog gates.

In the saturation region, $v_{GS} > V_t$ and $v_{GD} < V_t$:

$$i_D \approx K v_{OV}^2 \qquad v_{GS} > V_t \qquad \text{Saturation region} \tag{10.8}$$

Here, the transistor acts as a voltage-controlled current source. This relationship is made more exact by accounting for the **Early effect**, which describes the effect of v_{DS} on the effective length of the channel.

$$i_D = K v_{OV}^2 \left(1 + \frac{v_{DS}}{V_A}\right) \qquad v_{GS} > V_t \qquad \text{Saturation region} \tag{10.9}$$

where V_A is known as the Early voltage. When V_A is large compared to v_{DS}, as is often the case, the Early effect is small and equation 10.9 is well approximated by equation 10.8.

The three regions of operation can also be identified in the characteristic curves shown in Figure 10.5, which can be generated from the circuit of Figure 10.3(b) by varying the gate and drain voltages relative to the source voltage. Notice that for $v_{GS} < V_t$ the transistor is in the cutoff region and $i_D = 0$. The boundary between the saturation and triode regions is indicated by the curve $i_D = K v_{DS}^2$, which is the locus of all points where the slope of the characteristic curve first becomes zero as v_{DS} increases. (If the Early voltage V_A is not negligible, then the slope of the characteristic lines in saturation is not zero, but some small positive constant.) In the saturation Page 617region, the transistor drain current is nearly constant and independent of v_{DS}. In fact, its value is proportional to v_{GS}^2. Finally, in the triode region, the drain current is strongly dependent on v_{GS} and v_{VDS}. As $v_{DS} \rightarrow 0$ the slope of each characteristic curve becomes approximately constant, which is the characteristic of the ohmic region.

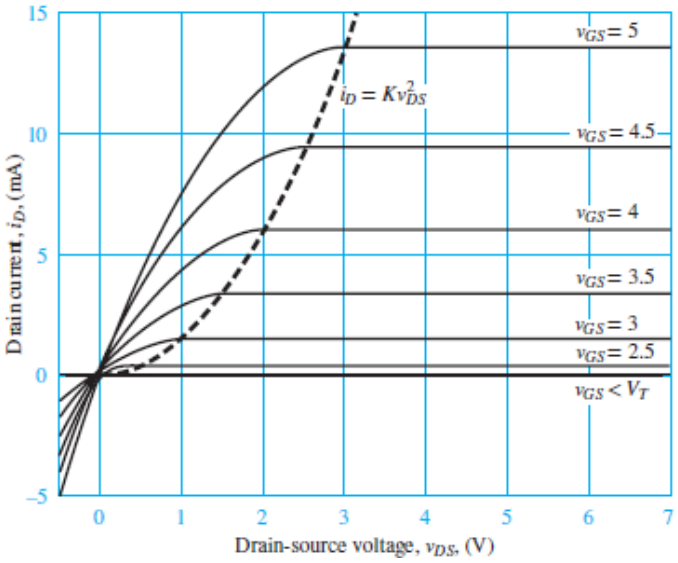


Figure 10.5 Characteristic drain curves for an NMOS transistor with V_t = 2 V and K = 1.5 mA/V^2

Operation of the *P*-channel Enhancement-Mode MOSFET

The operation of a PMOS enhancement-mode transistor is very similar in concept to that of an NMOS device. Figure 10.6 depicts a test circuit and a sketch of the device construction. Note that the roles of the *n*-type and *p*-type materials are reversed and that the charge carriers in the channel are holes, not electrons. Further, the threshold voltage V_t is now negative. However, if v_{GS} is replaced with v_{SG}, v_{GD} with v_{DG}, and v_{DS} with v_{SD}, and $|V_t|$ is used in place of V_t, then the analysis of the device is completely analogous to that of an NMOS transistor. In particular, Figure 10.7 depicts the behavior of a PMOS transistor in terms of the gate-to-drain and gate-to-source voltages, in analogy with Figure 10.4. The resulting equations for the three modes of operation of the PMOS transistor are summarized below:

Cutoff region: when $v_{SG} < |V_t|$.

$$i_D = 0 \qquad \text{Cutoff region} \tag{10.10}$$

Saturation region: when $v_{SG} > |V_t|>$ and $v_{DG} < |V_t|$.

$$i_D \cong K(v_{SG} - |V_t|)^2 \qquad \text{Saturation region} \tag{10.11}$$

Triode region: when $v_{SG} > |V_t|$ and $v_{DG} > |V_t|$.

$$i_D = K\left[2(v_{SG} - |V_t|)v_{SD} - v_{SD}^2\right] \quad \text{Triode or ohmic region} \tag{10.12}$$

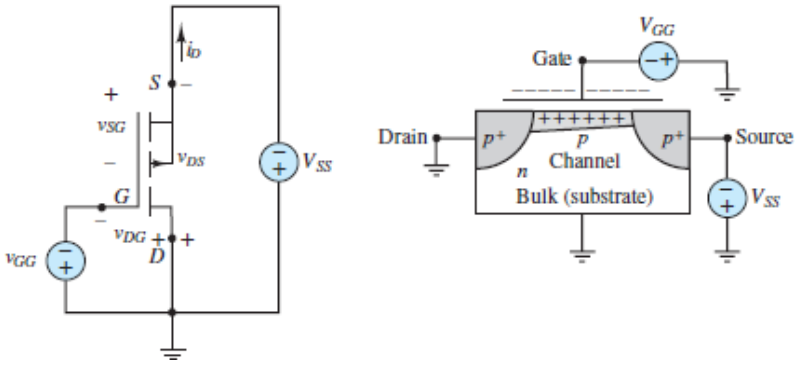


Figure 10.6 The *p*-channel enhancement-mode field-effect transistor (PMOS)

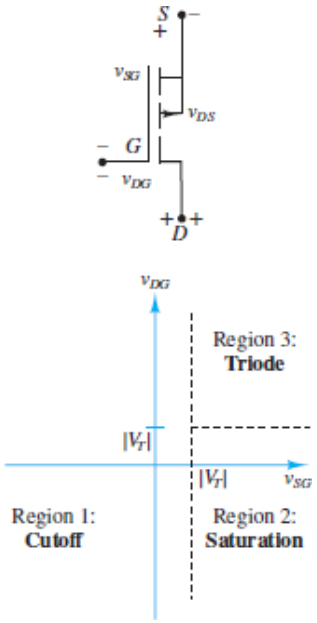


Figure 10.7 Regions of operation of PMOS transistor

EXAMPLE 10.1 Determining the Operating State of a MOSFET

Problem

Determine the operating state of the MOSFET shown in the circuit of Figure 10.8 for the given values of V_{DD} and V_{GG} if the ammeter and voltmeter shown read the following values:

Page 618

a. $V_{GG} = 1$ V; $V_{DD} = 10$ V; $v_{DS} = 10$ V; $i_D = 0$ mA; $R_D = 100\ \Omega$.
b. $V_{GG} = 4$ V; $V_{DD} = 10$ V; $v_{DS} = 2.8$ V; $i_D = 72$ mA; $R_D = 100\ \Omega$.
c. $V_{GG} = 3$ V; $V_{DD} = 10$ V; $v_{DS} = 1.5$ V; $i_D = 13.5$ mA; $R_D = 630\ \Omega$.

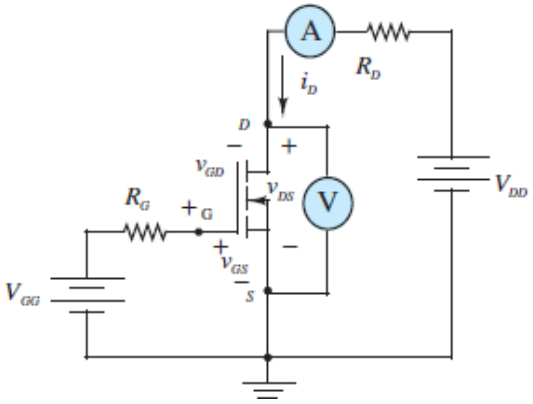


Figure P10.8

Solution

Known Quantities: MOSFET drain resistance; drain and gate supply voltages; MOSFET equations.

Find: MOSFET quiescent drain current i_{DQ} and quiescent drain-source voltage v_{DSQ}.

Schematics, Diagrams, Circuits, and Given Data: $V_t = 2$ V; $K = 18$ mA/V^2.

Assumptions: None.

Analysis: First, notice that the diode indicator in Figure 10.8 points from bulk to channel. These arrows always point from *p* to *n*; thus, the channel is *n*-type and

the transistor is an NMOS. The channel is also marked by a dashed line indicating enhancement mode.

a. Since the drain current is zero, the MOSFET is in the cutoff region. You should verify that both the conditions $v_{GS} < V_t$ and $v_{GD} < V_t$ are satisfied.

b. In this case, $v_{GS} = V_{GG} = 4\text{ V} > V_t$. On the other hand, $v_{GD} = v_G - v_D = 4 - 2.8 = 1.2\text{ V} < V_t$. Thus, the transistor is in the saturation region. We can calculate the drain current to be $i_D = K(v_{GS} - V_t)^2 = 18 \times (4 - 2)^2 = 72$ mA. Alternatively, the drain current can be calculated as:

$$i_D = \frac{V_{DD} - v_{DS}}{R_D} = \frac{10 - 2.8}{0.1\text{ k}\Omega} = 72\text{ mA}$$

c. In the third case, $v_{GS} = V_{GG} = v_G = 3\text{ V} > V_t$. The drain voltage is measured to be $v_{GD} = v_D = 1.5$ V, and therefore $v_{GD} = 3 - 1.5 = 1.5\text{ V} < V_t$. In this case, the MOSFET is in the ohmic, or triode, region. We can now calculate the current to be $i_D = K[2(v_{GS} - V_t)v_{DS} - v_{DS}^2] = 18 \times [2 \times (3 - 2) \times 1.5 - 1.5^2] = 13.5$ mA. The drain current can also be calculated as:

$$i_D = \frac{V_{DD} - v_{DS}}{R_D} = \frac{(10 - 1.5)\text{V}}{0.630\text{ k}\Omega} = 13.5\text{ mA}$$

Page 619

CHECK YOUR UNDERSTANDING

What is the operating state of the MOSFET of Example 10.1 for the following conditions?

$$V_{GG} = \frac{10}{3}\text{ V} \qquad V_{DD} = 10\text{ V} \qquad v_{DS} = 3.6\text{ V} \qquad i_D = 32\text{ mA} \qquad R_D = 200\,\Omega$$

Answer: Saturation

10.3 BIASING MOSFET CIRCUITS

Now that the basic characteristics of enhancement-mode MOSFETs and the means for identifying operating regions are known, it is time to develop systematic procedures for biasing a MOSFET. To bias a transistor is to set its DC operating voltages and currents. This section presents two bias circuits, which are identical to those presented for biasing BJTs. The first, illustrated in Examples 10.2 and 10.3, uses two distinct voltage supplies. This bias circuit is easier to understand, but not very practical—as was discussed in Chapter 9, it is preferable to have a single DC voltage supply and to enable the circuit to regulate its bias point. These features are presented in the second bias circuit, described in Examples 10.4 and 10.5.

EXAMPLE 10.2 MOSFET *Q*-Point Graphical Determination

Problem

Determine the *Q* point for the MOSFET in the circuit of Figure 10.9.

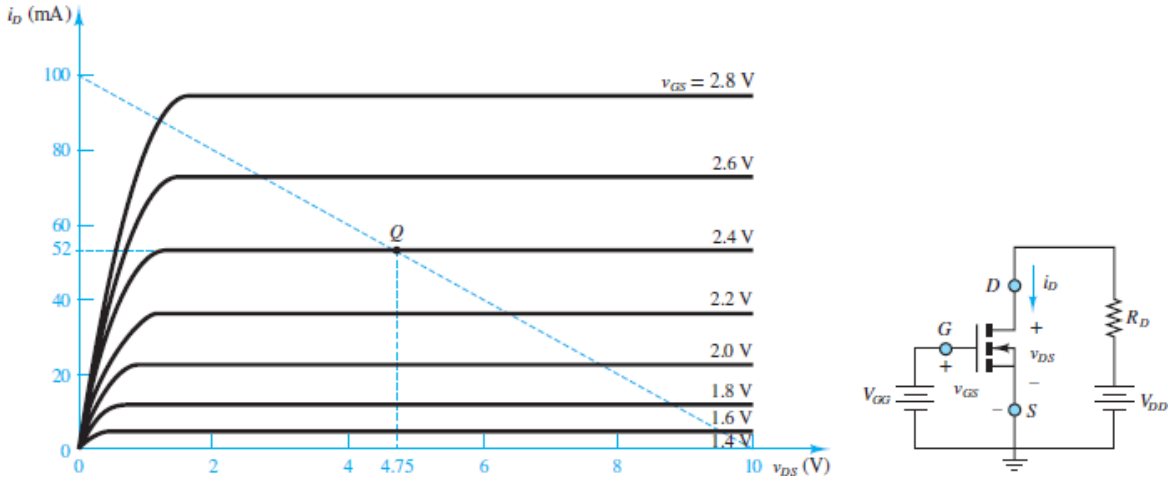


Figure 10.9 An *n*-channel enhancement MOSFET circuit and characteristics

Page 620

Solution

Known Quantities: MOSFET drain resistance; drain and gate supply voltages; MOSFET drain curves.

Find: MOSFET quiescent drain current i_{DQ} and quiescent drain-source voltage v_{DSQ}.

Schematics, Diagrams, Circuits, and Given Data: $V_{GG} = 2.4$ V; $V_{DD} = 10$ V; $R_D = 100\ \Omega$.

Assumptions: Use the characteristic curves of Figure 10.9.

Analysis: First, notice that the diode indicator in Figure 10.9 points from bulk to channel. These arrows always point from *p* to *n*; thus, the channel is *n*-type and the transistor is an NMOS. The channel is also marked by a dashed line indicating enhancement mode.

To determine the *Q* point, apply KVL and Ohm's law to the drain circuit.

$$V_{DD} = R_D i_D + v_{DS}$$
$$10 = 100\, i_D + v_{DS}$$

This equation is plotted as a dashed line on the drain curves of Figure 10.9. The drain current axis intercept is equal to $V_{DD}/R_D = 100$ mA and the drain-source voltage axis intercept is equal to $V_{DD} = 10$ V. The *Q* point is the intersection of the load line with the $v_{GS} = 2.4$ V curve. Thus, $i_{DQ} = 52$ mA and $v_{DSQ} = 4.75$ V.

Comments: The determination of a *Q* point for a MOSFET is easier than for a BJT because the gate current is essentially zero.

EXAMPLE 10.3 MOSFET *Q*-Point Calculation

Problem

Use the MOSFET characteristic curves shown in Figure 10.9 to determine the *Q* point for the conditions listed below.

Solution

Known Quantities: MOSFET drain resistance; drain and gate supply voltages; MOSFET equations.

Find: MOSFET quiescent drain current i_{DQ} and quiescent drain-source voltage v_{DSQ}.

Schematics, Diagrams, Circuits, and Given Data: $V_{GG} = 2.4$ V; $V_{DD} = 10$ V; $V_t = 1.4$ V; $K = 48.5$ mA/V^2; $R_D = 100\ \Omega$.

Assumptions: None.

Analysis: The gate supply V_{GG} ensures that $v_{GSQ} = V_{GG} = 2.4$ V. Thus, $v_{GSQ} > V_t$. Assume that the MOSFET is in the saturation region, and proceed to use equation 10.8 to calculate the drain current:

$$i_{DQ} = K(v_{GS} - V_t)^2 = 48.5 \times 10^{-3}(2.4 - 1.4)^2 = 48.5 \text{ mA}$$

Page 621

Apply KVL and Ohm's law to the drain circuit to calculate the quiescent drain-to-source voltage.

$$v_{DSQ} = V_{DD} - R_D i_{DQ} = 10 - 100 \times 48.5 \times 10^{-3} = 5.15 \text{ V}$$

The conditions required for saturation (region 2) are $v_{GS} > V_t$ and $v_{GD} < V_t$. The first condition is clearly satisfied. The second can be verified by recognizing that:

$$v_{GD} = v_{GS} + v_{SD} = v_{GS} - v_{DS} = -2.75 \text{ V}$$

Clearly, the condition $v_{GD} < V_t$ is also satisfied, and the MOSFET is indeed operating in the saturation region.

EXAMPLE 10.4 MOSFET Self-Bias Circuit

Problem

Figure 10.10(a) depicts a practical self-bias circuit for a MOSFET. Determine the Q point for the MOSFET by choosing R_S such that v_{DSQ} = 8 V. The values of all other parameters are given.

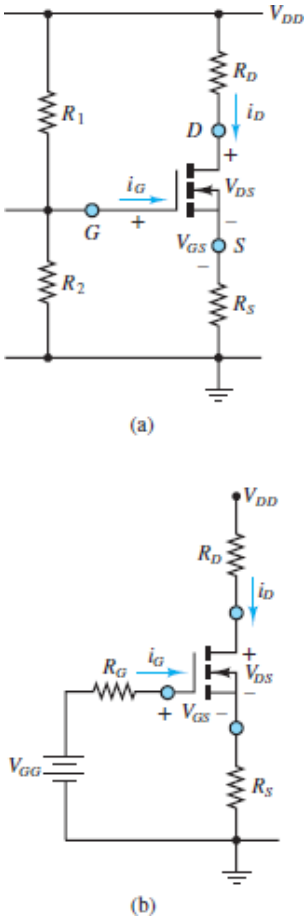


Figure 10.10 (a) Self-bias circuit; (b) equivalent circuit for part (a)

Solution

Known Quantities: MOSFET drain and gate resistances; drain supply voltage; MOSFET parameters V_t and K; desired drain-to-source voltage v_{DSQ}.

Find: MOSFET quiescent gate-source voltage v_{GSQ} , quiescent drain current i_{DQ} , and quiescent drain-source voltage v_{DSQ} .

Schematics, Diagrams, Circuits, and Given Data: V_{DD} = 30 V; R_D = 10 k Ω; R_1 = R_2 = 1.2 MΩ; V_t = 4 V; K = 0.2188 mA/V^2; v_{DSQ} = 8 V.

Assumptions: Assume operation in the saturation region.

Analysis: Apply Thévenin's theorem and voltage division to determine the equivalent network seen by the gate in Figure 10.10(a). The result is shown in Figure 10.10(b), where

$$V_{GG} = \frac{R_2}{R_1 + R_2} V_{DD} = 15 \text{ V} \qquad R_G = R_1 \| R_2 = 600 \text{ k}\Omega$$

Since the conductance parameter K is given in units of mA/V^2, assume that all currents are expressed in milliamps and all resistances in kilo-ohms. Applying KVL around the equivalent gate circuit of Figure 10.10(b) yields:

$$v_{GSQ} + i_{GQ}R_G + i_{DQ}R_S = V_{GG} = 15 \text{ V}$$

Since $i_{GQ} = 0$, due to the infinite input resistance of the MOSFET, the gate equation simplifies to:

$$v_{GSQ} + i_{DQ}R_S = V_{GG} = 15 \text{ V} \qquad \text{(a)}$$

The drain circuit equation is

$$v_{DSQ} + i_{DQ}R_D + i_{DQ}R_S = V_{DD} = 30 \text{ V} \qquad \text{(b)}$$

Page 622

Use equation 10.8 to obtain:

$$i_{DQ} = K(v_{GS} - V_t)^2 \qquad \text{(c)}$$

Use equation (a) to substitute for $i_{DQ}R_S$ in equation (b). The result is

$$V_{DD} = i_{DQ}R_D + v_{DSQ} + V_{GG} - v_{GSQ}$$

Since $V_{GG} = V_{DD}/2$ this equation can be rewritten as

$$i_{DQ} = \frac{1}{R_D}\left(\frac{V_{DD}}{2} - v_{DSQ} + v_{GSQ}\right)$$

Substitute the above equation for i_{DQ} into equation (c) to obtain a quadratic equation that can be solved for v_{GSQ} using the desired value of v_{DSQ}.

$$\frac{1}{R_D}\left(\frac{V_{DD}}{2} - v_{DSQ} + v_{GSQ}\right) = K(v_{GSQ} - V_t)^2$$

$$Kv_{GSQ}^2 - 2KV_t v_{GSQ} + KV_t^2 - \frac{1}{R_D}\left(\frac{V_{DD}}{2} - v_{DSQ}\right) - \frac{1}{R_D}v_{GSQ} = 0$$

$$v_{GSQ}^2 - \left(2V_t + \frac{1}{KR_D}\right)v_{GSQ} + V_t^2 - \frac{1}{KR_D}\left(\frac{V_{DD}}{2} - v_{DSQ}\right) = 0$$

$$v_{GSQ}^2 - 8.457\,v_{GSQ} + 12.8 = 0$$

The two solutions for the above quadratic equation are

$$v_{GSQ} = 6.48 \text{ V} \quad \text{and} \quad v_{GSQ} = 1.97 \text{ V}$$

Only the first of these two values is acceptable for operation in the saturation region, since the second root corresponds to a value of v_{GS} lower than the threshold voltage $V_t = 4$ V. Substitute the first value into equation (c) to compute the quiescent drain current:

$$i_{DQ} = 1.35 \text{ mA}$$

Use this value in the gate circuit equation (a) to compute the solution for the source resistance:

$$R_S = 6.32 \text{ k}\Omega$$

Comments: Two mathematical solutions are found from the quadratic equation. Only one of the solutions satisfies the physical constraints in the problem.

EXAMPLE 10.5 Analysis of a MOSFET Amplifier

Problem

Determine the gate and drain-source voltage and the drain current for the MOSFET amplifier of Figure 10.11.

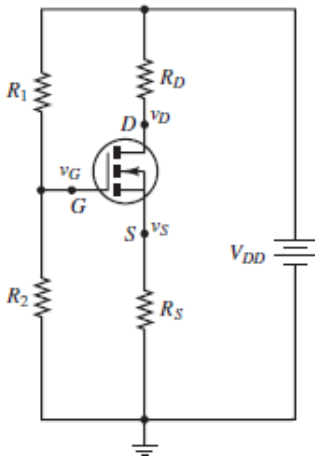


Figure 10.11

Solution

Known Quantities: Drain, source, and gate resistors; drain supply voltage; MOSFET parameters.

Find: v_{GS}; v_{DS}; i_D .

Page 623

Schematics, Diagrams, Circuits, and Given Data: $R_1 = R_2 = 1$ MΩ; $R_D = 6$ k Ω; $R_S = 6$ kΩ; $V_{DD} = 10$ V; $V_t = 1$ V; $K = 0.5$ mA/V^2.

Assumptions: The MOSFET is operating in the saturation region.

Analysis: Since the gate current is zero, the gate voltage can be computed by applying voltage division to the virtual series connection between resistors R_1 and R_2.

$$v_G = \frac{R_2}{R_1 + R_2} V_{DD} = \frac{1}{2} V_{DD} = 5 \text{ V}$$

Apply KVL and Ohm's law to write:

$$v_{GS} = v_G - v_S = v_G - R_S i_D = 5 - 6\, i_D$$

Also, since $R_D = R_S$ Ohm's law requires the voltage drop across R_D to be equal to the voltage drop across R_S. Consequently, v_D must be greater than $V_{DD}/2$ and v_{GD}

must be negative. Thus, $v_{GD} < V_t$, the transistor is operating in saturation and the drain current i_D is approximated by equation 10.8.

$$i_D = K(v_{GS} - V_t)^2 = 0.5(5 - 6i_D - 1)^2$$

or

$$36i_D^2 - 50i_D + 16 = 0$$

with solutions

$$i_D = 0.89 \text{ mA} \quad \text{and} \quad i_D = 0.5 \text{ mA}$$

Only one of these two values will satisfy the requirement $v_{GS} > V_t$. For $i_D = 0.89$ mA, $v_{GS} = 5 - 6i_D = -0.34$ V. For $i_D = 0.5$ mA, $v_{GS} = 5 - 6i_D = 2$ V. Thus, the only physically viable solution is

$$i_D = 0.5 \text{ mA} \quad v_{GS} = 2 \text{ V}$$

The corresponding drain voltage is therefore found to be

$$v_D = v_{DD} - R_D i_D = 10 - 6i_D = 7 \text{ V}$$

And

$$v_{DS} = v_D - v_S = v_D - i_D R_S = 7 - 3 = 4 \text{ V}$$

Comments: These results can be used to verify that the transistor is operating in saturation: $v_{GS} = 2 > V_t$ and $v_{GD} = v_{GS} - v_{DS} = 2 - 4 = -2 < V_t$.

CHECK YOUR UNDERSTANDING

Determine the operating region of the MOSFET of Example 10.2 when $v_{GS} = 3.5$ V.

Answer: The MOSFET is in the ohmic region.

CHECK YOUR UNDERSTANDING

Find the lowest value of R_D for the MOSFET of Example 10.3 that will place the MOSFET in the triode region.

Answer: $\approx 185.6\ \Omega$

CHECK YOUR UNDERSTANDING

Determine the appropriate value of R_S if we wish to move the operating point of the MOSFET of Example 10.4 to v_{DSQ} = 12 V. Also find the values of v_{GSQ} and i_{DQ}. Are these values unique?

Answer: The answer is unique. One of the two solutions is $v_{GS} = 2.42$ V, but this value is less than V_t so it is not valid. The other solution is $v_{GS} = 6.03$ V with $R_S = 9.9\ \text{k}\Omega$ and $i_D = 0.9$ mA.

10.4 MOSFET LARGE-SIGNAL AMPLIFIERS

The objective of this section is to illustrate how a MOSFET can be used as a large-signal amplifier, in applications similar to those illustrated in Chapter 9 for bipolar transistors. Equation 10.8 describes the approximate saturation region relationship between the drain current and gate-source voltage for the MOSFET in a large-signal amplifier application. Appropriate biasing, as explained in the preceding section, is used to ensure that the MOSFET is operating in saturation.

$$i_D \approx K(v_{GS} - V_t)^2 \tag{10.13}$$

MOSFET amplifiers are commonly found in one of two configurations: Page 625*common-source* and *source-follower*. Figure 10.12 depicts a basic common-source configuration. When the MOSFET is in saturation, this amplifier is essentially a voltage-controlled current source (VCCS), in which the drain current is controlled by the gate voltage. Thus, in saturation, the voltage v_o across the load R_o is

$$v_o = R_o i_D \approx R_o K(v_{GS} - V_t)^2 = R_o K(V_G - V_t)^2 \tag{10.14}$$

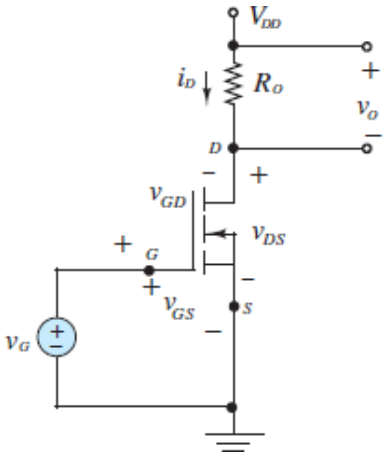


Figure 10.12 Common-source MOSFET amplifier

Notice that as v_o increases with V_G, the drain voltage relative to reference decreases. The result is that the drain voltage is inversely related to the gate voltage for a common-source amplifier.

A source-follower amplifier is shown in Figure 10.13(a). Note that the load is now connected between the source and reference. Once again, the load voltage is given by the expression $v_o = R_o i_D$, where

$$i_D \approx K(v_{GS} - V_t)^2 = K(v_G - v_o - V_t)^2 = K(\Delta v - R_o i_D)^2 \tag{10.15}$$

where $\Delta v = v_G - V_t$. Expand the quadratic term to obtain:

$$i_D = K(\Delta v - R_o i_D)^2 = K\Delta v^2 - 2K\Delta v R_o i_D + R_o^2 i_D^2 \tag{10.16}$$

This expression can be rearranged to yield:

$$i_D^2 - \frac{1}{R_o^2}(2K\Delta v R_o + 1) i_D + \frac{K}{R_o^2}\Delta v^2 = 0 \tag{10.17}$$

Use the quadratic equation to solve for the load current:

$$i_D = \frac{1}{2}\left\{-b \pm \sqrt{b^2 - 4c}\right\} \tag{10.18}$$

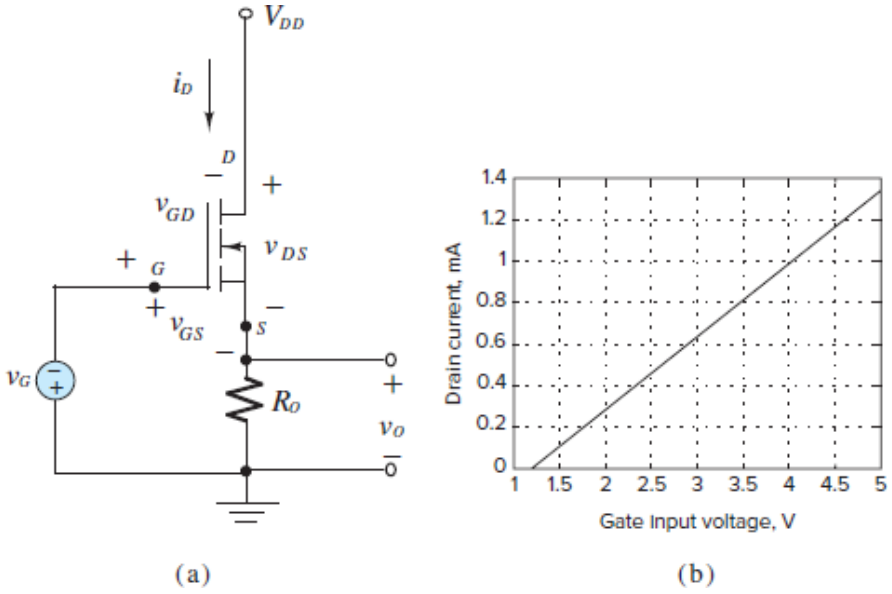


Figure 10.13 (a) Source-follower MOSFET amplifier; (b) drain current response of a source-follower amplifier for a 100-Ω load when K = 0.018 A/V^2 and V_t = 1.2 V

Page 626

where

$$-b = \frac{2K\Delta v R_o + 1}{R_o^2} \quad \text{and} \quad c = \frac{K\Delta v^2}{R_o^2} \tag{10.19}$$

Figure 10.13(b) depicts the linear drain current response of the source-follower MOSFET amplifier when the gate voltage varies between the threshold voltage and 5 V for a 100-Ω load when K = 0.018 A/V^2 and V_t = 1.2 V. Notice that v_o increases linearly with V_G in a noninverted fashion for a source-follower amplifier.

EXAMPLE 10.6 Using a MOSFET as a Current Source for Battery Charging

Problem

Analyze the two battery charging circuits shown in Figure 10.14 (a) and (b). Use the transistor parameters to determine the range of required gate voltages v_G to provide a variable charging current up to a maximum of 0.1 A. Assume that the terminal voltage of a discharged battery is 9 V, and of a charged battery is 10.5 V.

Solution

Known Quantities: Transistor large-signal parameters, lithium (Li) battery nominal voltage.

Find: V_{DD}, v_G , range of gate voltages leading to a maximum charging current of 0.1 A.

Schematics, Diagrams, Circuits, and Given Data. Figure 10.14(a) and (b). V_t = 1.2 V; K = 18 mA/V^2, V_B = 9 V discharged, V_B = 10.5 V charged.

Assumptions: Assume that the MOSFETs are operating in the saturation region.

Analysis: The conditions for the MOSFET to be in the saturation region are $v_{GS} > V_t$ and $v_{GD} < V_t$. The first condition is satisfied when $v_{GS} \leq 1.2$ V. The second condition is satisfied Page 627for the entire range of battery voltages as long as $V_{DD} > 10.5 + v_{GS} - V_t$. Assuming both conditions are satisfied such that the transistor is operating in saturation the drain current is

$$i_D = K(v_{GS} - V_t)^2 = 0.018 \times (v_{GS} - 1.2)^2 \text{ A}$$

The plot of Figure 10.14(c) depicts the battery charging (drain) current as a function of v_{GS}. The maximum charging current of 100 mA is generated at $v_{GS} \approx$ 3.5 V.

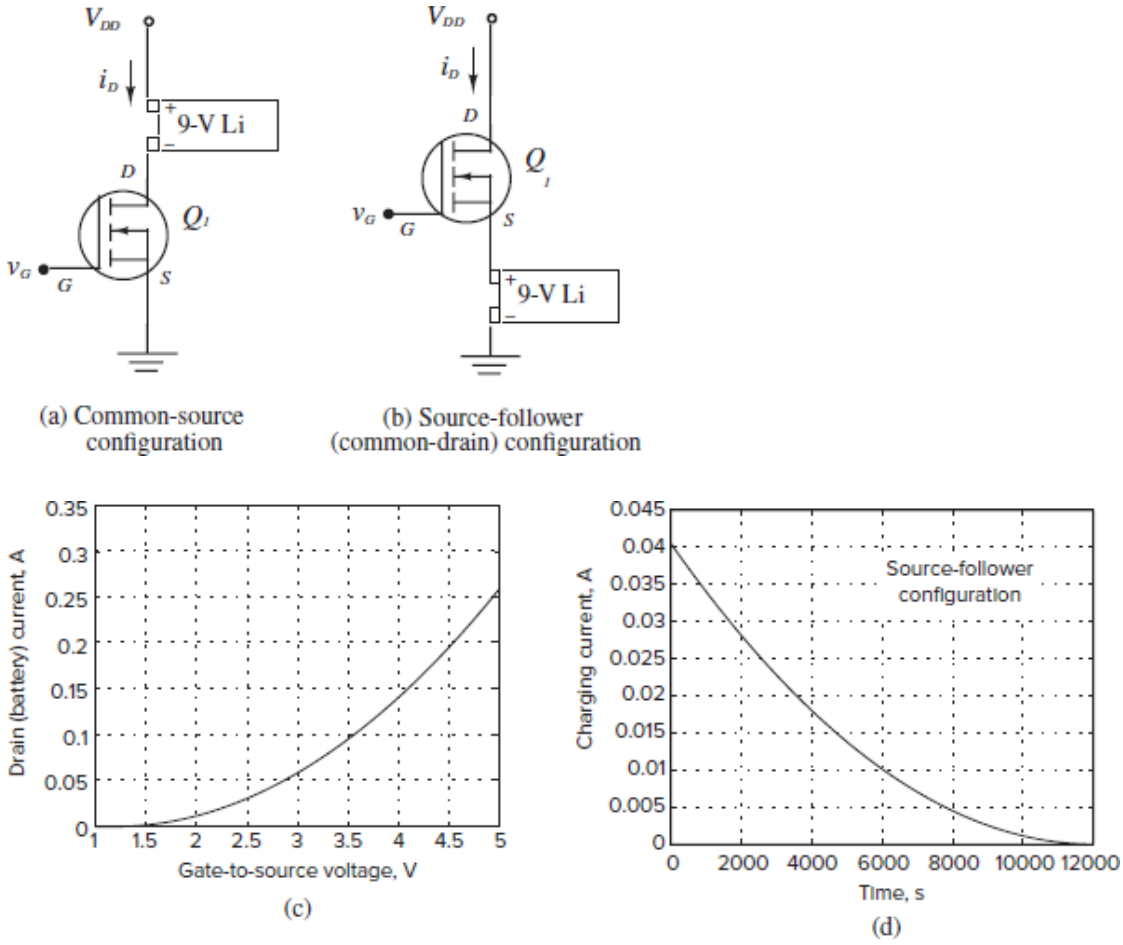


Figure 10.14 MOSFET battery charger

a. **Common-source configuration:** In this configuration $v_S = 0$ and so $v_G > 1.2$ V is required to begin charging the battery. The maximum charging current occurs when $v_G \approx 3.5$ V so to guarantee continuous operation in saturation mode it is necessary to set $V_{DD} > 10.5 + 3.5 - 1.2 = 12.8$ V.

b. **Source-follower (common-drain) configuration:** In this configuration $v_D = V_{DD}$ and $v_{GS} > 1.2$ V is required to begin charging the battery. The maximum source (battery) voltage is 10.5 V. Thus, $v_G > 11.7$ V is required to maintain current until the battery is fully charged. Again, the maximum charging current occurs when $v_{GS} \approx 3.5$ V so to guarantee continuous operation in saturation mode it is necessary to set $V_{DD} > 10.5 + 3.5 - 1.2 = 12.8$ V, which is the same requirement for the source-follower configuration. Assuming the battery is initially discharged so that $V_B = 9$ V

and the initial gate voltage is set to v_G = 11.7 V the initial charging current is

$$i_D = K(v_{GS} - V_t)^2 = K(v_G - V_B - V_t)^2 = 0.018 \times (11.7 - 9 - 1.2)^2$$
$$= 0.0405 \text{ A}$$

Also assume that during charging the battery voltage increases linearly from 9 to 10.5 V over a period of 20 min to calculate the charging current as the battery voltage increases. Note that when the battery is charged, v_{GS} is no longer larger than V_t and the transistor is cut off. A plot of the drain (charging) current as a function of time is shown in Figure 10.14(d). Note that the charging current tapers to zero as the battery voltage increases.

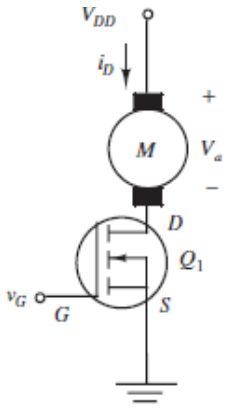
Comments: In the circuit of part b, the battery voltage is not likely to increase linearly. The voltage rise will taper off as the battery approaches full charge. In practice, the charging process will take longer than projected in Figure 10.14(d).

Page 628

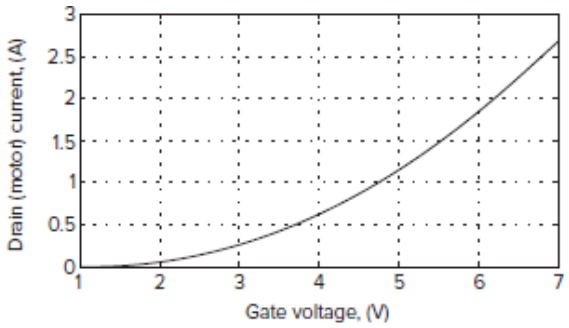
EXAMPLE 10.7 MOSFET DC Motor Drive Circuit

Problem

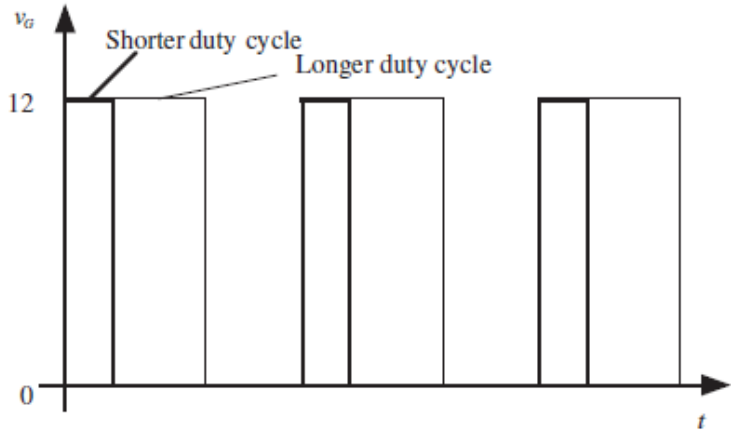
Design a common-source MOSFET driver for the **Lego® 9V XL motor, model 8882**. Figure 10.15 shows the driver circuit. Assume that the motor has a maximum (stall) current of 2,020 mA and a minimum startup current of 110 mA. The aim of the circuit is to control the current to the motor (and therefore the motor torque, which is proportional to the current) via the gate voltage.



(a) Common-source MOSFET DC motor driver circuit



(b) Drain current–gate voltage curve for the MOSFET in saturation



(c) Pulse-width modulation (PWM) gate voltage waveforms

Figure 10.15 DC motor drive circuit

Solution

Known Quantities: Transistor large-signal parameters, component values.

Find: R_1 and R_2, and the value of v_G needed to drive the motor.

Schematics, Diagrams, Circuits, and Given Data: Figure 10.15. $V_t = 1.2$ V; $K = 0.08$ A/V^2.

Assumptions: Assume that the MOSFET operates in the saturation region.

Analysis: The conditions for the MOSFET to be in the saturation region are $v_{GS} > V_t$ and $v_{GD} < V_t$. The first condition is satisfied whenever $v_{GS} = v_G \leq 1.2$ V. Thus the transistor will first begin to conduct when $v_G = 1.2$ V. The second condition is satisfied for the nominal 9 V motor voltage as long as $V_{DD} > 9 + v_{GS} - V_t$. Assuming both conditions are satisfied such that the transistor is operating in saturation the drain current is

$$i_D = K(v_{GS} - V_t)^2 = 0.08 \times (v_G - 1.2)^2 \text{ A}$$

The plot of Figure 10.15(b) depicts the DC motor (drain) current as a function of the gate voltage. The maximum current of 2,020 mA can be generated with a gate voltage of approximately 6.2 V. It would take approximately 2.4 V at the gate to generate the minimum required current of 110 mA.

Page 629

Comments: This circuit could be quite easily implemented in practice to drive the motor with a signal from a microcontroller. In practice, instead of trying to output an analog voltage, a microcontroller is better suited to the generation of a digital (on/off) signal. For example, the gate drive signal could be a *pulse-width modulated* (PWM) 0−12 V pulse train, in which the ratio of the *on* time to the period of the waveform time is called the *duty cycle*. Figure 10.15(c) depicts the possible appearance of a digital PWM gate voltage input.

CHECK YOUR UNDERSTANDING

What is the range of duty cycles needed to cover the current range of the Lego motor?

Answer: 20 to 52 percent

10.5 CMOS TECHNOLOGY AND MOSFET SWITCHES

The objective of this section is to illustrate how a MOSFET can be used as an analog or a digital switch (or gate). Most MOSFET switches are based upon a **complementary MOS**, or **CMOS**, technology, which makes use of the complementary characteristics of PMOS and NMOS devices to enable energy-efficient integrated circuits. Further, CMOS circuits are easily fabricated and require only a single supply voltage, which is a significant advantage.

Digital Switches and Gates

Consider the **CMOS inverter** of Figure 10.16, in which two enhancement-mode transistors, one PMOS and one NMOS, are connected to a single supply voltage (V_{DD} , relative to the reference node). Their gates share a common input voltage v_{in} . This device is known as an inverter because the output voltage $v_o \approx V_{DD}$ whenever $v_{\text{in}} \approx 0$, and vice versa. When used as a logic device, a voltage close to V_{DD} is known as a *logic high*, or a 1, whereas a voltage close 0 V is known as a *logic low*, or a 0.

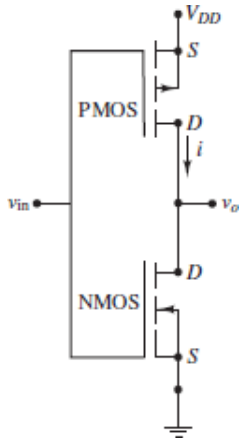


Figure 10.16 CMOS inverter

The objective is to use v_{in} to drive one of the two transistors into the ohmic region and at the same time drive the other transistor into cutoff. The result would be that $i = 0$ due to the transistor in cutoff and that, in turn, would result in $v_{DS} = 0$ across the other transistor. Thus, the circuit would behave like an open-circuit in

series with a short-circuit. When the NMOS transistor is acting as the short-circuit, $v_o \approx 0$. When the PMOS transistor is acting as the short-circuit, $v_o \approx V_{DD}$.

Consider the case when $v_{\text{in}} \approx V_{DD}$ is acting as a logic high input and assume that $V_{DD} \gg V_t$. The operating mode equations for the NMOS transistor are

$$v_{GS} = v_{\text{in}} \quad \text{and} \quad v_{GD} = v_{\text{in}} - v_o \quad \text{NMOS}$$

The operating mode equations for the PMOS transistor are

$$v_{SG} = V_{DD} - v_{\text{in}} \quad \text{and} \quad v_{DG} = v_o - v_{\text{in}} \quad \text{PMOS}$$

With $v_{\text{in}} \approx V_{DD}$ the PMOS transistor will be in cutoff since

$$v_{SG} < V_t \quad \text{(PMOS cutoff)}$$

With the PMOS is cutoff the drain current through the NMOS will be zero. Also, with $v_{\text{in}} \approx V_{DD}$ the operating conditions for the NMOS transistor are

$$v_{GS} = V_{DD} \gg V_t \quad \text{and} \quad v_{GD} = V_{DD} - v_o = V_{DD} \gg V_t \quad \text{(NMOS ohmic)}$$

where $v_o = 0$ because for the NMOS transistor $v_{DS} = i_D r_{DS} = 0$. The overall result can be represented as ideal open and closed switches, respectively, as shown in Figure 10.17(a).

The same analysis can be applied to the case when v_{in} is a logic low. In this case, the PMOS transistor sees a large negative gate-to-source voltage and a channel is formed in the triode state. On the other hand, the NMOS transistor sees a gate-to-source voltage near zero such that no channel is formed in the cutoff state. Figure 10.17(b) represents this situation in terms of ideal switches. Note that this circuit does not require the transistors to be biased. Also, note that the drain current i_D is zero in both cases such that a CMOS inverter consumes very little power.

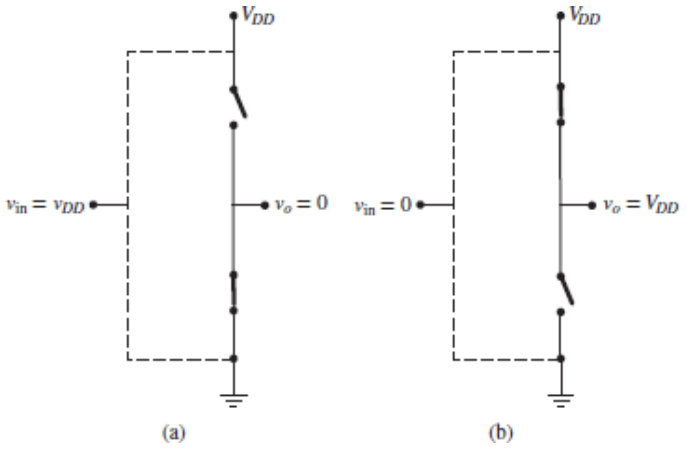


Figure 10.17 CMOS inverter approximated by ideal switches: (a) When v_{in} is high, v_o is tied to ground; (b) when v_{in} is low, v_o is tied to V_{DD}.

Analog Switches

A common analog gate employs a FET and takes advantage of the fact that its current can be bidirectional in the ohmic region. Recall that a MOSFET operating in the ohmic state acts very much as a linear resistor. For example, for an NMOS enhancement-mode transistor the conditions for the ohmic state can be defined as:

$$v_{GS} > V_t \quad \text{and} \quad |v_{DS}| \leq \frac{1}{4}(v_{GS} - V_t) \tag{10.20}$$

As long as the NMOS satisfies these conditions, it reasonably acts as a simple linear resistor with a channel resistance of:

$$r_{DS} = \frac{1}{2K(v_{GS} - V_t)} \tag{10.21}$$

Thus, the drain current can be simply represented as:

$$i_D \approx \frac{v_{DS}}{r_{DS}} \quad \text{for} \quad |v_{DS}| \leq \frac{1}{4}(v_{GS} - V_t) \quad \text{and} \quad v_{GS} > V_t \tag{10.22}$$

The most important feature of the MOSFET operating in the ohmic region is that it acts as a voltage-controlled resistor, with the gate-source voltage v_{GS} controlling the channel resistance r_{DS}. The use of the MOSFET as a switch in the ohmic region consists of providing a gate-source voltage that can either hold the MOSFET in the cutoff region ($v_{GS} \leq V_t$) or the ohmic region ($v_{GS} \gg v_{in}$).

Consider the circuit shown in Figure 10.18, where v_G can be varied externally and v_{in} is an analog input signal source that is to be connected to the load R_o at some appropriate time. When $v_{GS} \leq V_t$, the FET is in the cutoff region and acts as an open-circuit. If $v_{GS} \leq V_t$ such that the MOSFET is in the ohmic region, then $v_{GD} > V_t$ and the transistor acts as a linear resistance r_{DS} . If $r_{\text{DS}} \ll R_o$, then, by voltage division, $v_o \approx v_{\text{in}}$.

MOSFET analog switches are usually produced in integrated-circuit (IC) form and denoted by the symbol shown in Figure 10.19, where v_C is the controlling voltage (v_G in Figure 10.18).

Page 632

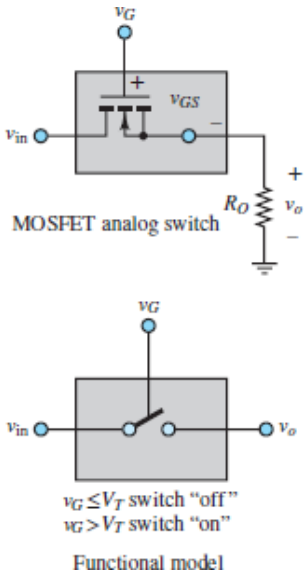


Figure 10.18 MOSFET analog switch

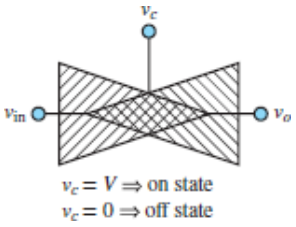
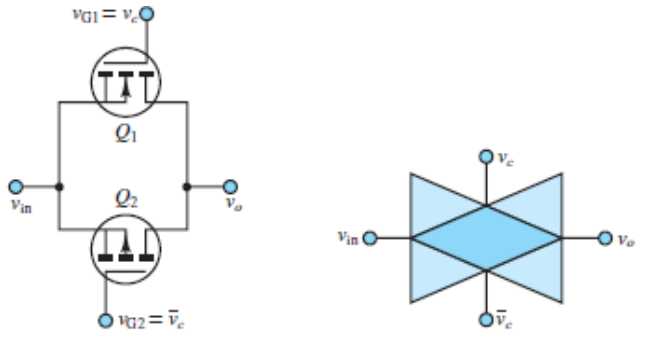


Figure 10.19 Symbol for a bilateral FET analog gate

FOCUS ON MEASUREMENTS

MOSFET Bidirectional Analog Gate

The variable-resistor feature of MOSFETs in the ohmic state finds application in the **analog transmission gate**. The circuit shown in Figure 10.20 depicts a circuit constructed using CMOS technology. The circuit operates on the basis of a control voltage v_C that can be either low (say, 0 V) or high ($v_C \gg V_t$), where V_t is the threshold voltage for the *n*-channel MOSFET and $-V_t$ is the threshold voltage for the *p*-channel MOSFET. The circuit operates in one of two modes. When the gate of Q_1 is connected to the high voltage and the gate of Q_2 is connected to the low voltage, the path between v_{in} and v_o has a relatively small resistance and the transmission gate conducts. When the gate of Q_1 is connected to the low voltage and the gate of Q_2 is connected to the high voltage, the transmission gate acts as a very large resistance and is an open-circuit for all practical purposes. A more precise analysis follows.



(a) CMOS transmission gate

(b) CMOS transmission gate circuit symbol

Figure 10.20 Analog transmission gate

Page 633

Let $v_C = V \gg V_t$ and $v_c = 0$. Assume that the input voltage v_{in} is in the range $0 \leq v_{in} \leq V$. To determine the state of the transmission gate, consider only the extreme cases $v_{in} = 0$ and $v_{in} = V$. When $v_{in} = 0$, $v_{GS1} = v_C - v_{in} = V - 0 = V \gg V_t$. Since V is above the threshold voltage, MOSFET Q_1 conducts (in the ohmic region). Further, $v_{GS2} = \bar{v}_C - v_{in} = 0 > -V_t$. Since the gate-source voltage is not more negative than the threshold voltage, Q_2 is in cutoff and does not conduct. Since one of the two possible paths between v_{in} and v_o is conducting, the transmission gate is on. Now consider the other extreme, where $v_{in} = V$. By reversing the previous argument, Q_1 is now off, since $v_{GS1} = 0 < V_t$. However, now Q_2 is in the ohmic state, because $v_{GS2} = \bar{v}_c - v_{in} = 0 - V \ll -V_t$. In this case, then, it is Q_2 that provides a conducting path between the input and the output of the transmission gate, and the transmission gate is also on. Thus, when $v_C = V$ and $v_c = 0$, the transmission gate conducts and provides a near-zero-resistance (typically tens of ohms) connection between the input and the output of the transmission gate, for values of the input ranging from 0 to V.

When the control voltages are reversed so that $v_C = 0$ and $\bar{v}_C = V \gg V_t$, it is straightforward to show that, regardless of the value of v_{in}, both Q_1 and Q_2 are always off; therefore, the transmission gate is essentially an open-circuit.

The analog transmission gate finds common application in *analog multiplexers* and *sample-and-hold* circuits.

EXAMPLE 10.8 NMOS Switch

Problem

Determine the operating points of the NMOS switch of Figure 10.21 when the input signal is equal to 0 and 2.5 V, respectively.

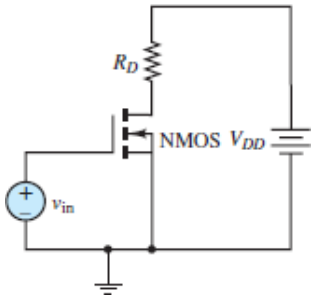


Figure 10.21

Solution

Known Quantities: Drain resistor; V_{DD}; input signal voltage.

Find: The Q point for each value of the input signal voltage.

Schematics, Diagrams, Circuits, and Given Data: $R_D = 125\ \Omega$; $V_{DD} = 10$ V; $v_{in} = 0$ V for $t < 0$; $v_{in} = 2.5$ V for $t \leq 0$.

Assumptions: NMOS drain characteristic curves (Figure 10.22).

Analysis: Apply KVL around the drain circuit to find its load line:

$$V_{DD} = R_D i_D + v_{DS} \qquad 10 = 125\, i_D + v_{DS}$$

If $i_D = 0$, then $v_{DS} = 10$ V. Likewise, if $v_{DS} = 0$, then $i_D = 10/125 = 80$ mA. These two results establish the load line for the drain circuit, as shown in Figure 10.22.

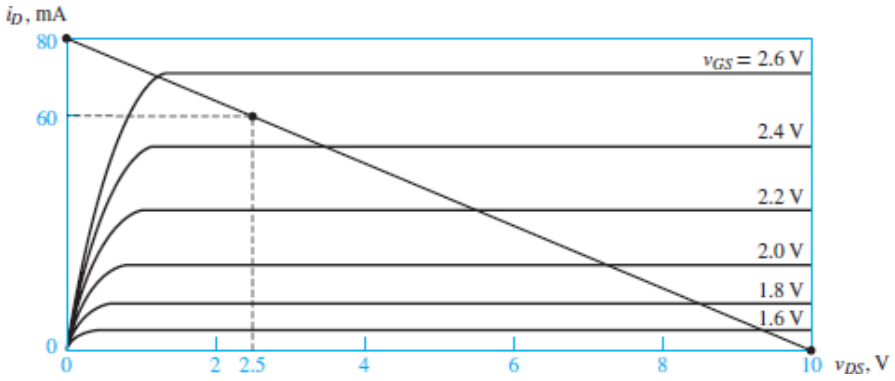


Figure 10.22 Drain curves for the NMOS of Figure 10.21

1. $t < 0$ s: When the input signal is zero, the gate voltage is zero and the NMOS is in the cutoff region. The Q point is

$$v_{GSQ} = 0 \text{ V} \qquad i_{DQ} = 0 \text{ mA} \qquad v_{DSQ} = 10 \text{ V}$$

2. $t \leq 0$ s: When the input signal is 2.5 V, the gate voltage is 2.5 V and the NMOS is in the saturation region as indicated by the intersection of the load line and the approximate location of the $v_{GS} = 2.5$ V curve. The Q point is

$$v_{GSQ} = 2.5 \text{ V} \qquad i_{DQ} = 60 \text{ mA} \qquad v_{DSQ} = 2.5 \text{ V}$$

This result satisfies KVL because $R_D i_D = 0.06 \times 125 = 7.5$ V

EXAMPLE 10.9 CMOS Gate

Problem

Determine the logic function implemented by the CMOS gate of Figure 10.23. Use the table below to summarize the behavior of the circuit.

v_1 (V)	v_2 (V)	State of M_1	State of M_2	State of M_3	State of M_4	v_o
0	0					
0	5					
5	0					
5	5					

Solution

Find: The logic value of v_{out} for each combination of v_1 and v_2 .

Schematics, Diagrams, Circuits, and Given Data: $V_t = 1.7$ V; $V_{DD} = 5$ V.

Assumptions: Treat the MOSFETs as open-circuits when off and as linear resistors when on.

Page 635

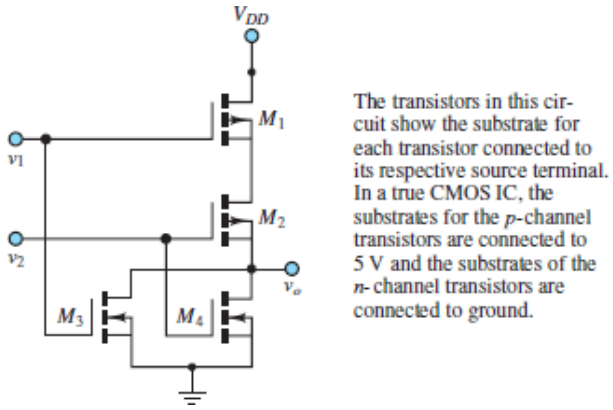


Figure 10.23

Analysis: Note that the state of an NMOS transistor for a high (5-V) gate input is the same as the state of a PMOS transistor for a low (0-V) gate input; both result in the formation of a channel with the transistors in the triode (ohmic) state. In these two cases, the transistors can be represented by simple linear resistors.

On the other hand, the state of an NMOS transistor for a low (0-V) gate input is the same as the state of a PMOS transistor for a high (5-V) gate input; both result in no channel formation with the transistors in the cutoff state. In these two cases, the transistors can be represented by open-circuits.

a. $v_1 = v_2 = 0$ **V**: With both input voltages equal to zero, M_3 and M_4 are in cutoff and are off since $v_{GS} < V_t$ for both transistors. On the other hand, both M_1 and M_2 form channels, are on, and act as simple linear resistors. However,

because both M_3 and M_4 act as open-circuits, there is no current through M_1 and M_2, which act as *pull-up* resistors; that is, with no current through M_1 and M_2, there is no voltage drop across either transistor and, thus, $v_o = V_{DD} =$ 5 V, which is a logic high. This situation is depicted in Figure 10.24(a).

b. $v_1 = 0$ **V**; $v_2 =$ **5 V**: With $v_1 = 0$, M_1 forms a channel, is on, and acts as a linear resistor. However, M_3 does not form a channel, is off, and acts as an open-circuit. With $v_2 = 5$ V, M_2 does not form a channel, is off, and acts as an open-circuit, whereas M_4 forms a channel, is on, and acts as a linear resistor. This situation is depicted in Figure 10.24(b). Notice that there can be no current through M_4 because M_2 prevents M_4 from seeing the 5-V source. The result is that $v_o = 0$, which is a logic low.

c. $v_1 = 5$ **V**; $v_2 =$ **0 V**: By symmetry with case b, when the values of v_1 and v_2 are inverted, the states of the four transistors are also inverted. As a result, M_1 and M_4 are off and act as open-circuits, whereas M_2 and M_3 are on and act as linear resistors, as depicted in Figure 10.24(c). Again, an open-circuit, this time M_1, prevents M_3 from seeing the 5-V source such that there is no current through M_3. The result is that once again $v_o = 0$, which is a logic low.

d. $v_1 = v_2 = 5$ **V**: Finally, with both input voltages equal to 5 V, M_1 and M_2 do not form channels, are off, and act as open-circuits. Although M_3 and M_4 both form channels, are on, and act as linear resistors, as depicted in Figure 10.24(d), they are unable to see the 5-V supply voltage and, thus, their currents are zero. Therefore, $v_o = 0$, which is a logic low. Notice that this situation is the inverse of that in case a.

Page 636

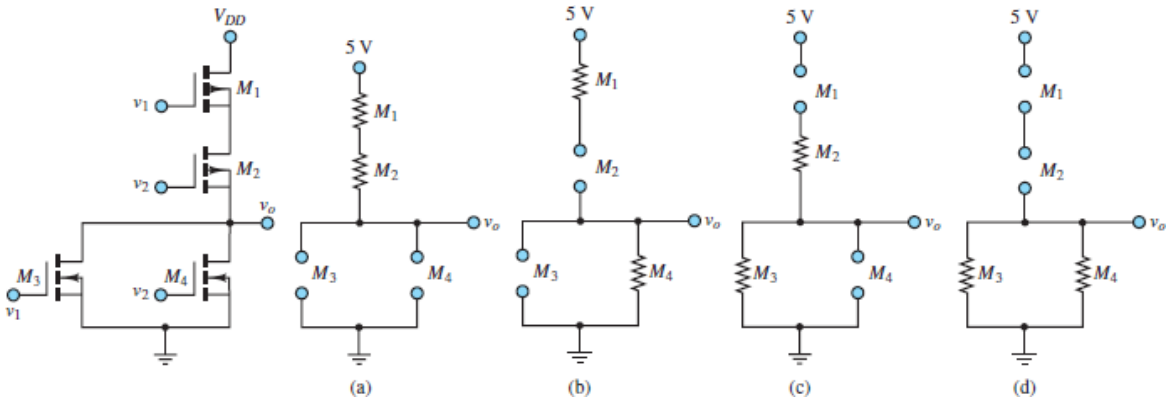


Figure 10.24 When $v_1 = v_2 = 0$, the gate-to-source voltage is low for all four transistors. The result is that the NMOS transistors M_3 and M_4 are off, while the PMOS transistors M_1 and M_2 are on.

These results are summarized in the table below.

v_1 (V)	v_2 (V)	M_1	M_2	M_3	M_4	v_o (V)
0	0	On	On	Off	Off	5
0	5	On	Off	Off	On	0
5	0	Off	On	On	Off	0
5	5	Off	Off	On	On	0

Columns v_1, v_2, and v_o represent a two-variable *truth table* when 0 V and 5 V are interpreted as FALSE and TRUE conditions, respectively. The results indicate that the output variable v_o is TRUE if and only if both input variables are FALSE. Otherwise, the output is FALSE. Such a truth table describes a two-input NOR gate.

CHECK YOUR UNDERSTANDING

What value of R_D would ensure a drain-to-source voltage v_{DS} of 5 V when v_{in} = 2.5 V in the circuit of Example 10.8?

Answer: 83.3 Ω

CHECK YOUR UNDERSTANDING

Analyze the CMOS gate of Figure 10.25, and find the output voltages for the following conditions: (a) $v_1 = 0$, $v_2 = 0$; (b) $v_1 = 5$ V, $v_2 = 0$; (c) $v_1 = 0$, $v_2 = 5$ V; (d) $v_1 = 5$ V, $v_2 = 5$ V. Show that the behavior is equivalent to a logical NAND gate.

Answer:

NAND gate

v_1 (V)	v_2 (V)	v_o (V)
0	0	5
0	5	5
5	0	5
5	5	0

Figure 10.25 CMOS NAND gate

CHECK YOUR UNDERSTANDING

Show that the CMOS bidirectional gate described in the Focus on Measurements box, "MOSFET Bidirectional Analog Gate," is off for all values of v_{in} between 0 and V whenever $v_C = 0$ and $v_C = V > V_t$.

Conclusion

This chapter has introduced field-effect transistors, focusing primarily on metal-oxide semiconductor enhancement-mode *n*-channel devices to explain the operation of FETs as amplifiers. A brief introduction to *p*-channel devices is used as the basis to introduce CMOS technology and to present analog and digital switches and logic gate applications of MOSFETs. Upon completing this chapter, you should have mastered the following learning objectives:

1. *Understand the classification of field-effect transistors.* FETs include three major families; the enhancement-mode family is the most commonly used and is the one explored in this chapter. Depletion-mode and junction FETs are only mentioned briefly.
2. *Learn the basic operation of enhancement-mode MOSFETs by understanding their i-υ curves and defining equations.* MOSFETs can be described by the *i-υ* drain characteristic curves and by a set of nonlinear equations linking the drain current to the gate-to-source and drain-to-source voltages. MOSFETs can operate in one of four regions: *cutoff*, in which the transistor does not conduct current; *triode*, in which the transistor can act as a voltage-controlled resistor under certain conditions; *saturation*, in which the transistor acts as a voltage-controlled current source and can be used as an amplifier; and *breakdown* when the limits of operation are exceeded.
3. *Learn how enhancement-mode MOSFET circuits are biased.* MOSFET circuits can be biased to operate around a certain operating point, known as the *Q* point, when appropriate supply voltages and resistors are selected.
4. *Understand the concept and operation of FET large-signal amplifiers.* Once a MOSFET circuit is properly biased in the saturation region, it can serve as an amplifier by virtue of its voltage-controlled current source property: small changes in the gate-to-source voltages are translated to proportional changes in drain current.
5. *Understand the concept and operation of FET switches.* MOSFETs can serve as analog and digital switches: by controlling the gate voltage, a MOSFET can be turned on and off (digital switch), or its resistance can be modulated (analog switch).
6. *Analyze FET switches and digital gates.* These devices find application in CMOS circuits as digital logic gates and analog transmission gates.

HOMEWORK PROBLEMS

Section 10.2: Enhancement-Mode MOSFETs

10.1 The transistors shown in Figure P10.1 have $|V_t| = 3$ V. Determine the operating region.

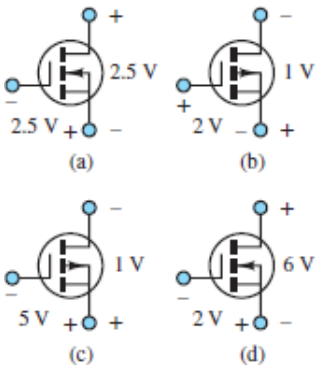


Figure P10.1

10.2 The three terminals of an *n*-channel enhancement-mode MOSFET are at potentials of 4, 5, and 10 V with respect to ground. Draw the circuit symbol, with the appropriate voltages at each terminal, if the device is operating

a. In the ohmic region.

b. In the saturation region.

10.3 An enhancement-type NMOS transistor with V_t = 2 V has its source grounded and a 3-VDC source connected to the gate. Determine the operating state if

a. $v_D = 0.5$ V

b. $v_D = 1$ V

c. $v_D = 5$ V

10.4 In the circuit shown in Figure P10.4, the PMOS transistor has $|V_t| = 2$ V and $K = 10$ mA/V^2. Find R and v_S for $i_S = 0.4$ mA.

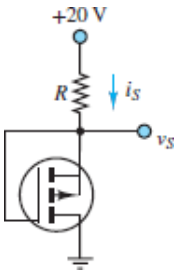


Figure P10.4

10.5 An enhancement-type NMOS transistor has V_t = 2.5 V and i_D = 0.8 mA when $v_{GS} = v_{DS} = 4$ V. Find the value of i_D for $v_{GS} = 5$ V.

10.6 The NMOS transistor shown in Figure P10.6 has V_t = 1.5 V and K = 0.4 mA/V^2. If v_G is a pulse with 0 to 5 V, find the voltage levels of the pulse signal at the drain.

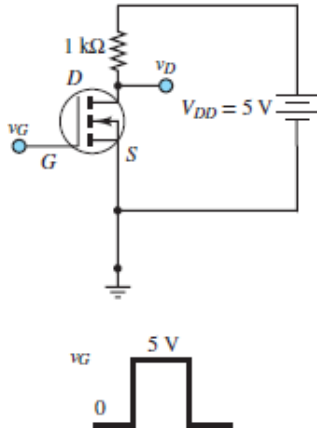


Figure P10.6

10.7 In the circuit shown in Figure P10.7, a drain voltage of 0.1 V is established. Find the current i_D for $V_t = 1$ V and $K = 0.5$ mA/V^2.

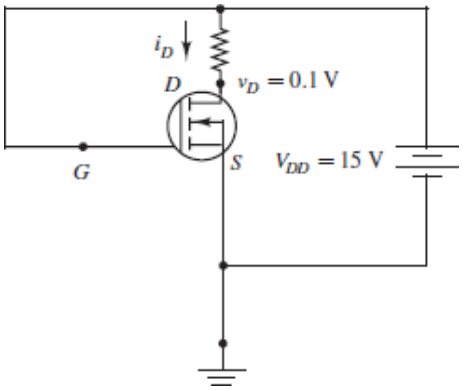


Figure P10.7

Section 10.3: Biasing MOSFET Circuits

10.8 An *n*-channel enhancement-mode MOSFET, shown in Figure P10.8, is operated in the ohmic Page 639region. Size the resistors so that the quiescent drain current $I_{DQ} = 4$ mA. Let $V_{DD} = 15$ V, $K = 0.3$ mA/V^2, and $V_t = 3.3$ V.

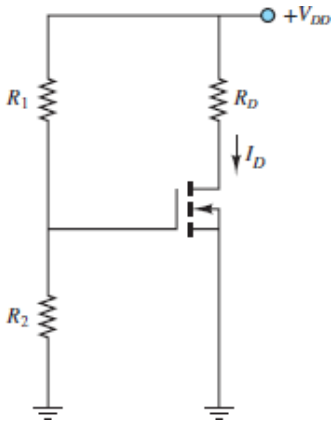


Figure P10.8

10.9 Compute the power dissipated by the circuit in Figure P10.9. Let $V_{DD} = V_{SS} = 15$ V, $R_1 = R_2 = 90$ k Ω, $R_D = 0.1$ k Ω, $V_t = 3.5$ V, $K = 0.816$ mA/V^2.

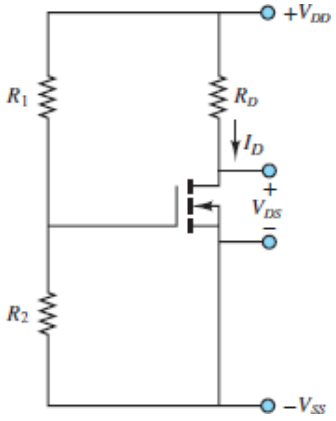


Figure P10.9

10.10 Find the operating region of the enhancement-type NMOS transistor shown in Figure P10.8. Let $V_{DD} = 20$ V, $K = 0.2$ ma/V^2, $V_t = 4$ V, $R_1 = 4$ MΩ, $R_2 = 3$ MΩ, and $R_D = 3$ kΩ.

10.11 Find the operating region of the enhancement-type NMOS transistor shown in Figure P10.11. Let $V_{DD} = 18$ V, $K = 0.3$ ma/V^2, $V_t = 3$ V, $R_1 = 5.5$ MΩ, $R_2 = 4.5$ MΩ, $R_D = 2$ k Ω, and $R_S = 1$ k Ω.

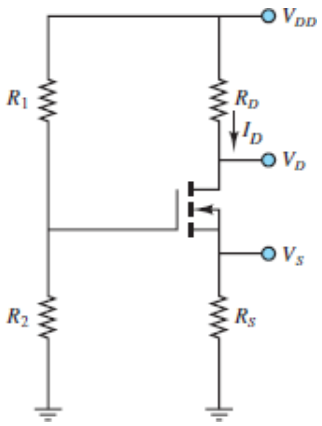


Figure P10.11

10.12 In the circuit shown in Figure P10.12, the MOSFET operates in the saturation region, for $I_S = 0.5$ mA and $V_S = 3$ V. This enhancement-type PMOS has $V_t = -1$ V, and $K = 0.5$ mA/V^2. Find:

a. R_S and the ratio R_1/R_2.

b. The largest allowable value of R_S for the MOSFET to remain in the saturation region.

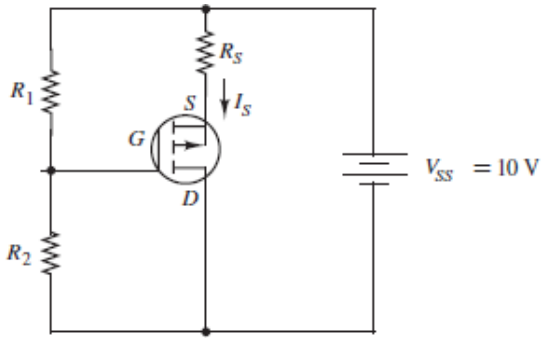


Figure P10.12

10.13 The *i-v* characteristic of an *n*-channel enhancement MOSFET is shown in Figure P10.13(a); a standard bias circuit based on the *n*-channel MOSFET is shown in Figure P10.13(b). Determine the quiescent current I_{DQ} and drain-to-source voltage V_{DS} when $V_{DD} = 10$ V and $R_D = 5$ Ω. In what region is the transistor operating?

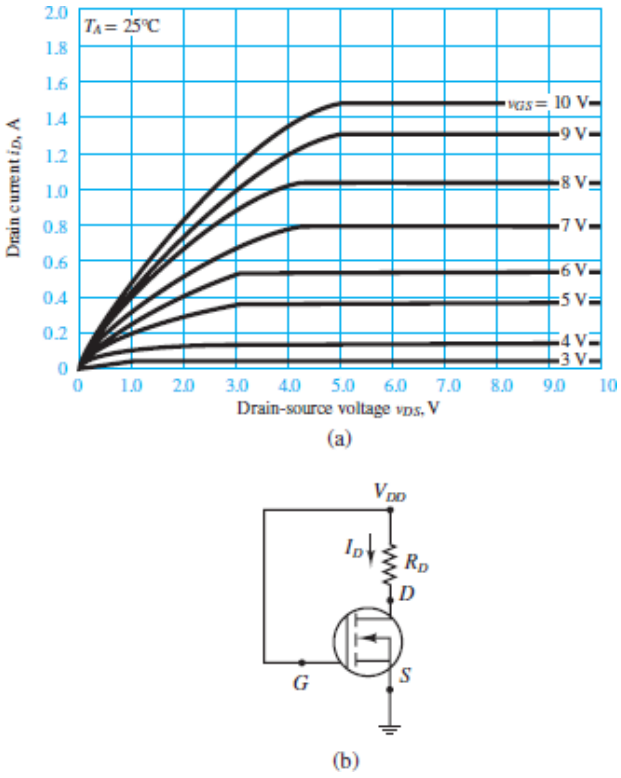


Figure P10.13

10.14 Given the enhancement-type NMOS transistor and drain characteristic shown in Figure P10.14, compute R_S and V_{DD}. Let R_1 = 200 kΩ and R_2 = 100 k Ω.

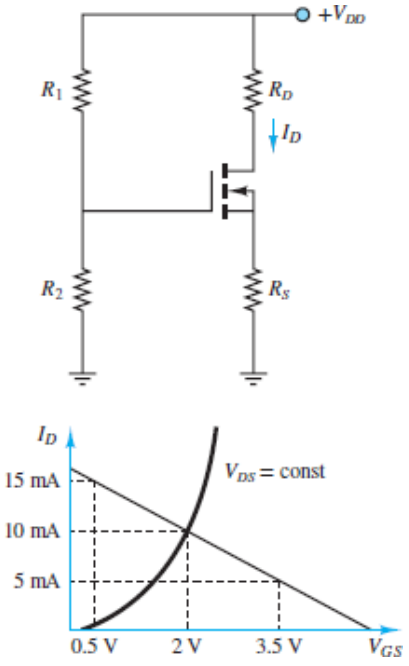


Figure P10.14

10.15 Given the enhancement-type NMOS transistor shown in Figure P10.8, compute R_1, R_2 and R_D. Let $I_D = 2$ mA, $V_t = 4$ V, $V_{DS} = 8$ V, $V_{DD} = 16$ V, $K = 0.375$ mA/V^2, and total dissipated power $P_T = 35$ mW.

10.16 Given the enhancement-type NMOS transistor shown in Figure P10.11, compute R_1, R_2, R_S and R_D. Let $I_D = 4$ mA, $V_D = 9$ V, $V_{DS} = 4.5$ V, $V_{DD} = 18$ V, $V_t = 4$ V, $K = 0.625$ mA/V^2, and maximum total dissipated power $P_{T,\,\max} = 75$ mW.

Section 10.4: MOSFET Large-Signal Amplifiers

10.17 The power MOSFET circuit of Figure P10.17 is configured as a voltage-controlled current source (VCCS). Let $K = 1.5$ A/V^2 and $V_t = 3$ V.

a. If $V_G = 5$ V, find the range of R for which the VCCS will operate.

b. If $R = 1\ \Omega$, determine the range of V_G for which the VCCS will operate.

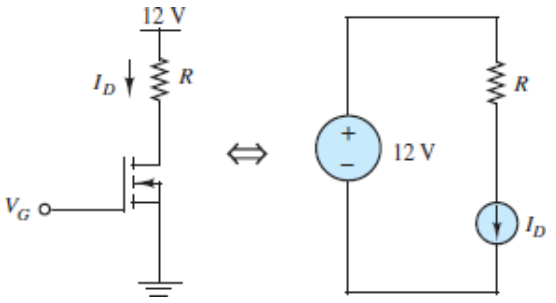


Figure P10.17

10.18 The circuit of Figure P10.18 is called a *source follower* and acts as a voltage-controlled current source (VCCS).

a. Determine I_S if $V_G = 10$ V, $R = 2\ \Omega$, $K = 0.5$ A/V^2 and $V_t = 4$ V.

b. If the power rating of the MOSFET is 50 W, how small can R be?

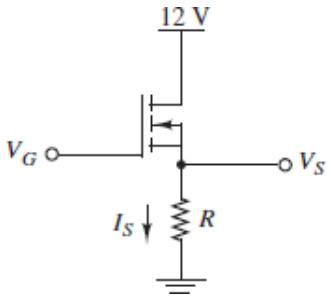


Figure P10.18

Page 641

10.19 The circuit of Figure P10.19 is a class A amplifier.

a. Determine the output current for the given biased audio tone input v_G = 10 + 0.1 cos(500t) V. Let K = 2 mA/V^2 and V_t = 3 V.

b. Determine the output voltage v_o.

c. Determine the voltage gain of the cos(500 t) signal.

d. Determine the DC power consumption of the resistor and the MOSFET.

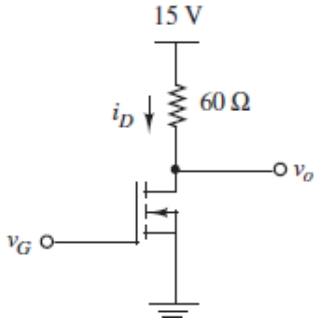


Figure P10.19

10.20 The circuit of Figure P10.20 is a source-follower amplifier. Let K = 30 mA/V^2, V_t = 4 V, and v_G = 9 + 0.1 cos(500 t) V.

a. Determine the source current i_S.

b. Determine the output voltage v_o.

c. Determine the voltage gain for the cos(500 t) signal.

d. Determine the DC power consumption of the MOSFET and the 4-Ω resistor.

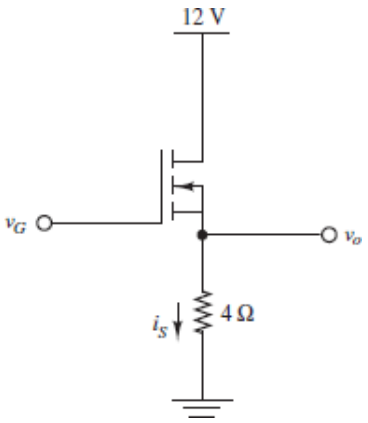


Figure P10.20

10.21 Sometimes it is necessary to discharge batteries before recharging. To do this, an electronic load can be used. A high-power electronic load is shown in Figure P10.21, for the battery discharge application. With $K = 4$ A/V^2, $V_t = 3$ V, and $V_G = 8$ V, determine the discharging current I_D and the required MOSFET power rating.

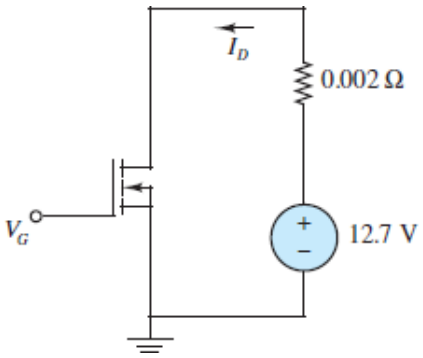


Figure P10.21

10.22 A precision voltage source can be created by driving the drain of a MOSFET. Figure P10.22 shows a circuit that will accomplish this function. With $I_{\text{Ref}} = 0.01$ A, determine the output V_G. Let $K = 0.006$ A/V^2 and $V_t = 1.5$ V.

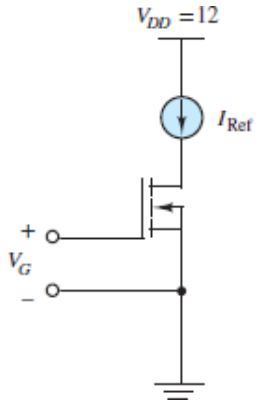


Figure P10.22

10.23 To allow more current in a MOSFET amplifier, several MOSFETs can be connected in parallel. Determine the currents I_D and I_S in the circuits of Figure P10.23. Let $K = 0.2$ A/V^2, $V_t = 3$ V, and $V_G = V_{DD}$.

Page 642

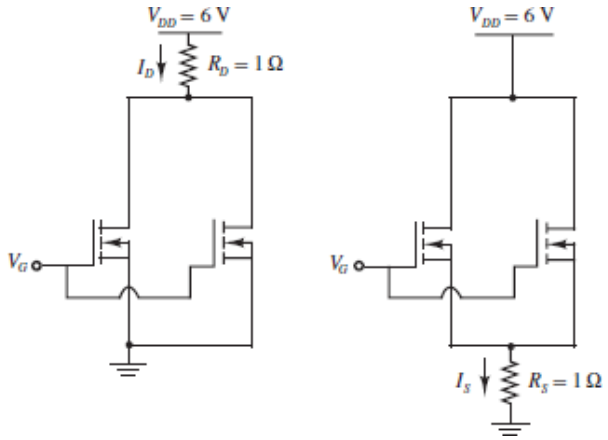


Figure P10.23

10.24 A push-pull amplifier can be constructed from matched n and p-channel MOSFETs, as shown in Figure P10.24. Let $K_n = K_p = 0.5$ A/V^2, $V_{in} = +3$ V, $V_{ip} = -3$ V, and $v_{\text{in}} = 0.8 \cos(1{,}000\, t)$ V. Determine v_o and i_o.

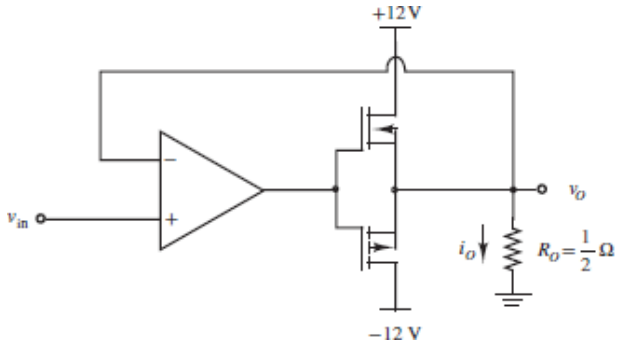


Figure P10.24

10.25 Show that the NMOS shown in Figure P10.25 cannot be in triode mode. Determine its *i*-*v* characteristic to show that it acts as a VCCS.

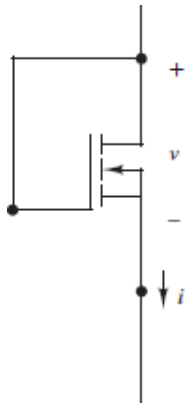


Figure P10.25

10.26 Determine v_o and i_o for the two-stage amplifier shown in the circuit of Figure P10.26, with identical MOSFETs having $K = 1$ A/V^2 and $V_t = 3$ V, for

a. $v_G = 4$ V.

b. $v_G = 5$ V.

c. $v_G = 4 + 0.1\cos(750\,t)$.

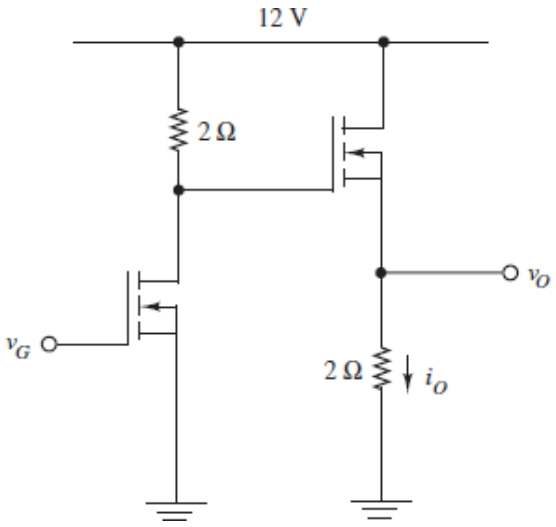


Figure P10.26

Section 10.5: CMOS Technology and MOSFET Switches

10.27 For the CMOS NOR gate of Figure 10.23 identify the state of each transistor for $v_1 = v_2 = 5$ V. Assume $V_{DD} = 5$ V.

10.28 Repeat Problem 10.27 for $v_1 = 5$ V and $v_2 = 0$ V.

10.29 Draw the schematic diagram of a two-input CMOS OR gate.

10.30 Draw the schematic diagram of a two-input CMOS AND gate.

10.31 Draw the schematic diagram of a two-input CMOS NOR gate.

10.32 Draw the schematic diagram of a two-input CMOS NAND gate.

10.33 Draw the schematic diagram of a three-input CMOS OR gate.

10.34 Draw the schematic diagram of a three-input CMOS AND gate.

10.35 Draw the schematic diagram of a three-input CMOS gate that realizes the logic function $\overline{A(B+C)}$.

Page 643

10.36 Show that the circuit of Figure P10.36 functions as a logic inverter.

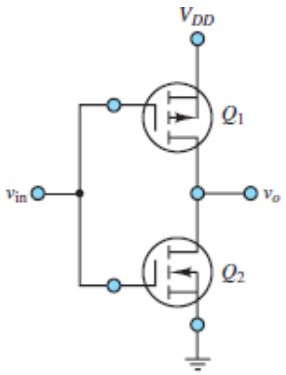


Figure P10.36

10.37 Show that the circuit of Figure P10.37 functions as a NOR gate.

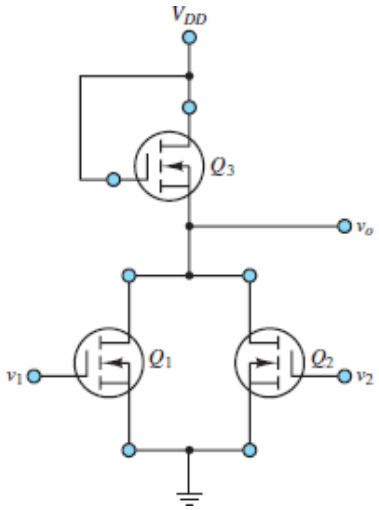


Figure P10.37

10.38 Show that the circuit of Figure P10.38 functions as a NAND gate.

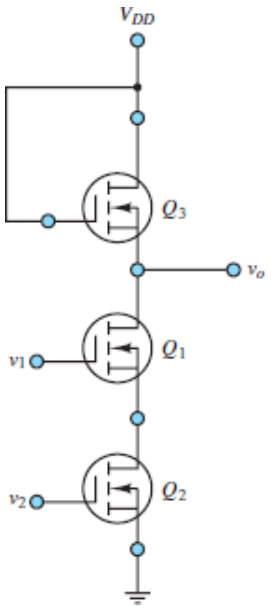


Figure P10.38

10.39 Determine the logic function implemented by the CMOS gate of Figure P10.39. Use a table to summarize the behavior of the circuit.

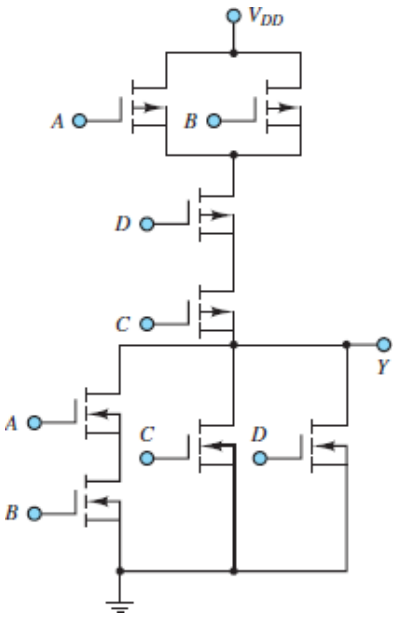


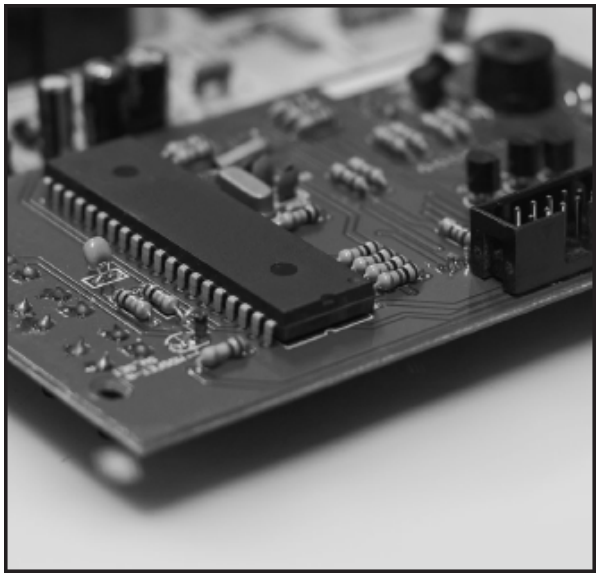
Figure P10.39

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]The bulk is also known as the **substrate, body**, or **base**.

[2]In the past, a metal oxide was used, which explains the terminology *metal-oxide semiconductor* MOS.

PART IV
DIGITAL ELECTRONICS



yuyanga/Getty Images

C H A P T E R
11

DIGITAL LOGIC CIRCUITS

Digital computers have played a prominent role in engineering and science for over half a century, performing a number of essential functions such as numerical computations and data acquisition. The elements of all digital computers are *combinational* and *sequential* logic circuits, built up from basic *logic gates*. The inputs, operations, and outputs of these circuits are described in terms of the binary number system and boolean algebra. Several practical examples are presented to demonstrate that even simple combinations of logic gates can perform useful functions in engineering practice. A number of logic modules are introduced that are described in terms of simple logic gates and yet provide more advanced functions, such as read-only memory, multiplexing, and decoding. Throughout the chapter, simple examples are given to demonstrate the usefulness of digital logic circuits in various engineering applications.

Learning Objectives

Students will learn to...

1. Apply concepts of analog and digital signals and of quantization. *Section 11.1*.
2. Convert between decimal and binary number systems and use the hexadecimal system and BCD and Gray codes. *Section 11.2*.

3. Write truth tables, and realize logic functions from truth tables by using logic gates. *Section 11.3*.
4. Systematically design logic functions using Karnaugh maps. *Section 11.4*.
5. Apply various combinational logic modules, including multiplexers, memory and decoder elements, and programmable logic arrays. *Section 11.5*.

11.1 ANALOG AND DIGITAL SIGNALS

One of the fundamental distinctions in the study of signals derived from physical measurements is that between analog and digital signals. An **analog signal** is one that varies in analogy with a physical quantity (e.g., temperature, force, or acceleration). For example, most electronic sensors produce a voltage proportional to some other measured quantity, such as a variable pressure or vibration. Figure 11.1 depicts an analog function of time $f(t)$, which can take any value in the given range.

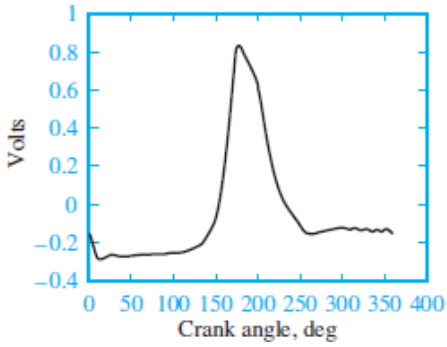


Figure 11.1 Voltage analog of internal combustion engine in-cylinder pressure

A **digital signal**, on the other hand, can take only a *finite number of values*. This is an extremely important distinction. An example of a digital signal is one that allows display of a temperature measurement on a digital readout. Assume that the digital readout is three digits long and can display numbers from 0 to 100, that the temperature sensor is correctly calibrated to measure temperatures from 0 to 100°C, and that the output of the sensor ranges from 0 to 5 V, where 0 V corresponds to 0°C and 5 V to 100°C. Then, the calibration constant of the sensor is

$$k_T = \frac{100°\text{C} - 0°\text{C}}{5\ \text{V} - 0\ \text{V}} = 20°\text{C/V}$$

Clearly, the output of the sensor is an analog signal; however, the display can show only a finite number of readouts (101, to be precise). Because the display can only take a value out of a discrete set of states—the integers from 0 to 100—it is called a digital display, indicating that the variable displayed is expressed in digital form.

It is important to realize that each temperature on the display corresponds to a *range of voltages:* each digit on the display represents one-hundredth of the 5-V range of the sensor, or 0.05 V = 50 mV. Thus, the display will read 0 if the sensor voltage is between 0 and 49 mV, 1 if it is between 50 and 99 mV, and so on. Figure 11.2 depicts the staircase function relationship between the analog voltage and the digital readout. This **quantization** of the sensor output voltage is in effect an approximation. To know the temperature with greater precision, a greater number of display digits could be employed.

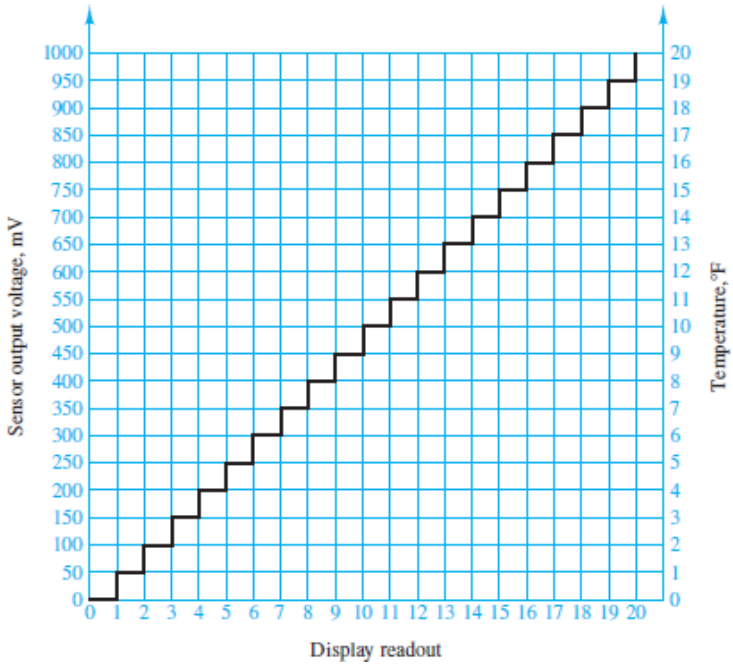


Figure 11.2 Digital representation of an analog signal

The most common digital signals are binary signals. A **binary signal** is a signal that can take only one of two discrete values and is therefore characterized by transitions between two states. Figure 11.3 displays a typical binary signal. In binary arithmetic (which we discuss in Section 11.2), the two discrete values f_1 and f_0 are represented, respectively, by the numbers 1 and 0. In binary voltage waveforms, these values are represented by two voltage levels. For example, in the TTL

convention (see Chapter 9), these values are (nominally) 5 and 0 V, respectively; in CMOS circuits, these values can vary substantially. Other conventions are also used, including reversing the assignment, for example, by letting a 0-V level represent a logic 1 and a 5-V level represent a logic 0. Note that in a binary Page 648waveform, knowledge of the transition between one state and another (e.g., from f_0 to f_1 at $t = t_2$) is equivalent to knowledge of the state. Thus, digital logic circuits can operate by detecting transitions between voltage levels. The transitions are often called **edges** and can be positive (f_0 to f_1) or negative (f_1 to f_0). They are also referred to as*leading* and *trailing* edges, respectively. Virtually all the signals handled by a computer are binary.

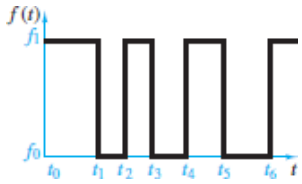


Figure 11.3 A binary signal

11.2 THE BINARY NUMBER SYSTEM

The binary number system is a natural choice for representing the behavior of circuits that operate in one of two states (on or off, 1 or 0, or the like). Diode and transistor gates and switches fall into this category. Table 11.1 shows the correspondence between decimal and binary number systems for integer decimal numbers up to 16.

Table 11.1 Conversion from decimal to binary

Decimal number n_{10}	Binary number n_2
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111
16	10000

Binary numbers are based on powers of 2, whereas the decimal system is based on powers of 10. For example, the number 372 in the decimal system can be expressed as

$$372 = (3 \times 10^2) + (7 \times 10^1) + (2 \times 10^0)$$

while the binary number 10110 corresponds to the following combination of powers of 2:

$$10110 = (1 \times 2^4) + (0 \times 2^3) + (1 \times 2^2) + (1 \times 2^1) + (0 \times 2^0)$$

A subscript is often used to indicate the base of a number and thus clarify its value, as shown below. The absence of a subscript often implies base 10, although at times a sequence of 1's and 0's is understood from its context to be base 2.

$$10110_2 = 16 + 0 + 4 + 2 + 0 = 22_{10}$$

Note that a fractional number can also be similarly represented. For example, the number 3.25 in the decimal system may be represented as

$$3.25_{10} = 3 \times 10^0 + 2 \times 10^{-1} + 5 \times 10^{-2}$$

while in the binary system the number 10.011 corresponds to

$$10.011_2 = 1 \times 2^1 + 0 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}$$

$$= 2 + 0 + 0 + \frac{1}{4} + \frac{1}{8} = 2.375_{10}$$

Table 11.1 shows that it takes four binary digits, also called **bits**, to represent the decimal numbers 0 to 15. Usually, the rightmost bit is called the **least significant bit**, or **LSB**, and the leftmost bit is called the **most significant bit**, or **MSB**. Since binary numbers clearly require a larger number of digits than decimal numbers to represent a value, the digits are usually grouped into sets of 4, 8, or 16. Four bits are a **nibble** and eight bits are a **byte**. A **word** is the basic unit of data in a digital system. A word may be two or more bytes depending upon the particular digital architecture.

Addition and Subtraction

The binary operations of addition and subtraction are based on the simple rules shown in Tables 11.2 and 11.3. Note that, just as is done in the decimal system, a Page 649carry is generated whenever the sum of two digits exceeds the largest single-digit number in the given number system, which is 1 in the binary system. The carry is treated exactly as in the decimal system. A few examples of binary addition and subtraction are shown in Figures 11.4 and 11.5, with their decimal counterparts.

Table 11.2 Rules for addition

0 + 0 = 0
0 + 1 = 1
1 + 0 = 1
1 + 1 = 0 (with a carry of 1)

Table 11.3 Rules for subtraction

0 − 0 = 0
1 − 0 = 1
1 − 1 = 0
0 − 1 = 1 (with a borrow of 1)

Decimal	Binary	Decimal	Binary	Decimal	Binary
5	101	15	1111	3.25	11.01
+6	+110	+20	+10100	+5.75	+101.11
11	1011	35	100011	9.00	1001.00

(Note that in this example, $3.25 = 3\frac{1}{4}$ and $5.75 = 5\frac{3}{4}$.)

Figure 11.4 Examples of binary addition

Decimal	Binary	Decimal	Binary	Decimal	Binary
9	1001	16	10000	6.25	110.01
−5	−101	−3	−11	−4.50	−100.10
4	0100	13	01101	1.75	001.11

Figure 11.5 Examples of binary subtraction

Multiplication and Division

Whereas in the decimal system the multiplication table consists of $10^2 = 100$ entries, in the binary system we only have $2^2 = 4$ entries (see Table 11.4).

Table 11.4 Rules for multiplication

$0 \times 0 = 0$
$0 \times 1 = 0$
$1 \times 0 = 0$
$1 \times 1 = 1$

Division in the binary system is also based on rules analogous to those of the decimal system, with the two basic laws given in Table 11.5. Once again, there are only two cases, and just as in the decimal system, division by zero is not contemplated.

Table 11.5 Rules for division

$0 \div 1 = 0$
$1 \div 1 = 1$

Conversion From Decimal to Binary

The conversion of a decimal number to its binary equivalent is performed by successive division of the decimal number by 2, checking for the remainder each time. Figure 11.6 illustrates this idea with an example. The result obtained in Figure 11.6 may be easily verified by performing the opposite conversion, from binary to decimal:

$$110001 = 2^5 + 2^4 + 2^0 = 32 + 16 + 1 = 49$$

Remainder
$49 \div 2 = 24 + 1$
$24 \div 2 = 12 + 0$
$12 \div 2 = 6 + 0$
$6 \div 2 = 3 + 0$
$3 \div 2 = 1 + 1$
$1 \div 2 = 0 + 1$
$49_{10} = 110001_2$

Figure 11.6 Example of conversion from decimal to binary

The same technique can be used for converting decimal fractional numbers to their binary form, provided that the whole number is separated from the fractional part and each is converted to binary form (separately), with the results added at the end. Figure 11.7 outlines this procedure by converting the number 37.53 to binary form. The procedure is outlined in two steps. First, the integer part is converted; then, to convert the fractional part, one simple technique consists of multiplying the decimal fraction by 2 in successive stages. If the result exceeds 1, a 1 is needed to the right of the binary fraction being formed (100101 . . . , in our example). Otherwise, a 0 is added. This procedure is continued until no fractional terms are left. In this case, the decimal part is 0.53_{10}, and Figure 11.7 illustrates the succession of calculations. Stopping the procedure after 11 digits results in the following approximation:

$$37.53_{10} = 100101.10000111101$$

Greater precision could be attained by continuing to add binary digits, at the expense of added complexity. Note that an infinite number of binary digits may be required to represent a decimal number *exactly*.

Remainder
$37 \div 2 = 18 + 1$
$18 \div 2 = 9 + 0$
$9 \div 2 = 4 + 1$
$4 \div 2 = 2 + 0$
$2 \div 2 = 1 + 0$
$1 \div 2 = 0 + 1$
$37_{10} = 100101_2$

$2 \times 0.53 = 1.06 \rightarrow 1$
$2 \times 0.06 = 0.12 \rightarrow 0$
$2 \times 0.12 = 0.24 \rightarrow 0$
$2 \times 0.24 = 0.48 \rightarrow 0$
$2 \times 0.48 = 0.96 \rightarrow 0$
$2 \times 0.96 = 1.92 \rightarrow 1$
$2 \times 0.92 = 1.84 \rightarrow 1$
$2 \times 0.84 = 1.68 \rightarrow 1$
$2 \times 0.68 = 1.36 \rightarrow 1$
$2 \times 0.36 = 0.72 \rightarrow 0$
$2 \times 0.72 = 1.44 \rightarrow 1$
$0.53_{10} = 0.10000111101_2$

Figure 11.7 Conversion of 37.53_{10} from decimal to binary

Complements and Negative Numbers

Complements are used to simplify the operation of subtraction in digital computers. In practice, the operation $X - Y$ is replaced with the operation $X + (-Y)$. This

procedure results in considerable simplification since the computer hardware need include only adding circuitry. Two types of complements are used with binary numbers: the **ones complement** and the **twos complement**.

The ones complement of an n-bit binary number is obtained by subtracting the number itself from $2^n - 1$. Two examples are as follows:

$$\begin{aligned} a &= 0101 \\ \text{Ones complement of } a &= (2^4 - 1) - a \\ &= (1111) - (0101) \\ &= 1010 \end{aligned}$$

$$\begin{aligned} b &= 101101 \\ \text{Ones complement of } b &= (2^6 - 1) - b \\ &= (111111) - (101101) \\ &= 010010 \end{aligned}$$

The twos complement of an n-bit binary number is obtained by subtracting the number itself from 2^n. Twos complements of the same numbers a and b used in the preceding illustration are computed as follows:

$$\begin{aligned} a &= 0101 \\ \text{Twos complement of } a &= 2^4 - a \\ &= (10000) - (0101) \\ &= 1011 \end{aligned}$$

$$\begin{aligned} b &= 101101 \\ \text{Twos complement of } b &= 2^6 - b \\ &= (1000000) - (101101) \\ &= 010011 \end{aligned}$$

Page 651

A simple rule that may be used to obtain the twos complement directly from a binary number is the following: Starting at the least significant (rightmost) bit, copy each bit *until the first 1 has been copied*, and then replace each successive 1 by a 0 and each 0 by a 1. You may wish to try this rule on the two previous examples to verify that it is much easier to use than subtraction from 2^n.

Different conventions exist in the binary system to represent whether a number is negative or positive. One convention, called the **sign-magnitude convention**, designates a *sign bit*, usually positioned at the beginning of the number, for which a value of 1 represents a minus sign and a value of 0 represents a plus sign. Thus, an 8-bit binary number would consist of 1 sign bit followed by 7 *magnitude bits*, as shown in Figure 11.8(a). In a digital system that uses 8-bit signed integer words, we could represent integer numbers (decimal) in the range

$$-2^7 \leq N \leq +(2^7 - 1)$$

or

$$-128 \leq N \leq +127$$

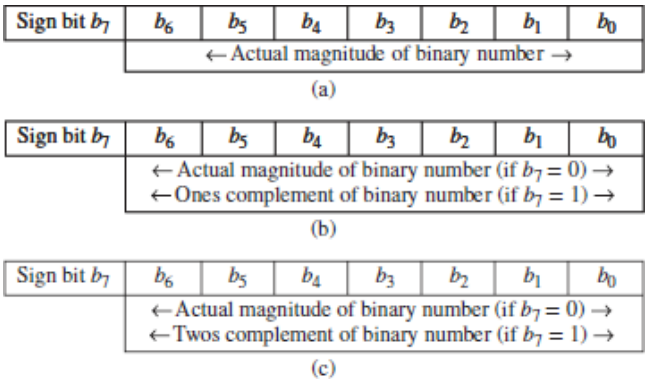


Figure 11.8 (a) An 8-bit sign-magnitude binary number; (b) an 8-bit 1s complement binary number; (c) an 8-bit 2s complement binary number

A second convention uses the ones complement notation. In this convention, a sign bit is also used to indicate whether the number is positive (sign bit = 0) or negative (sign bit = 1). However, the binary number is represented by its true magnitude if the number is positive and by its *ones complement* if the number is negative. Figure 11.8(b) illustrates the convention. For example, the number 91_{10} would be represented by the 7-bit binary number 1011011_2 with a leading 0 (the sign bit): $\mathbf{0}1011011_2$. On the other hand, the number -91_{10} would be represented by the 7-bit ones complement binary number 0100100_2 with a leading 1 (the sign bit): $\mathbf{1}0100100_2$.

Most digital computers use the twos complement convention in performing integer arithmetic operations. The twos complement convention represents positive numbers by a sign bit of 0, followed by the binary magnitude; negative numbers are represented by a sign bit of 1, *followed by the twos complement of the binary magnitude*, as shown in Figure 11.8(c). The advantage of the twos complement convention is that the algebraic sum of twos complement binary numbers is carried out very simply by adding the two numbers *including the sign bit*.

The Hexadecimal System

It should be apparent by now that representing numbers in base-2 and base-10 systems is simply a matter of convenience, given a specific application. Another frequently used base is 16, which results in the **hexadecimal system**. In the hexadecimal (or hex) code, the bits in a binary number are gathered into groups of 4. Since there are 16 possible distinct bit combinations for a 4-bit number, 16 distinct symbols are needed to fully represent a hex digit. Thus, the symbols 0–9 and A–F are used, as shown in Table 11.6. In hex code, an 8-bit word corresponds to just two digits; for example,

$$1010\ 0111_2 = A7_{16}$$
$$0010\ 1001_2 = 29_{16}$$

Table 11.6 Hexadecimal code

code	
0	0000
1	0001
2	0010
3	0111
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
A	1010
B	1011
C	1100
D	1101
E	1110
F	1111

The **ASCII**[1] character code represents all alphanumeric characters, and others, commonly used in printed documents as hexadecimal values. This code is used, for example, to define the visual output associated with **char** type variables found in all computer programming languages. The 128 members of the standard ASCII character set are listed in Appendix D along with their hexadecimal equivalents.

Binary Codes

Variations on the standard binary code are used in certain applications for practical reasons. Two of the most commonly used variations are the **Gray code** and **binary-**

coded decimal, or **BCD**, **representation**. The most basic BCD representation consists of the first 10 entries of the standard 4-bit binary code, as shown in Table 11.7. There are also other BCD codes, all reflecting the same principle: Each decimal digit is represented by a fixed-length binary word. Although this method is attractive because of its direct correspondence with the decimal system, it is not efficient. Consider, for example, the decimal number 68. Its binary representation by direct conversion is the 7-bit number 1000100. However, the corresponding BCD representation would require 8 bits:

$$68_{10} = 01101000_{BCD}$$

Table 11.7 BCD code

0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001

Gray code is simply a reshuffled binary code with the property that any two consecutive numbers differ by only 1 bit. Table 11.8 illustrates the 3-bit Gray code. Page 653The Gray code is useful in encoding applications because a single bit reading error results in an off-by-one counting error. Thus, the impact of bit reading errors is more likely to be marginal than when using other encoding schemes.

Table 11.8 Three-bit Gray code

Binary	Gray
000	000
001	001
010	011
011	010
100	110
101	111
110	101
111	100

FOCUS ON MEASUREMENTS

Digital Position Encoders

Position encoders are devices that output a digital signal proportional to their (linear or angular) position. These devices are very useful in measuring instantaneous position in *motion control* applications. Motion control is a technique used when it is necessary to accurately control the motion of a moving object; examples are found in robotics, machine tools, and servomechanisms. For example, in positioning the arm of a robot to pick up an object, it is very important to know its exact position at all times. Since one is usually interested in both rotational and translational motion, two types of encoders are discussed in this example: *linear* and *angular* position encoders.

An optical position encoder consists of an *encoder pad*, which is either a strip (for translational motion) or a disk (for rotational motion) with alternating black and white areas. These areas are arranged to reproduce some binary code, as shown in Figure 11.9, where both the conventional binary and Gray codes are depicted for a 4-bit linear encoder pad. A fixed array of photodiodes (see Chapter 8) senses the reflected light from each of the cells across a row of the encoder path; depending on the amount of light reflected, each photodiode circuit will output a voltage corresponding to a binary 1 or 0. Thus, a different 4-bit word is generated for each row of the encoder.

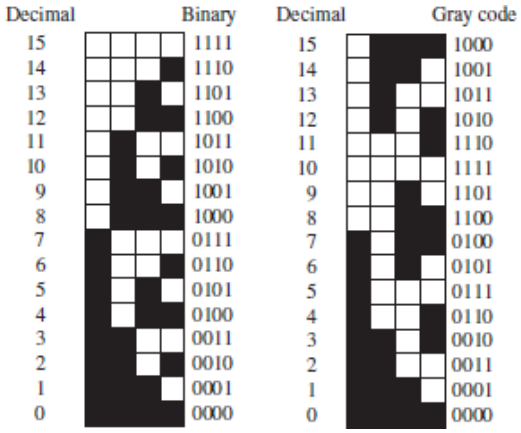


Figure 11.9 Binary and Gray code patterns for linear position encoders

Suppose the encoder pad is 100 mm in length. Then its resolution can be computed as follows. The pad will be divided into $2^4 = 16$ segments, and each segment corresponds to an increment of 100/16 mm = 6.25 mm. If greater resolution were necessary, more bits could be employed: an 8-bit pad of the same length would attain a resolution of 100/256 mm = 0.39 mm.

A similar construction can be employed for the 5-bit angular encoder of Figure 11.10. In this case, the angular resolution can be expressed in degrees of rotation, where $2^5 = 32$ sections correspond to 360°. Thus, the resolution is 360°/32 = 11.25°. Once again, greater angular resolution could be obtained by employing a larger number of bits.

Page 654

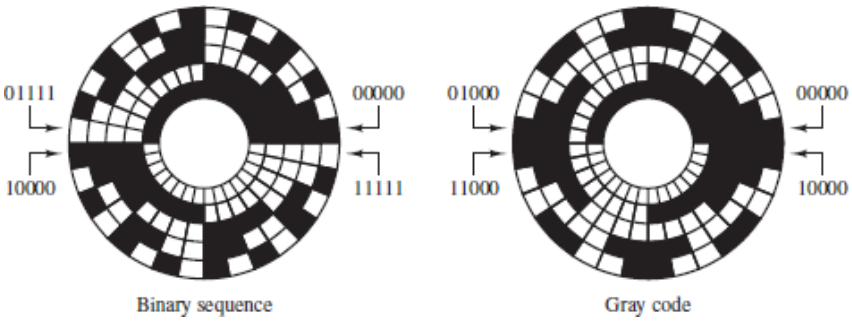


Figure 11.10 Binary and Gray code patterns for angular position encoders

EXAMPLE 11.1 Twos Complement Operations

Problem

Perform the following subtractions, using twos complement arithmetic.

1. $X - Y = 1011100 - 1110010$ (7-bits)
2. $X - Y = 10101111 - 01110011$ (8-bits)

Solution

Analysis: The twos complement subtractions are performed by replacing the operation $X - Y$ with the operation $X + (-Y)$. Thus, for the difference of the two 7-bit numbers, find the twos complement of Y and add the result to X.

$$\begin{aligned} X - Y &= 1011100 - 1110010 = 1011100 + (2^7 - 1110010) \\ &= 1011100 + 0001110 = 1101010 \end{aligned}$$

Next, add the *sign bit* (in boldface type) in front of the result since the difference $X - Y$ is a negative number.

$$X - Y = \mathbf{1}1101010$$

Repeat the procedure for the difference of the two 8-bit numbers.

$$\begin{aligned} X - Y &= 10101111 - 01110011 = 10101111 + (2^8 - 01110011) \\ &= 10101111 + 10001101 = 00111100 \\ &= \mathbf{0}00111100 \end{aligned}$$

where the first digit is a 0 because $X - Y$ is a positive number.

Page 655

EXAMPLE 11.2 Conversion From Binary to Hexadecimal

Problem

Convert the following binary numbers to hexadecimal form.

1. 100111
2. 1011101
3. 11001101
4. 101101111001
5. 100110110
6. 1101011011

Solution

Analysis: A simple method for binary to hexadecimal conversion consists of grouping each binary number into 4-bit groups and then performing the conversion for each 4-bit word following Table 11.6:

1. $100111_2 = 0010_20111_2 = 27_{16}$
2. $1011101_2 = 0101_21101_2 = 5D_{16}$
3. $11001101_2 = 1100_21101_2 = CD_{16}$
4. $101101111001_2 = 1011_20111_21001_2 = B79_{16}$
5. $100110110_2 = 0001_20011_20110_2 = 136_{16}$
6. $1101011011_2 = 0011_20101_21011_2 = 35B_{16}$

Comments: To convert from hexadecimal to binary, replace each hexadecimal number with the equivalent 4-bit nibble.

CHECK YOUR UNDERSTANDING

Convert the following decimal numbers to binary form.

a. 39
b. 59
c. 512
d. 0.4475
e. $\frac{25}{32}$
f. 0.796875
g. 256.75
h. 129.5625
i. 4,096.90625

Convert the following binary numbers to decimal.

a. 1101

b. 11011
c. 10111
d. 0.1011
e. 0.001101
f. 0.001101101
g. 111011.1011
h. 1011011.001101
i. 10110.0101011101

Answer: (a) 10011, (b) 11011, (c) 100000000, (d) 0.0111001010000, (e) 0.11001, (f) 0.110011, (g) 100000000.11, (h) 10000001.1001, (i) 10000000000000.11101; (a) 13, (b) 27, (c) 23, (d) 0.6875, (e) 0.203125, (f) 0.212890625, (g) 59.6875, (h) 91.203125, (i) 22.340820312 5

Page 656

CHECK YOUR UNDERSTANDING

Perform the following additions and subtractions. Express the answer in decimal form for (a) through (d) and in binary form for (e) through (h).

a. $1001.1_2 + 1011.01_2$
b. $100101_2 + 100101_2$
c. $0.1011_2 + 0.1101_2$
d. $1011.01_2 + 1001.11_2$
e. $64_{10} - 32_{10}$
f. $127_{10} - 63_{10}$
g. $93.5_{10} - 42.75_{10}$
h. $(84\frac{9}{32})_{10} - (48\frac{5}{16})_{10}$

Answer: (a) 20.75_{10}, (b) 74_{10}, (c) 1.5_{10}, (d) 21_{10}, (e) 100000_2, (f) 1000000_2, (g) 1100010.11_2, (h) 10001.11111_2

CHECK YOUR UNDERSTANDING

How many possible numbers can be represented in a 12-bit word?

If we use an 8-bit word with a sign bit (7 magnitude bits plus 1 sign bit) to represent voltages −5 and +5 V, what is the smallest increment of voltage that can be represented?

Answer: 4,096; 39 mV

CHECK YOUR UNDERSTANDING

Find the twos complement of the following binary numbers.

a. 11101001
b. 10010111
c. 1011110

Answer: (a) 00010111, (b) 01101001, (c) 01000010

CHECK YOUR UNDERSTANDING

Convert the following numbers from hexadecimal to binary or from binary to hexadecimal.

a. F83
b. 3C9
c. A6
d. 110101110_2
e. 10111001_2
f. 11011101101_2

Convert the following numbers from hexadecimal to binary, and find their twos complements.

a. F43

a. 2B9

c. A6

Answer: (a) 1111000011, (b) 0011100100, (c) 10100110, (d) 1AE, (e) B9, (f) 6ED; (a) 0000 1011 1101, (b) 1101 0100 0111, (c) 0101 1010

Page 657

11.3 BOOLEAN ALGEBRA AND LOGIC GATES

The mathematics associated with the binary number system (and with the more general field of logic) is called *boolean*, in honor of the English mathematician George Boole, who published a treatise in 1854 entitled "An Investigation of the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities." The development of a *logical algebra*, as Boole called it, is one of the results of his investigations. The variables in a boolean, or logic, expression can take only one of two values, usually represented by the numbers 0 and 1. These variables are sometimes referred to as true (1) and false (0). This convention is normally referred to as **positive logic**. (There is also a **negative logic** convention in which the roles of logic 1 and logic 0 are reversed.)

Analysis of **logic functions**, that is, functions of logical (boolean) variables, can be carried out in terms of truth tables. A truth table is a listing of all the possible values that each of the boolean variables can take and of the corresponding value of the desired function. **Logic gates** represent these functions and can be constructed from transistors and used to implement logic functions.

AND and OR Gates

The basis of **boolean algebra** lies in the operations of **logical addition**, or the **OR** operation, and **logical multiplication**, or the **AND** operation. Both of these find a correspondence in simple logic gates. Logical addition, although represented by the symbol +, differs from conventional algebraic addition, as shown in the last rule listed in Table 11.9. Note that this rule also differs from the last rule of binary addition studied in Section 11.2. Logical addition can be represented by the logic

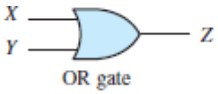
gate called an **OR gate**, whose symbol, inputs, and outputs are shown in Figure 11.11. The OR gate represents the following logical statement:

If either X or Y is true (1), then Z is true (1). Logical OR (11.1)

This rule is embodied in electronic gates, in which a logic 1 corresponds, say, to a 5-V signal and a logic 0 to a 0-V signal.

Table 11.9 Rules for logical addition (OR)

logical addition (OR)
0 + 0 = 0
0 + 1 = 1
1 + 0 = 1
1 + 1 = 1



OR gate

X	Y	Z
0	0	0
0	1	1
1	0	1
1	1	1

Truth table

Figure 11.11 Logical addition and the OR gate

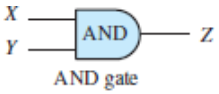
Logical multiplication is denoted by the center dot · , is defined by the rules of Table 11.10, and is represented by the **AND gate**, which is shown in Figure 11.12. The AND gate corresponds to the following logical statement:

If both X and Y are true (1), then Z is true (1). Logical AND (11.2)

One can easily envision logic gates (AND and OR) with an arbitrary number of inputs; three- and four-input gates are not uncommon.

Table 11.10 Rules for logical multiplication (AND)

$0 \cdot 0 = 0$
$0 \cdot 1 = 0$
$1 \cdot 0 = 0$
$1 \cdot 1 = 1$



AND gate

X	Y	Z
0	0	0
0	1	0
1	0	0
1	1	1

Truth table

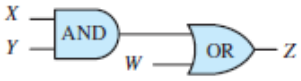
Figure 11.12 Logical multiplication and the AND gate

The rules that define a logic function are often represented in a tabular form known as a **truth table**. Truth tables for the AND and OR gates are shown in Page 658Figures 11.11 and 11.12. A truth table is nothing more than a tabular summary of all possible outputs of a logic gate, given all possible input values. If the number of inputs is 3, the number of possible combinations grows from 4 to 8, but the basic idea is unchanged. Truth tables are very useful in defining logic functions. A typical logic design problem might specify requirements such as "the output Z shall be logic 1 only when the condition ($X = 1$ AND $Y = 1$) OR ($W = 1$) occurs, and shall be logic 0 otherwise." The truth table for this particular logic function is shown in Figure 11.13.

Logic gate realization of the statement "the output Z shall be logic 1 only when the condition $(X = 1$ AND $Y = 1)$ OR $(W = 1)$ occurs, and shall be logic 0 otherwise."

X	Y	W	Z
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

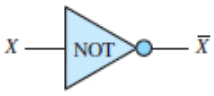
Truth table



Solution using logic gates

Figure 11.13 Example of logic function implementation with logic gates

The AND and OR gates form the basis of all logic design in conjunction with the **NOT gate**. The NOT gate is essentially a single-input, single-output inverter that provides the complement of the input logic variable. The complement of a logic variable X is denoted by $\overline{X}$, as shown in Figure 11.14.



NOT gate

X	$\overline{X}$
1	0
0	1

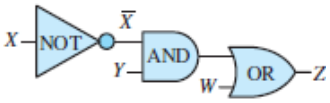
Truth table for NOT gate

Figure 11.14 Complements and the NOT gate

To illustrate the use of the NOT gate, or inverter, return to the design example of Figure 11.13, where the output of a logic circuit is $Z = 1$ only if $X = 0$ AND $Y = 1$ OR if $W = 1$. This example could also be stated as follows: "The output Z shall be logic 1 only when the condition $(\overline{X} = 1$ AND $Y = 1)$ OR $(W = 1)$ occurs, and shall be logic 0 otherwise." The formal solution to this elementary design exercise is illustrated in Figure 11.15.

X	$\bar{X}$	Y	W	Z
0	1	0	0	0
0	1	0	1	1
0	1	1	0	1
0	1	1	1	1
1	0	0	0	0
1	0	0	1	1
1	0	1	0	0
1	0	1	1	1

Truth table



Solution using logic gates

Figure 11.15 Solution of a logic problem using logic gates

Table 11.11 lists some of the rules of boolean algebra; each of these can be proved by using a truth table, as will be shown in examples and exercises. An example of **proof by perfect induction** for rule 16 is given in Figure 11.16 in the form of a truth table. This technique can be employed to prove any of the laws of Table 11.11, which in turn can be used to simplify logic expressions.

Page 659

Table 11.11 Rules of boolean algebra

1.	$0 + X = X$	
2.	$1 + X = 1$	
3.	$X + X = X$	
4.	$X + \overline{X} = 1$	
5.	$0 \cdot X = 0$	
6.	$1 \cdot X = X$	
7.	$X \cdot X = X$	
8.	$X \cdot \overline{X} = 0$	
9.	$\overline{\overline{X}} = X$	
10.	$X + Y = Y + X$	Commutative law
11.	$X \cdot Y = Y \cdot X$	
12.	$X + (Y + Z) = (X + Y) + Z$	Associative law
13.	$X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$	
14.	$X \cdot (Y + Z) = X \cdot Y + X \cdot Z$	Distributive law
15.	$X + X \cdot Z = X$	Absorption law
16.	$X \cdot (X + Y) = X$	
17.	$(X + Y) \cdot (X + Z) = X + Y \cdot Z$	
18.	$X + \overline{X} \cdot Y = X + Y$	
19.	$X \cdot Y + Y \cdot Z + \overline{X} \cdot Z = X \cdot Y + \overline{X} \cdot Z$	

X	Y	$X+Y$	$X\cdot(X+Y)$
0	0	0	0
0	1	1	0
1	0	1	1
1	1	1	1

Figure 11.16 Proof of rule 16 by perfect induction

De Morgan's Laws

Two very important logic rules are known as **De Morgan's laws**. These laws state that AND and OR functions can be interchanged by making appropriate NOT operations. In terms of Boolean algebra these theorems are

$$\overline{(X+Y)} = \overline{X} \cdot \overline{Y} \tag{11.3}$$

De Morgan's laws

$$\overline{(X \cdot Y)} = \overline{X} + \overline{Y} \tag{11.4}$$

Notice the **duality** that exists between AND and OR operations.

Any logic function can be implemented using only OR and NOT gates or only AND and NOT gates.

De Morgan's laws can be visualized in terms of logic gates and the associated truth tables, as shown in Figure 11.17.

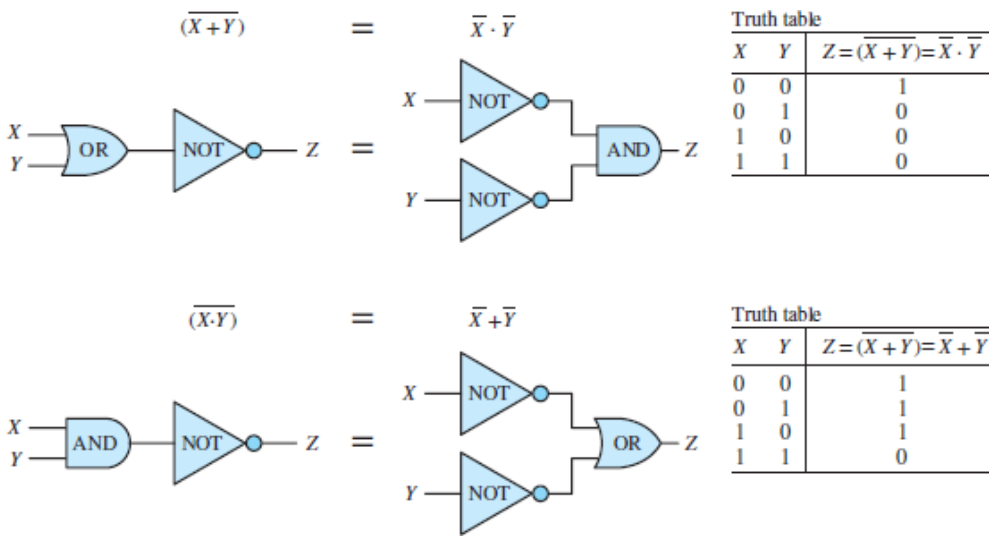


Figure 11.17 De Morgan's laws

An important consequence of De Morgan's laws is the ability to express any logic function as a **sum of products** (SOP) and/or as a **product of sums** (POS), as shown in Figure 11.18. The two forms are logically equivalent; however, one may be simpler to implement with logic gates.

Page 660

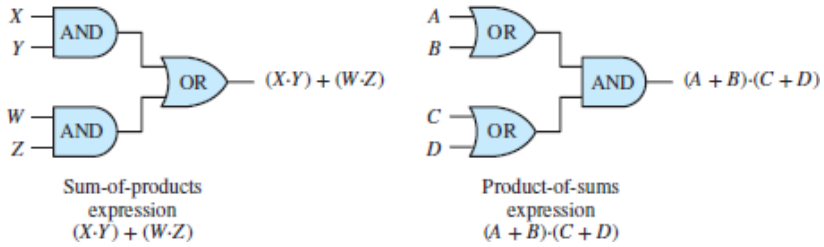
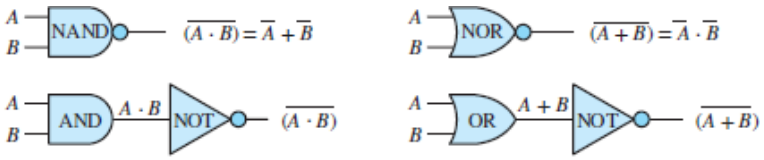


Figure 11.18 Sum-of-products and product-of-sums logic functions

NAND and NOR Gates

In addition to AND and OR gates, the complementary forms of these gates, called NAND and NOR, are commonly used in practice. In fact, NAND and NOR gates form the basis of most practical logic circuits. Figure 11.19 depicts these two gates Page 661and illustrates how they can be easily interpreted in terms of AND, OR, and NOT gates by virtue of De Morgan's laws. It is easy to verify that the logic function

implemented by the NAND and NOR gates corresponds, respectively, to AND and OR gates followed by an inverter. It is very important to note that, by De Morgan's laws, the NAND gate performs a *logical addition* on the *complements* of the inputs, while the NOR gate performs a *logical multiplication* on the *complements* of the inputs. Consequently, any logic function could be implemented with either NOR or NAND gates only.



A	B	$\overline{A}$	$\overline{B}$	$\overline{(A \cdot B)}$
0	0	1	1	1
0	1	1	0	1
1	0	0	1	1
1	1	0	0	0

NAND gate

A	B	$\overline{A}$	$\overline{B}$	$\overline{(A + B)}$
0	0	1	1	1
0	1	1	0	0
1	0	0	1	0
1	1	0	0	0

NOR gate

Figure 11.19 Equivalence of NAND and NOR gates with AND and OR gates

The XOR (Exclusive OR) Gate

It is rather common practice for a manufacturer of integrated circuits to provide common combinations of logic circuits in a single integrated-circuit (IC) package. An example of this idea is provided by the **exclusive OR (XOR) gate**, which provides a logic function similar, but not identical, to the OR gate we have already studied. The XOR gate acts as an OR gate, except when its inputs are all logic 1s; in this case, the output is a logic 0 (thus the term *exclusive*). Figure 11.20 shows the logic circuit symbol adopted for this gate and the corresponding truth table. The logic function implemented by the XOR gate is the following: either X or Y, but not both. This description can be extended to an arbitrary number of inputs.

X	Y	Z
0	0	0
0	1	1
1	0	1
1	1	0

Truth table

Figure 11.20 XOR gate

The symbol adopted for the exclusive OR operation is $\oplus$. The XOR gate can be obtained by a combination of basic gates. For example, the XOR function can be expressed as $Z = X \oplus Y = (X + Y) \cdot (\overline{X \cdot Y})$, and realized by means of the circuit shown in Figure 11.21.

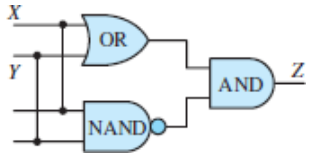


Figure 11.21 Realization of an XOR gate

Common IC logic gate configurations are typically available in both of the two more common device families, TTL and CMOS.

FOCUS ON MEASUREMENTS

Fail-Safe Autopilot Logic

This example aims to illustrate the significance of De Morgan's laws and of the duality of the sum-of-products and product-of-sums forms. Suppose that a fail-safe autopilot system in a commercial aircraft requires that, prior to initiating a takeoff or landing maneuver, the following check be passed: two of three possible pilots must be available. The three possibilities are the pilot, the copilot, and the autopilot. Imagine further that there exist switches in the pilot and copilot seats that are turned on by the weight of the crew, and that a self-check circuit exists to verify the proper

operation of the autopilot system. Let the variable X denote the pilot state (1 if the pilot is sitting at the controls), Y denote the same condition for the copilot, and Z denote the state of the autopilot, where $Z = 1$ indicates that the autopilot is functioning. Then since we wish two of these conditions to be active before the maneuver can be initiated, the logic function corresponding to "system ready" is

$$f = X \cdot Y + X \cdot Z + Y \cdot Z$$

This can also be verified by a truth table.

Page 662

Pilot	Copilot	Autopilot	System ready
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

The function f defined above is based on the notion of a *positive check*; that is, it indicates when the system is ready. Apply De Morgan's laws to the function f, which is in sum-of-products form:

$$\bar{f} = g = \overline{X \cdot Y + X \cdot Z + Y \cdot Z} = (\bar{X} + \bar{Y}) \cdot (\bar{X} + \bar{Z}) \cdot (\bar{Y} + \bar{Z})$$

The function g, in product-of-sums form, conveys exactly the same information as the function f, but it performs a negative check; in other words, g verifies the *system not ready condition*. Clearly, whether one chooses to implement the function in one form or another is simply a matter of choice; the two forms give exactly the same information.

EXAMPLE 11.3 Simplification of Logical Expression

Problem

Using the rules of Table 11.11, simplify the following function.

$$f(A, B, C, D) = \overline{A} \cdot \overline{B} \cdot D + \overline{A} \cdot B \cdot D + B \cdot C \cdot D + A \cdot C \cdot D$$

Solution

Find: Simplified expression for logical function of four variables.

Analysis:

$$\begin{aligned}
f &= \overline{A} \cdot \overline{B} \cdot D + \overline{A} \cdot B \cdot D + B \cdot C \cdot D + A \cdot C \cdot D \\
&= \overline{A} \cdot D \cdot (\overline{B} + B) + B \cdot C \cdot D + A \cdot C \cdot D && \text{Rule 14} \\
&= \overline{A} \cdot D + B \cdot C \cdot D + A \cdot C \cdot D && \text{Rule 4} \\
&= (\overline{A} + A \cdot C) \cdot D + B \cdot C \cdot D && \text{Rule 14} \\
&= (\overline{A} + C) \cdot D + B \cdot C \cdot D && \text{Rule 18} \\
&= \overline{A} \cdot D + C \cdot D + B \cdot C \cdot D && \text{Rule 14} \\
&= \overline{A} \cdot D + C \cdot D \cdot (1 + B) && \text{Rules 6 and 14} \\
&= \overline{A} \cdot D + C \cdot D = (\overline{A} + C) \cdot D && \text{Rules 2 and 14}
\end{aligned}$$

Page 663

EXAMPLE 11.4 Realizing Logic Functions From Truth Tables

Problem

Realize the logic function described by the truth table below.

A	*B*	*C*	*y*
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

Solution

Known Quantities: Value of function y (A, B, C) for each possible combination of logical variables A, B, C.

Find: Logical expression realizing the function y.

Analysis: Express y as the sum of the products of the three variables for each combination that yields $y = 1$. If the value of a variable is 1, use the uncomplemented variable. If it's 0, use the complemented variable. For example, the second row (first instance of $y = 1$) would yield the term $\overline{A} \cdot \overline{B} \cdot C$. Thus,

$$\begin{aligned} y &= \overline{A} \cdot \overline{B} \cdot C + \overline{A} \cdot B \cdot C + A \cdot \overline{B} \cdot \overline{C} + A \cdot \overline{B} \cdot C + A \cdot B \cdot \overline{C} + A \cdot B \cdot C \\ &= \overline{A} \cdot C(\overline{B} + B) + A \cdot \overline{B} \cdot (\overline{C} + C) + A \cdot B \cdot (\overline{C} + C) \\ &= \overline{A} \cdot C + A \cdot \overline{B} + A \cdot B = \overline{A} \cdot C + A \cdot (\overline{B} + B) = \overline{A} \cdot C + A = A + C \end{aligned}$$

Thus, the function is a two-input OR gate, as shown in Figure 11.22.

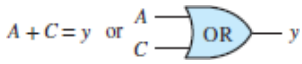


Figure 11.22

Comments: The derivation above has made use of two rules from Table 11.11: rules 4 and 18. Notice that the variable B does not appear in the final realization.

EXAMPLE 11.5 De Morgan's Laws and Product-of-Sums Expressions

Problem

Realize the logic function $y = A+B \cdot C$ in product-of-sums form. Implement the solution, using AND, OR, and NOT gates.

Solution

Known Quantities: Logical expression for the function $y(A, B, C)$.

Find: Physical realization using AND, OR, and NOT gates.

Analysis: We use the fact that $\overline{\overline{y}} = y$ and apply De Morgan's laws as follows:

$$\bar{y} = \overline{A + (B \cdot C)} = \bar{A} \cdot \overline{(B \cdot C)} = \bar{A} \cdot (\bar{B} + \bar{C})$$
$$\bar{\bar{y}} = y = \overline{\bar{A} \cdot (\bar{B} + \bar{C})}$$

The original sum-of-products function is realized using complements of each variable (obtained using NOT gates) and is finally complemented as shown in Figure 11.23.

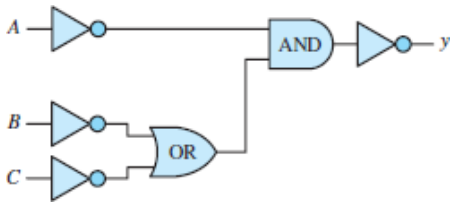


Figure 11.23

Comments: It should be evident that the original sum-of-products expression, which could be implemented with just one AND and one OR gate, has a much more efficient realization.

EXAMPLE 11.6 Realizing the AND Function With NAND Gates

Problem

Use a truth table to show that the AND function can be realized using only NAND gates, and show the physical realization.

Solution

Known Quantities: AND and NAND truth tables.

Find: AND realization using NAND gates.

Assumptions: Consider two-input functions and gates.

Analysis: The truth table below summarizes the two functions:

A	B	NAND $\overline{A \cdot B}$	AND $A \cdot B$
0	0	1	0
0	1	1	0
1	0	1	0
1	1	0	1

To realize the AND function, simply invert the output of a NAND gate. Observe that a NAND gate with its inputs tied together acts as an inverter, which can be verified in the above truth table by looking at the NAND output for the input combinations 0–0 and 1–1, or by referring to Figure 11.24. The final realization is shown in Figure 11.25.

A	B (= A)	A·B	$\overline{(A \cdot B)}$
0	0	0	1
1	1	1	0

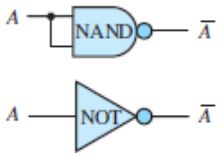


Figure 11.24 NAND gate as an inverter

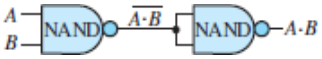


Figure 11.25

Page 665

Comments: NAND (and NOR) gates are well suited to implement functions that contain complemented products. Complementary logic gates arise naturally from the inverting characteristics of transistor switches.

EXAMPLE 11.7 Realizing the AND Function With NOR Gates

Problem

Show that the AND function can be realized using only NOR gates, and determine the physical realization.

Solution

Known Quantities: AND and NOR functions.

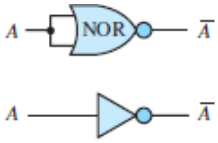
Find: AND realization using NOR gates.

Assumptions: Consider two-input functions and gates.

Analysis: This problem can be solved using De Morgan's laws. The output of an AND gate can be expressed as $f = A \cdot B$. Using De Morgan's theorem, write

$$f = \overline{\overline{f}} = \overline{\overline{A \cdot B}} = \overline{\overline{A} + \overline{B}}$$

The above function is implemented by noting that a NOR gate with its input tied together acts as a NOT gate (Figure 11.26). Thus, the logic circuit of Figure 11.27 provides the desired answer.



A	$B\ (=A)$	$A + B$	$\overline{(A + B)}$
0	0	0	1
1	1	1	0

Figure 11.26 NOR gate as an inverter

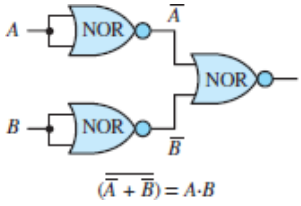


Figure 11.27

Comments: NOR (and NAND) gates are well suited to implement functions that contain complemented products. Complementary logic gates arise naturally from the inverting characteristics of transistor switches. As a result, such gates are commonly employed in practice.

EXAMPLE 11.8 Realizing a Function With NAND and NOR Gates

Problem

Realize the following function, using only NAND and NOR gates:

$$y = \overline{(A \cdot B) + C}$$

Page 666

Solution

Known Quantities: Logical expression for y.

Find: Realization of y using only NAND and NOR gates.

Assumptions: Consider two-input functions and gates.

Analysis: Refer to Examples 11.6 and 11.7 and realize the term $Z = \overline{(A \cdot B)}$ using a two-input NAND gate, and the term $\overline{Z + C}$ using a two-input NOR gate. The solution is shown in Figure 11.28.

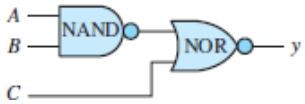


Figure 11.28

EXAMPLE 11.9 Half Adder

Problem

Analyze the half adder circuit of Figure 11.29.

Solution

Known Quantities: Logic circuit.

Find: Truth table, functional description.

Schematics, Diagrams, Circuits, and Given Data: Figure 11.29.

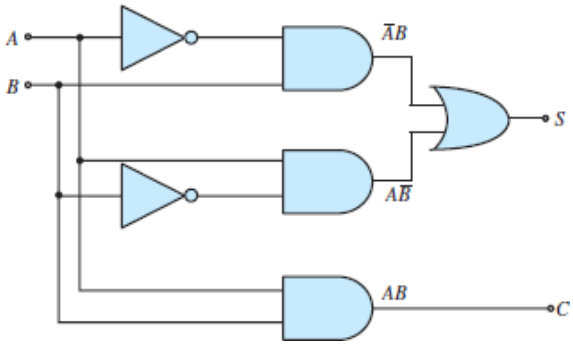


Figure 11.29 Logic circuit realization of a half adder

Analysis: The addition of two binary digits was summarized in Table 11.2. It is important to observe that when both A and B are equal to 1, the sum requires two digits: the lower digit is a 0, and there also is a carry of 1. Thus, the circuit representing this operation must give an output consisting of two digits. Figure 11.29 shows a circuit called a *half adder* that performs binary addition providing two output bits: the sum S and the carry C.

A logic statement for the rule of addition can be written as follows: S is 1 if A is 0 and B is 1, or if A is 1 and B is 0; C is 1 if A and B are 1. In terms of a logic function, we can express this statement with the following logical expressions:

$$S = \bar{A}B + A\bar{B} \quad \text{and} \quad C = AB$$

The circuit of Figure 11.29 implements this function using NOT, AND, and OR gates.

EXAMPLE 11.10 Full Adder

Problem

Analyze the full adder circuit of Figure 11.30.

Solution

Known Quantities: Logic circuit.

Find: Truth table, functional description.

Schematics, Diagrams, Circuits, and Given Data: Figure 11.30.

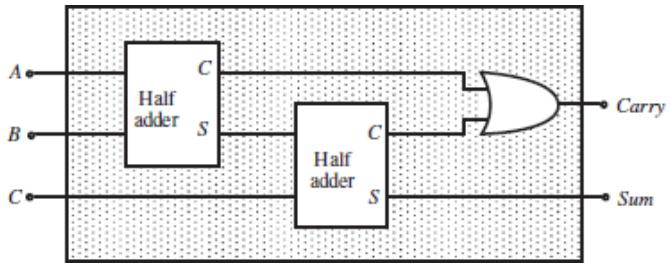


Figure 11.30 Logic circuit realization of a full adder

Analysis: A *full adder* is a circuit capable of performing a complete 2-bit addition, including taking a carry from a preceding operation. The circuit of Figure 11.30 uses two half adders, such as the one described in Example 11.9, and an OR gate to process the addition of 2 bits, *A* and *B*, plus the possible carry from a preceding addition from another (half or full) adder circuit. The truth table below illustrates this operation.

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Sum	0	0	0	1	0	1	1	1
Carry	0	1	1	0	1	0	0	1

Truth table for full adder

Comments: To perform the addition of two 4-bit nibbles, we would need a half adder for the first column (LSB), and a full adder for each additional column, that is, three full adders.

CHECK YOUR UNDERSTANDING

Prepare a step-by-step truth table for the following logic expressions.

a. $\overline{(X+Y+Z)} + (X \cdot Y \cdot Z) \cdot \overline{X}$

b. $\overline{X} \cdot Y \cdot Z + Y \cdot (Z + W)$

c. $(X \cdot \overline{Y} + Z \cdot \overline{W}) \cdot (W \cdot X + \overline{Z} \cdot Y)$

(*Hint:* Your truth table must have 2^n entries, where n is the number of logic variables.)

Page 668

Answer:

(a)

X	Y	Z	Result
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

(b)

X	Y	Z	W	Result
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	0

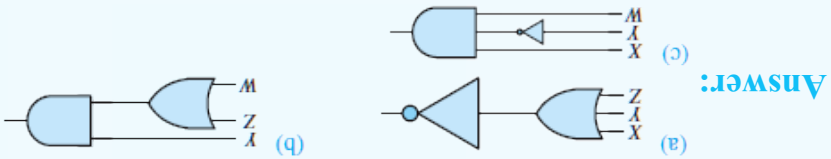
(c)

X	Y	Z	W	Result
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	0

CHECK YOUR UNDERSTANDING

Implement the three logic functions of the previous Check Your Understanding exercise using the smallest number of AND, OR, and NOT gates only.

Answer:

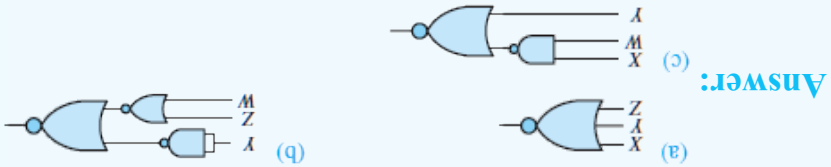


Page 669

CHECK YOUR UNDERSTANDING

Implement the three logic functions of the previous Check Your Understanding exercise using the least number of NAND and NOR gates only. (*Hint:* Use De Morgan's laws and the fact that $\overline{\overline{f}} = f$.)

Answer:



CHECK YOUR UNDERSTANDING

Show that one can obtain an OR gate by using NAND gates only. (*Hint:* Use three NAND gates.)

CHECK YOUR UNDERSTANDING

Show that the XOR function can also be expressed as $Z = X \cdot \overline{Y} + Y \cdot \overline{X}$. Realize the corresponding function using NOT, AND, and OR gates. [*Hint:* Use truth tables for the logic function Z (as defined in the exercise) and for the XOR function.]

11.4 KARNAUGH MAPS AND LOGIC DESIGN

In the design of logic functions by means of logic gates, more than one solution is usually available for the implementation of a given logic expression. Some combinations of gates can implement a given function more efficiently than others. Fortunately, there is a procedure that utilizes a map describing all possible combinations of the variables present in a logic function. This map is called a **Karnaugh map**, after its inventor. Figure 11.31 depicts the appearance of Karnaugh maps for two-, three-, and four-variable expressions in two different forms. As can be seen, the row and column assignments for two or more variables are arranged so that all adjacent terms change by only 1 bit. For example, in the three-variable map, the columns next to column 01 are columns 00 and 11. Also note that each map consists of 2^N **cells**, where N is the number of logic variables.

Each cell in a Karnaugh map contains a **minterm**, that is, a product of the N variables that appear in our logic expression (perhaps in complemented form). For example, for the case of three variables ($N = 3$), there are $2^3 = 8$ such combinations, or minterms, as shown in Figure 11.31. The content of each cell—that is, the minterm—is the product of the variables appearing at the corresponding vertical Page 670and horizontal coordinates. For example, in the three-variable map, $X \cdot Y \cdot \overline{Z}$ appears at the intersection of $X \cdot Y$ and $\overline{Z}$. The map is filled by placing a value of 1 for any combination of variables for which the desired output is a 1. For example, consider the function of three variables for which an output of 1 is desired whenever variables X, Y, and Z have the following values:

$X = 0$	$Y = 1$	$Z = 0$
$X = 0$	$Y = 1$	$Z = 1$
$X = 1$	$Y = 1$	$Z = 0$
$X = 1$	$Y = 1$	$Z = 1$

The same truth table is shown in Figure 11.32 together with the corresponding Karnaugh map.

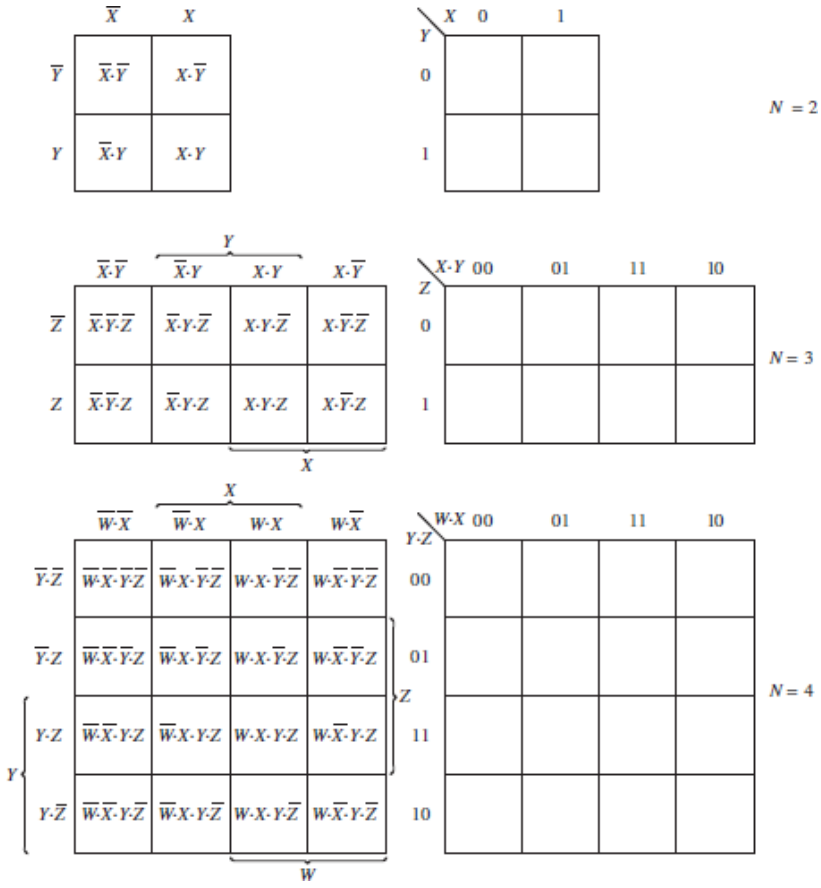
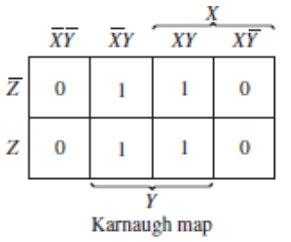


Figure 11.31 Two-, three-, and four-variable Karnaugh maps



Karnaugh map

X	Y	Z	Desired function
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	1

Truth table

Figure 11.32 Truth table and Karnaugh map representations of a logic function

The arrangement of the cells in the Karnaugh map is such that any two adjacent cells contain minterms that vary in only one variable. This property is quite useful in the design of logic functions by means of logic gates, especially if the map is considered to be continuously wrapping around itself, as if the top and bottom, and right and left, edges were touching. For the three-variable map given Page 671in Figure 11.31, for example, the cell $X \cdot \overline{Y} \cdot \overline{Z}$ is adjacent to $X \cdot \overline{Y} \cdot \overline{Z}$ if the map is "rolled" so that the right edge touches the left. Note that these two cells differ only in the variable X.[2]

Shown in Figure 11.33 is a more complex, four-variable logic function. First, define a *subgroup* as a set of 2^m adjacent cells with logical value 1, for m = 1,2,3,. . . ,N. Thus, a subgroup can consist of 1, 2, 4, 8, 16, 32, . . . cells. All possible subgroups for the four-variable map of Figure 11.33 are shown in Figure 11.34. Note that there are no four-cell subgroups in this particular case. Note also that there is some overlap between subgroups. Examples of four-cell and eight-cell subgroups are shown in Figure 11.35.

X	W	Y	Z	Desired function
0	0	0	0	1
0	0	0	1	1
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	1
0	1	1	0	1
0	1	1	1	0
1	0	0	0	0
1	0	0	1	1
1	0	1	0	1
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

Truth table for four-variable expression

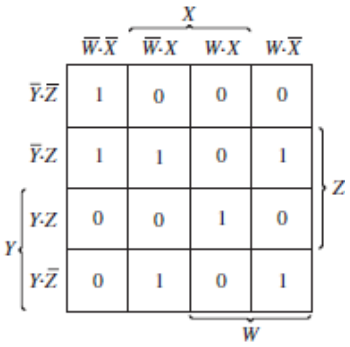
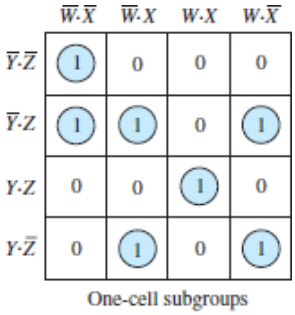
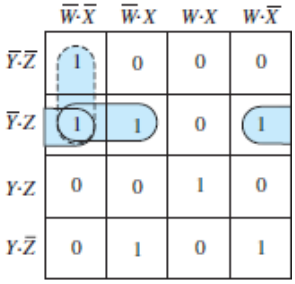


Figure 11.33 Karnaugh map for a four-variable expression



One-cell subgroups



Two-cell subgroups

Figure 11.34 One- and two-cell subgroups for the Karnaugh map of Figure 11.31

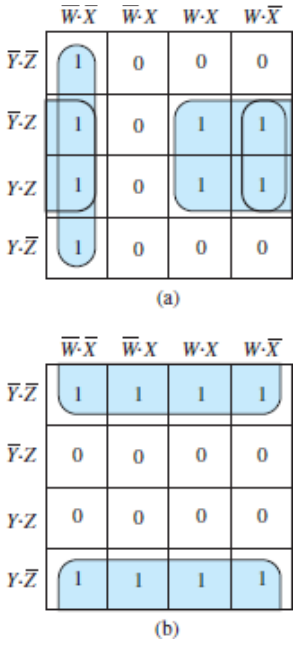


Figure 11.35 Examples of four- and eight-cell subgroups

In general, the goal is to find the largest possible subgroups to cover all the 1 entries in the map. The use of maps and subgroups in minimizing logic expressions is best explained by considering the following rule of boolean algebra:

$$Y \cdot X + Y \cdot \overline{X} = Y$$

where the variable Y could represent a product of logic variables [e.g., we could similarly write $(Z \cdot W) \cdot X + (Z \cdot W) \cdot \overline{X} = Z \cdot W$ with $Y = Z \cdot W$]. This rule is easily proved by factoring Y

$$Y \cdot (X + \overline{X})$$

and observing that $X + \overline{X} = 1$ always. Then it should be clear that variable X need not appear in the expression at all.

Consider the logic expression

$$\overline{W} \cdot X \cdot \overline{Y} \cdot Z + \overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot Z + W \cdot \overline{X} \cdot \overline{Y} \cdot Z + W \cdot X \cdot \overline{Y} \cdot Z$$

and factor it as follows:

$$\begin{aligned}\overline{W} \cdot Z \cdot \overline{Y} \cdot (X + \overline{X}) + W \cdot \overline{Y} \cdot Z \cdot (\overline{X} + X) &= \overline{W} \cdot Z \cdot \overline{Y} + W \cdot \overline{Y} \cdot Z \\ &= \overline{Y} \cdot Z \cdot (\overline{W} + W) = \overline{Y} \cdot Z\end{aligned}$$

That is quite a simplification! Consider a map in which a 1 is placed in the cells corresponding to the minterms $\overline{W} \cdot X \cdot \overline{Y} \cdot Z$, $\overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot Z$, $W \cdot \overline{X} \cdot \overline{Y} \cdot Z$, and $W \cdot X \cdot \overline{Y} \cdot Z$. It can easily be verified that the map of Figure 11.36 shows a single four-cell subgroup corresponding to the term $\overline{Y} \cdot Z$.

	$\overline{W}\cdot\overline{X}$	$\overline{W}\cdot X$	$W\cdot X$	$W\cdot\overline{X}$
$\overline{Y}\cdot\overline{Z}$	0	0	0	0
$\overline{Y}\cdot Z$	1	1	1	1
$Y\cdot Z$	0	0	0	0
$Y\cdot\overline{Z}$	0	0	0	0

Figure 11.36 Karnaugh map for the function $\overline{W} \cdot X \cdot \overline{Y} \cdot Z + \overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot Z + W \cdot \overline{X} \cdot \overline{Y} \cdot Z + W \cdot X \cdot \overline{Y} \cdot Z$

In any subgroup, one or more of the variables present will appear in both complemented *and* uncomplemented forms in all their combinations with the other variables. These variables can be eliminated. As an illustration, in the *eight-cell* subgroup case of Figure 11.37, the full-blown expression would be

$$\overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + \overline{W} \cdot X \cdot \overline{Y} \cdot \overline{Z} + W \cdot X \cdot \overline{Y} \cdot \overline{Z} + W \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + \overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + \overline{W} \cdot X \cdot \overline{Y} \cdot \overline{Z} + W \cdot X \cdot \overline{Y} \cdot \overline{Z} + W \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z}$$

However, consider the eight-cell subgroup and note that the three variables *X*, *W*, and *Z* appear in both complemented and uncomplemented form in all their combinations with the other variables, and thus can be removed from the expression. Thus, the seemingly unwieldy expression simplifies to $\overline{Y}$! In logic design terms, a simple inverter with *Y* input is sufficient to implement the expression.

	$\overline{W}\cdot\overline{X}$	$\overline{W}\cdot X$	$W\cdot X$	$W\cdot\overline{X}$
$\overline{Y}\cdot\overline{Z}$	1	1	1	1
$\overline{Y}\cdot Z$	1	1	1	1
$Y\cdot Z$	0	0	0	0
$Y\cdot\overline{Z}$	0	0	0	0

Figure 11.37

Sum-of-Products and Product-of-Sums Realizations

Logic functions can be expressed in either of two forms: sum of products (SOP) or product of sums (POS). For example, the following logic expression is in SOP form:

$$\overline{W} \cdot X \cdot \overline{Y} \cdot Z + \overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot Z + W \cdot \overline{X} \cdot \overline{Y} \cdot Z + W \cdot X \cdot \overline{Y} \cdot Z$$

A Karnaugh map can be used to determine a minimal sum-of-products expression.

FOCUS ON PROBLEM SOLVING

SUM-OF-PRODUCTS REALIZATIONS

The following steps describe the process of using an existing N variable Karna map to determine the minimal sum-of-products (SOP) realization of a logic funct

1. Identify all 2^{N-1} cell subgroups.
2. Determine the minimal logic expression for each such subgroup.
3. Identify all 2^{N-2} cell subgroups not already included in a higher dimen subgroup.
4. Determine the minimal logic expression for each such subgroup.
5. Continue this process until only isolated cells remain.
6. Determine the logic expression for each isolated cell.
7. Sum all of the above logic expressions.

Page 673

De Morgan's laws state that every SOP expression has an equivalent POS form. A simple example of a POS expression is $(W + Y) \cdot (Y + Z)$. For any particular logical expression one of the two forms may lead to a realization involving a smaller number of gates.

FOCUS ON PROBLEM SOLVING

PRODUCT-OF-SUMS REALIZATIONS

1. Group 0s in subgroups exactly as is done for 1s when seeking an expression.
2. Produce a complemented Karnaugh map by swapping X with $\overline{X}$, Y with $\overline{Y}$, a with $\overline{Z}$.
3. Each subgroup of 0s represents a *sum* of the complemented Karnaugh m elements.
4. Form the product of those sums.

An alternate POS realization method is to represent each subgroup of 0s as the product of the Karnaugh map elements, form the sum of these products, and complement the entire summation. After some manipulation using DeMorgan's laws, the result will yield an equivalent POS form. Examples 11.16 and 11.17 illustrate how one form may result in a more efficient solution than the other.

Don't-Care Conditions

Another simplification technique may be employed whenever the value of a logic function is permitted to be either 1 or 0 for certain combinations of the input variables. This situation often arises in problem specifications. A good example is the binary-coded decimal (BCD) system, in which the six 4-bit combinations [1010], [1011], [1100], [1101], [1110], and [1111] are not permitted. The algorithm used to determine the value of the BCD nibble should be indifferent to these six combinations. On the other hand, an error checking algorithm should not be; it should detect an erroneous input nibble!

Whenever it does not matter whether a position in the map is filled by a 1 or a 0, a **don't-care** entry is used, denoted by an ***x***. When forming subgroups in the Karnaugh map, each don't-care entry can be treated as either a 1 or a 0 as necessary to yield the smallest number of subgroups and therefore the greatest simplification.

FOCUS ON MEASUREMENTS

Safety Circuit for Operation of a Stamping Press

Problem:

To operate a stamping press, an operator must press two buttons (b_1 and b_2) 1 m apart from each other and away from the press (this ensures that the operator's hands cannot be caught in the press). When the buttons are pressed, the logical variables b_1 and b_2 are equal to 1. Thus, define a new variable $A = b_1 \cdot b_2$; when $A = 1$, the operator's hands are safely away from the press. In addition to this safety requirement, however, other Page 674conditions must be satisfied before the operator can activate the press. The press is designed to operate on one of two workpieces, part I and part II, but not both. Thus, acceptable logic states for the press to be operated are "part I is in the press, but not part II" and "part II is in the press, but not part I." Denote the presence of part I in the press by the logical variable $B = 1$ and the presence of part II by the logical variable $C = 1$, so that, for example, a robot used to place either part in the press could activate a pair of switches (corresponding to logical variables B and C) indicating which part, if any, is in the press. Finally, for the press to be operable, it must be "ready," meaning that it has to have completed any previous stamping operation. Let the logical variable $D = 1$ represent the ready condition. The operation of the press is now represented in terms of four logical variables, summarized in the truth table of Table 11.12. Note that only two combinations of the logical variables will result in operation of the press: $ABCD = 1011$ and $ABCD = 1101$. Using a Karnaugh map, realize the logic circuitry required to implement the truth table shown.

Table 11.12 Conditions for operation of stamping press

(A) $b_1 \cdot b_2$*	(B) Part I is in press	(C) Part II is in press	(D) Press is operable	Press operation 1 = pressing 0 = not pressing
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	1
1	1	0	0	0
1	1	0	1	1
1	1	1	0	0
1	1	1	1	0

*Both buttons (b_1, b_2) must be pressed for this to be a 1.

Solution:

Table 11.12 can be converted to a Karnaugh map, as shown in Figure 11.38. Since there are many more 0s than 1s in the table, the use of 0s in covering the map will lead to greater simplification. This will result in a product-of-sums expression. The four subgroups shown in Figure 11.38 yield the equation

$$A \cdot D \cdot (C + B) \cdot (\overline{C} + \overline{B})$$

By De Morgan's law, this equation is equivalent to

$$A \cdot D \cdot (C + B) \cdot \overline{(C \cdot B)}$$

which can be realized by the circuit of Figure 11.39.

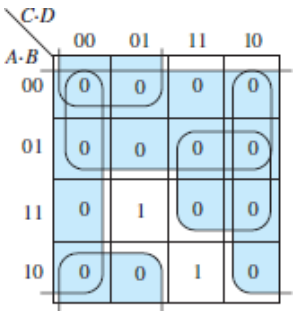


Figure 11.38

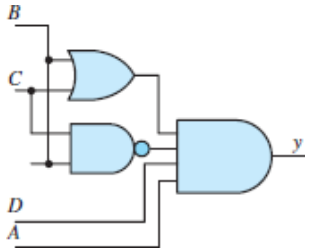


Figure 11.39

Page 675

For the purpose of comparison, the corresponding sum-of-products circuit is shown in Figure 11.40. Note that this circuit employs a greater number of gates and will therefore lead to a more expensive design.

Page 676

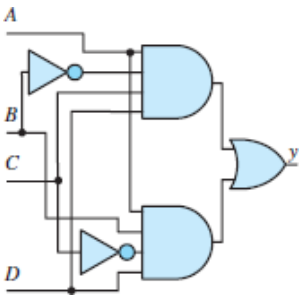


Figure 11.40

EXAMPLE 11.11 Logic Circuit Design Using Karnaugh Maps

Problem

Design a logic circuit that implements the truth table of Figure 11.41.

A	B	C	D	y
0	0	0	0	1
0	0	0	1	1
0	0	1	0	1
0	0	1	1	0
0	1	0	0	0
0	1	0	1	1
0	1	1	0	0
0	1	1	1	0
1	0	0	0	1
1	0	0	1	1
1	0	1	0	0
1	0	1	1	1
1	1	0	0	0
1	1	0	1	1
1	1	1	0	0
1	1	1	1	1

Figure 11.41

Solution

Known Quantities: Truth table for y (A, B, C, D).

Find: Realization of y.

Assumptions: Two-, three-, and four-input gates are available.

Analysis: The truth table is represented in the Karnaugh map of Figure 11.42, which is shown with values of 1 and 0 already in place. There are four subgroups in the map; three are four-cell subgroups, and one is a two-cell subgroup. The expressions for the subgroups are $\overline{A} \cdot \overline{B} \cdot \overline{D}$ for the two-cell subgroup: $\overline{B} \cdot \overline{C}$ for the subgroup that wraps around the map; $\overline{C} \cdot D$ for the 4-by-1 subgroup; and $A \cdot D$ for the square subgroup at the bottom of the map. Thus, the expression for y is

$$y = \overline{A} \cdot \overline{B} \cdot \overline{D} + \overline{B} \cdot \overline{C} + \overline{C}D + AD$$

Page 677

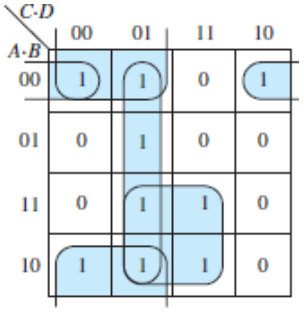


Figure 11.42 Karnaugh map for Example 11.11

The implementation of the above function with logic gates is shown in Figure 11.43.

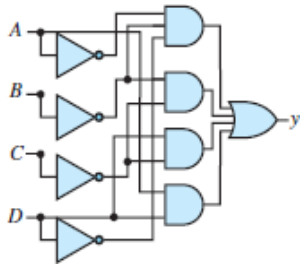


Figure 11.43 Logic circuit realization of Karnaugh map of Figure 11.42

Comments: Notice the OR gate at the far right. It produces the sum of all the individual AND products. The Karnaugh map highlighting of Figure 11.42 yields an SOP expression because all the 1s were highlighted.

EXAMPLE 11.12 Deriving a Sum-of-Products Expression from a Logic Circuit

Problem

Derive the truth table and minimum sum-of-products expression for the circuit of Figure 11.44.

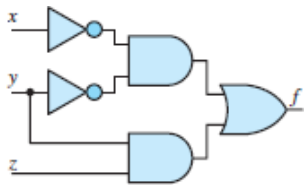


Figure 11.44

Solution

Known Quantities: Logic circuit representing $f(x, y, z)$.

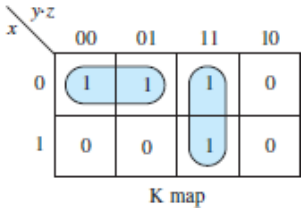
Find: Expression for f and corresponding truth table.

Analysis: The logic function corresponding to the logic circuit of Figure 11.44 is

$$f = \bar{x} \cdot \bar{y} + y \cdot z$$

The truth table corresponding to this expression and the corresponding Karnaugh map with sum-of-products covering are shown in Figure 11.45.

x	y	z	f
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1



K map

Figure 11.45

Comments: If the 0s in the Karnaugh map had been highlighted, the resulting expression would be a POS. Verify that the complexity of the circuit would be unchanged. Note also that the subgroup ($x = 0$, $yz = 01, 11$) is not used because it does not further minimize the solution.

Page 678

EXAMPLE 11.13 Realizing a Sum of Products Using Only NAND Gates

Problem

Realize the following three-input POS function in SOP form, using only two-input NAND gates. Keep in mind that a NAND gate with identical inputs acts as a NOT gate.

$$f = (\bar{x} + \bar{y}) \cdot (y + \bar{z})$$

Solution

Known quantities: $f(x, y, z)$.

Find: Logic circuit for f using only NAND gates.

Analysis: The first step is to convert the expression for *f* into an expression that can be easily implemented with NAND gates. Observe that direct application of De Morgan's laws yields

$$\bar{x} + \bar{y} = \overline{x \cdot y}$$
$$y + \bar{z} = \overline{z \cdot \bar{y}}$$

Thus, the function can be written as

$$f = (\overline{x \cdot y}) \cdot (\overline{z \cdot \bar{y}})$$

and implemented with five NAND gates, as shown in Figure 11.46.

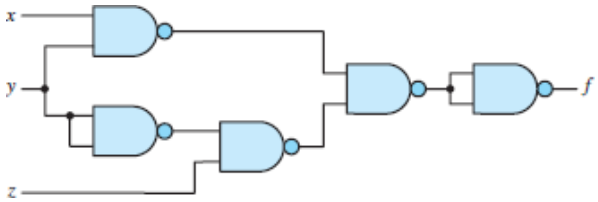


Figure 11.46

The entire function *f* can be negated twice (leaving it unchanged) before applying one of De Morgan's laws to produce the SOP form of *f*.

$$f = \overline{\overline{(x \cdot y) \cdot (\overline{z \cdot \bar{y}})}} = \overline{(x \cdot y) + (z \cdot \bar{y})}$$

The SOP form requires one NOT gate, two AND gates and one NOR gate. However, NAND technology tends to be faster than other gates and so the implementation shown in Figure 11.46 may be preferable in practice.

EXAMPLE 11.14 Simplifying Expressions by Using Karnaugh Maps

Problem

Simplify the following expression by using a Karnaugh map.

$$f = x \cdot y + \bar{x} \cdot z + y \cdot z$$

Solution

Known Quantities: $f(x, y, z)$.

Find: Minimal expression for f.

Analysis: A three-term Karnaugh map is highlighted, as shown in Figure 11.47. The 1s can be covered using just two subgroups: $f = x \cdot y + \bar{x} \cdot z$. Thus, the term $y \cdot z$ is redundant.

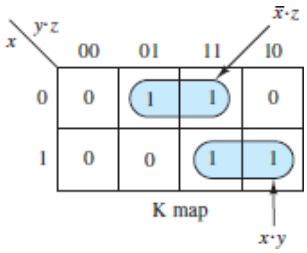


Figure 11.47

Comments: Notice that the term $y \cdot z$ is covered by the two subgroups. Thus, the expression $x \cdot y + \bar{x} \cdot z + y \cdot z$ is equivalent to, but more complicated than, $x \cdot y + \bar{x} \cdot z$.

LO

EXAMPLE 11.15 Simplifying a Logic Circuit by Using the Karnaugh Map

Problem

Derive the Karnaugh map for the circuit of Figure 11.48 and use the resulting map to produce a minimized logic expression.

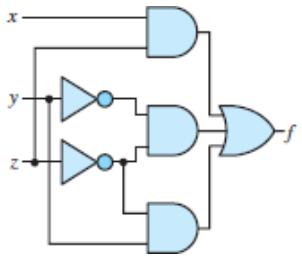


Figure 11.48

Solution

Known Quantities: Logic circuit.

Find: Minimized logic circuit.

Analysis: Determine the expression $f(x, y, z)$ from the logic circuit.

$$f = (x \cdot z) + (\bar{y} \cdot \bar{z}) + (y \cdot \bar{z})$$

This expression leads to the Karnaugh map shown in Figure 11.49, where the 1s are covered by three two-cell subgroups. However, the 1s in the Karnaugh map can be covered more efficiently by two four-cell subgroups, resulting in the simpler function $f = x + \bar{z}$ and the logic circuit shown in Figure 11.50.

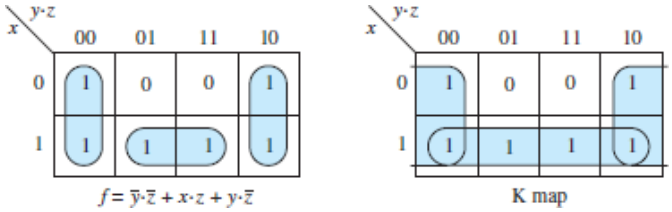


Figure 11.49

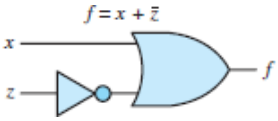


Figure 11.50

Comments: In general, a smaller number of subgroups results in a smaller number of terms in the logic expression.

EXAMPLE 11.16 Product-of-Sums Design

Problem

Realize the function *f* described by the accompanying truth table in minimal product-of-sums form. Draw the corresponding Karnaugh map.

x	*y*	*z*	*f*
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	0

Solution

Known Quantities: Truth table for logic function.

Find: Realization in minimal product-of-sums forms.

Analysis: Cover the Karnaugh map of Figure 11.51 using 0s to obtain the following function:

$$f = \bar{z} \cdot (\bar{x} + \bar{y})$$

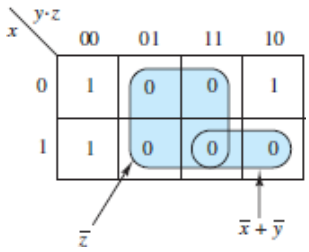


Figure 11.51

Comments: What is the equivalent SOP solution? Find it! Is it simpler?

EXAMPLE 11.17 Comparison of Sum-of-Products and Product-of-Sums Designs

Problem

Realize the function *f* described by the accompanying truth table, using both 0 and 1 coverings in the Karnaugh map.

x	*y*	*z*	*f*
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

Solution

Known Quantities: Truth table for logic function.

Find: Realization in both sum-of-products and product-of-sums forms.

Analysis:

1. *Product-of-sums expression*. Product-of-sums expressions use 0s to determine the logical expression from a Karnaugh map. Figure 11.52 depicts the Karnaugh map covering with 0s, leading to the expression

$$f = (x + y + z) \cdot (\bar{x} + \bar{y})$$

2. *Sum-of-products expression*. Sum-of-products expressions use 1s to determine the logical expression from a Karnaugh map. Figure 11.53 depicts the Karnaugh map covering with 1s, leading to the expression

$$f = (\bar{x} \cdot y) + (x \cdot \bar{y}) + (\bar{y} \cdot z) = (\bar{x} \cdot y) + (x \cdot \bar{y}) + (\bar{x} \cdot z)$$

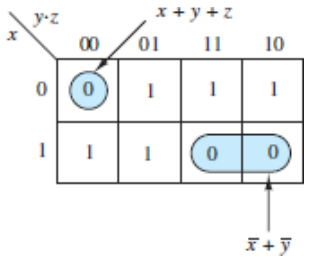


Figure 11.52

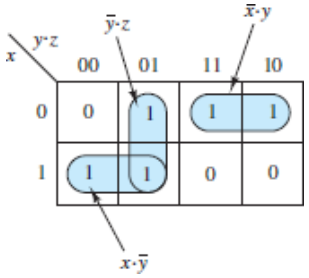


Figure 11.53

Comments: It is worthwhile to verify that both SOP forms of f are correct. The POS solution requires five gates (one three-input OR, one two-input OR, two NOT, and one AND), while the SOP solution requires six gates (one three-input OR, two NOT, and three AND). Thus, the POS solution leads to a simpler design.

EXAMPLE 11.18 Using Don't-Care Conditions to Simplify Expressions —1

Problem

Find a minimum sum-of-products realization for the expression $f(A,B,C)$.

Solution

Known Quantities: Logical expression, don't-care conditions.

Find: Minimal realization.

Schematics, Diagrams, Circuits, and Given Data:

$$f(A, B, C) = \begin{cases} 1 & \text{for } \{A, B, C\} = \{000, 010, 011\} \\ x & \text{for } \{A, B, C\} = \{100, 101, 110\} \end{cases}$$

Analysis: Cover the Karnaugh map of Figure 11.54 using 1s and don't-care entries to obtain the following minimal expression from one four-cell subgroup and one two-cell subgroup.

$$f(A, B, C) = \overline{A} \cdot B + \overline{C}$$

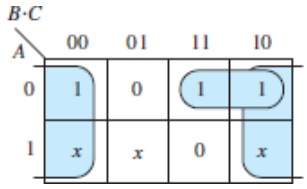


Figure 11.54

Comments: Note that one of the don't-care entries was not used, since doing so would not lead to any further simplification.

EXAMPLE 11.19 Using Don't-Care Conditions to Simplify Expressions —2

Problem

Use don't-care entries to simplify the expression

$$f(A, B, C, D) = \overline{A} \cdot \overline{B} \cdot \overline{C} \cdot D + \overline{A} \cdot \overline{B} \cdot C \cdot \overline{D} + \overline{A} \cdot \overline{B} \cdot C \cdot D + \overline{A} \cdot B \cdot \overline{C} \cdot D + A \cdot \overline{B} \cdot C \cdot D + A \cdot B \cdot \overline{C} \cdot \overline{D}$$

Solution

Known Quantities: Logical expression; don't-care conditions.

Find: Minimal realization.

Schematics, Diagrams, Circuits, and Given Data: Don't-care conditions: $f(A,B,C,D) = \{0100,0110,1010,1110\}$.

Analysis: Cover the Karnaugh map of Figure 11.55 using 1s and don't-care entries to obtain the following simplified expression from two four-cell subgroups and one two-cell subgroup.

$$f(A, B, C, D) = B \cdot \overline{D} + \overline{B} \cdot C + \overline{A} \cdot \overline{C} \cdot D$$

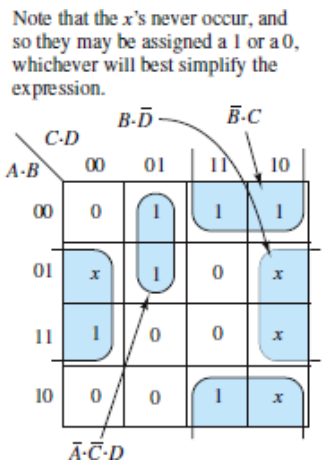


Figure 11.55

Comments: The don't-care entries can be interpreted as 0s to find the POS form. Verify that the SOP expression is simpler.

CHECK YOUR UNDERSTANDING

Simplify the following expression, using a Karnaugh map.

$$\overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + \overline{W} \cdot \overline{X} \cdot Y \cdot \overline{Z} + W \cdot X \cdot \overline{Y} \cdot \overline{Z} + W \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + W \cdot \overline{X} \cdot Y \cdot \overline{Z} + W \cdot X \cdot Y \cdot \overline{Z}$$

Simplify the following expression, using a Karnaugh map.

$$\overline{W} \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + \overline{W} \cdot \overline{X} \cdot Y \cdot \overline{Z} + W \cdot X \cdot \overline{Y} \cdot \overline{Z} + W \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} + W \cdot \overline{X} \cdot Y \cdot \overline{Z} + \overline{W} \cdot X \cdot \overline{Y} \cdot \overline{Z}$$

Answer: $W \cdot \overline{Z} + X \cdot \overline{Z}$; $\overline{Y} \cdot \overline{Z} + \overline{Z} \cdot X + Z \cdot \overline{X}$

CHECK YOUR UNDERSTANDING

Would a sum-of-products realization for Example 11.16 require fewer gates?

Answer: No

Page 682

CHECK YOUR UNDERSTANDING

Verify that the product-of-sums expression for Example 11.17 can be realized with fewer gates than the sum-of-products expression.

CHECK YOUR UNDERSTANDING

Show that the circuit of Figure 11.53 can also be obtained from the sum of products.

CHECK YOUR UNDERSTANDING

In Example 11.18, assign a value of 0 to the don't-care terms and derive the corresponding minimal expression. Is the new function simpler than the one obtained in Example 11.18?

In Example 11.18, assign a value of 1 to all don't-care terms and derive the corresponding minimal expression. Is the new function simpler than the one obtained in Example 11.18?

Answer: $f = \overline{A} \cdot B + \overline{A} \cdot \overline{C}$; No; $f = \overline{A} \cdot B + A \cdot \overline{B} + \overline{C}$; No

CHECK YOUR UNDERSTANDING

In Example 11.19, assign a value of 0 to the don't-care terms and derive the corresponding minimal expression. Is the new function simpler than the one obtained in Example 11.19?

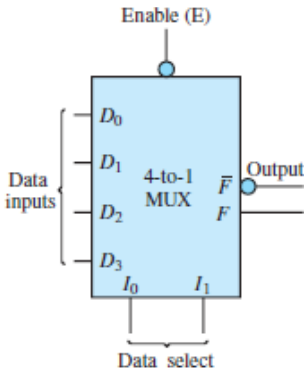
Answer: $f = A \cdot B \cdot \overline{C} \cdot \overline{D} + \overline{A} \cdot \overline{C} \cdot D + \overline{A} \cdot \overline{B} \cdot C + \overline{B} \cdot C \cdot D$; No

11.5 COMBINATIONAL LOGIC MODULES

The basic logic gates described in the previous section are used to implement more advanced functions and are often combined to form logic modules, which, thanks to modern technology, are available in compact integrated-circuit packages. In this section a few of the more common **combinational logic modules** are discussed, illustrating how these can be used to implement advanced logic functions.

Multiplexers

Multiplexers, or **data selectors**, are combinational logic circuits. A typical multiplexer (MUX) has 2^n **data lines**, n **address** (or **data select**) **lines**, and one output. In addition, other control inputs (e.g., enables) may exist. Standard, commercially available MUXs allow for n up to 4; however, two or more MUXs can be combined if a greater range is needed. The MUX allows for one of 2^n inputs to be selected as the data output. The address or data select lines determine which input is selected as the output. Figure 11.56 depicts the block diagram of a four-input ($n = 2$) MUX. Page 683The input data lines are labeled D_0, D_1, D_2, and D_3; the **data select**, or address, **lines** are labeled I_0 and I_1; and the output is available in both complemented and uncomplemented form and is thus labeled F or $\overline{F}$. Finally, an **enable** input, labeled E, is also provided, as a means of enabling or disabling the MUX: if $E = 1$, the MUX is disabled; if $E = 0$, it is enabled. The negative logic (MUX off when $E = 1$ and on when $E = 0$) is indicated by the small "bubble" at the enable input, which represents a complement operation (just as at the output of NAND and NOR gates). The enable input is useful when using a cascade of MUXs, to select a line from a large number, say, $2^8 = 256$, inputs.



Functional diagram of 4-to-1 MUX

I_1	I_0	F
0	0	D_0
0	1	D_1
1	0	D_2
1	1	D_3

Truth table of 4-to-1 MUX

Figure 11.56 4-to-1 MUX

Figure 11.57 shows the internal logic of a 4:1 MUX using exclusively NAND gates and inverters. Figure 11.58 depicts the concept of a cascade of MUXs. (In practice, one 8:1 MUX would be used instead of cascading two 4:1 MUXs.) When there are more than 16 data inputs it may be necessary to cascade MUXs.

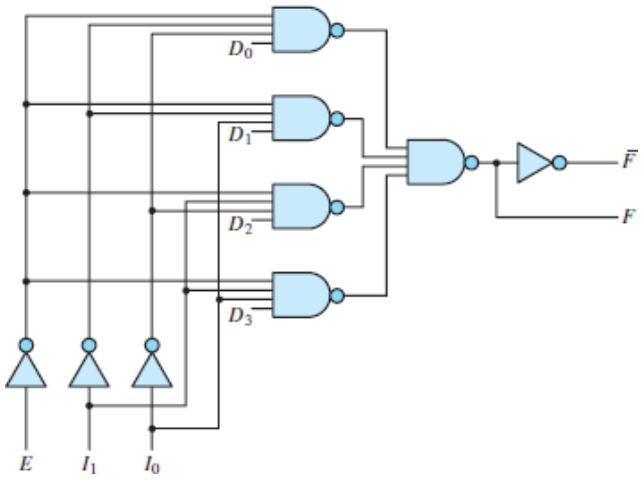


Figure 11.57 Internal logic of a 4-to-1 MUX with inverted enable logic (E = 0 enables the MUX.)

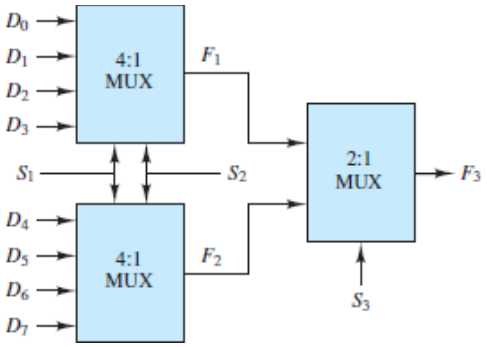


Figure 11.58 Cascade of two 4:1 MUXs with a third 2:1 MUX to select from the other two. The data select lines S_1, S_2, and S_3 determine which data input will appear at F_3.

In the design of digital systems (e.g., microprocessors), a single line is often required to carry two or more different digital signals. However, only one signal Page 684at a time can be placed on the line. A MUX allows different signals to be selected at different instants on the single output line by using clock inputs on the data select lines.

The data selector function of a 4:1 MUX is best understood in terms of Table 11.13. In this truth table, each *x* represents a don't-care entry. As can be seen from the truth table, the output selects one of the data lines depending on the values of I_1 and I_0, assuming that I_0 is the least significant bit. For example, $I_1\ I_0 = 10$ selects D_2, which means that the output F will be the same as input D_2. Similar tables can be constructed for larger MUXs.

Table 11.13

I_1	I_0	D_3	D_2	D_1	D_0	F
0	0	*x*	*x*	*x*	0	0
0	0	*x*	*x*	*x*	1	1
0	1	*x*	*x*	0	*x*	0
0	1	*x*	*x*	1	*x*	1
1	0	*x*	0	*x*	*x*	0
1	0	*x*	1	*x*	*x*	1
1	1	0	*x*	*x*	*x*	0
1	1	1	*x*	*x*	*x*	1

Read-Only Memory (ROM)

Another common technique for implementing logic functions uses **read-only memory**, or **ROM**. As the name implies, a ROM is a logic circuit that holds information in storage ("memory")—in the form of binary numbers—that cannot be altered but can be "read" by a logic circuit. A ROM is an array of memory cells, each of which can store either a 1 or a 0. The array consists of $2m \times n$ cells, where n is the number of bits in each word[3], m is the number of address lines and 2m is the number of words stored in a ROM. When an address is selected, in a fashion similar to the operation of the MUX, the binary word corresponding to the address selected appears at the output, which consists of n bits, that is, the same number of bits as the stored words. In some sense, a ROM can be thought of as a MUX that has an output consisting of a word instead of a single bit. A ROM is often used to store a *lookup table* where results of complicated computations are stored so that the results can be "looked up" instead of being computed in real time by a microprocessor.

Figure 11.59 depicts the conceptual arrangement of a ROM with $n = 4$ and $m = 2$. As an illustration, the ROM lookup table has been filled with arbitrary 4-bit words. In Figure 11.59, for an enable input of 0 (i.e., on) and values for the address lines of $I_0 = 0$ and $I_1 = 1$, the output word would be $W_2 = 0110$, so that $b_0 = 0$, $b_1 = 1$, $b_2 = 1$, $b_3 = 0$. Depending on the content of the ROM and the number of address and output lines, one could implement an arbitrary logic function.

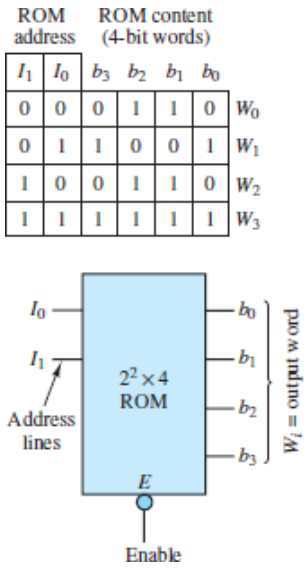


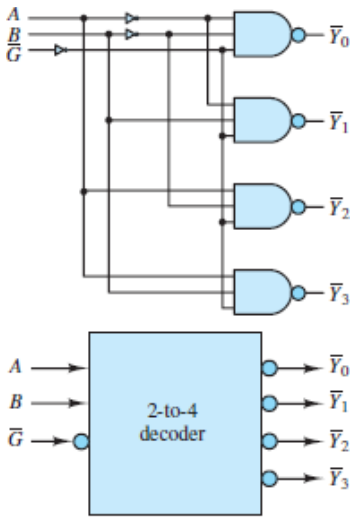
Figure 11.59 Read-only memory

Unfortunately, the data stored in read-only memories must be entered during fabrication and cannot be altered later. A much more convenient type of read-only memory is the **erasable programmable read-only memory (EPROM)**, the Page 685content of which can be easily programmed and stored and may be changed if

needed. EPROMs find use in many practical applications, because of their flexibility in content and ease of programming. The Focus on Measurements box illustrates the use of an EPROM to store a *lookup table* (LUT).

Decoders and SRAM

Decoders, which are commonly used for applications such as address decoding or memory expansion, are combinational logic circuits. A **decoder** converts a m-bit input code to a unique n-bit output code, where $m \leq n \leq 2^m$. Figure 11.60 shows the logic and block diagrams and the truth table for a 2-to-4 decoder with two inputs A and B and four outputs $\overline{Y}_0$, $\overline{Y}_1$, $\overline{Y}_2$ and $\overline{Y}_3$. The decoder has an active-low enable input $\overline{G}$. When $\overline{G}$ is logic 1, all decoder outputs are forced to logic 1 regardless of the select inputs. The decoder shown here employs active low logic such that when $\overline{G}$ is logic 0 only one of the outputs is logic 0. The other outputs are logic 1. The particular logic 0 output is determined by the select inputs A and B. Decoders are also available with active high logic such that when $\overline{G}$ is logic 0 only one of the outputs is logic 1 while all other outputs are logic 0. Decoders are also available with an active-high enable input.



Inputs			Outputs			
Enable	Select					
$\overline{G}$	A	B	$\overline{Y}_0$	$\overline{Y}_1$	$\overline{Y}_2$	$\overline{Y}_3$
1	x	x	1	1	1	1
0	0	0	0	1	1	1
0	0	1	1	0	1	1
0	1	0	1	1	0	1
0	1	1	1	1	1	0

Figure 11.60 A 2-to-4 decoder

This simple description of decoders permits a brief discussion of the internal organization of **static random-access memory (SRAM)**, which provides high-speed memory, a large bit capacity, and low cost. The memory array has a column length equal to the number of words 2m and a row length equal to the number of bits per word *N*. To select a word, an *m*-to-2m decoder is needed. The decoder inputs select one word in the memory array. To choose the desired word from the memory array, the proper address inputs are required. As an example, if the number of words in the memory array is eight, a 3-to-8 decoder is needed. Figure 11.61 shows the internal organization of a typical SRAM.

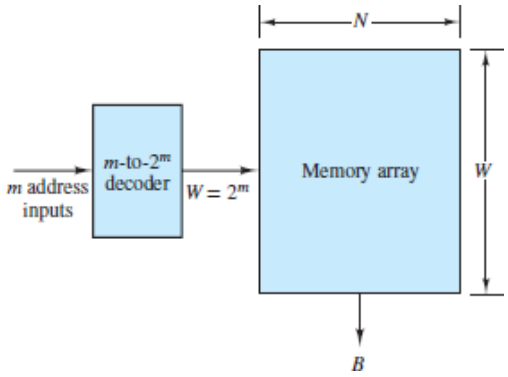


Figure 11.61 Internal organization of a SRAM

Gate Arrays and Programmable Logic Devices

Many of today's digital logic designs are deployed using **programmable logic devices** (**PLDs**). Early PLDs consisted of arrays of AND and OR gates connected by way of a programmable interconnect. Specific combinational logic functions Page 686are implemented onto these devices by programming the interconnect to create the required logical connections. More recent PLDs consist of arrays of programmable logic blocks, in lieu of basic gate arrays, combined with a programmable interconnect. These devices also incorporate *flip-flops* making them suitable for *sequential logic* (see chapter 12) designs. All PLDs can be programmed using special programming languages called **hardware description languages** (**HDLs**). Four types of PLDs are defined here:

1. ***PAL (programmable array logic):*** An early PLD containing an array of programmable AND gates whose outputs are connected to a fixed array of OR gates.
2. ***PLA (programmable logic array):*** An early PLD containing an array of programmable AND gates whose outputs are connected to an array of programmable OR gates. These devices provide more programming flexibility than PAL devices.
3. ***CPLD (complex programmable logic device):*** A modern device that combines multiple PLDs into a single package allowing for more complex logic designs. The PLDs are connected to each other via a programmable interconnect. Each PLD is paired with a flip-flop making these devices suitable for both combinational and sequential logic designs. Configuration information for these reprogrammable devices is stored in a nonvolatile **electrically erasable**

programmable read-only memory (**EEPROM**) and thus they retain their configuration between power cycles.

4. ***FPGA (field-programmable gate array):*** A modern PLD containing an array of programmable logic blocks connected by way of a programmable interconnect. Programmable logic blocks contain one or more *n*-input **look-up tables** (**LUTs**), a variety of multiplexers, and one or more flip-flops. Each *n*-input LUT can implement any arbitrary *n*-input Boolean function. These devices often have tens to hundreds of thousands of programmable logic blocks making them the most flexible of the PLDs presented here. FPGAs also include **random-access memory** (**RAM**) blocks into the programmable interconnect that can be incorporated into the logic design. Programming information for these reprogrammable devices is typically stored in volatile memory and thus they must be reprogrammed each time power is applied.

To illustrate the concept of logic design using a PLA, Figure 11.62 shows a diagram of a small PLA architecture with three inputs and two outputs. The circles represent programmable interconnects. Filled circles represent a location where a connection exists, whereas empty circles represent a location where no connection exists. The PLA shown in this figure implements two logical functions:

$$O_1 = \overline{I}_1 \cdot I_2 + \overline{I}_3$$
$$O_2 = I_1 \cdot \overline{I}_2 \cdot I_3 + I_2 \cdot \overline{I}_3$$

An equivalent logic circuit and the corresponding HDL code to program these logical functions into a PLA is shown in Figure 11.63. The HDL first defines the inputs and outputs. Next, a set of equations describes the logical functions to be implemented and to which output the result is assigned. Note that the Page 687symbol & represents the logical function AND. The symbol | represents the logical function OR.

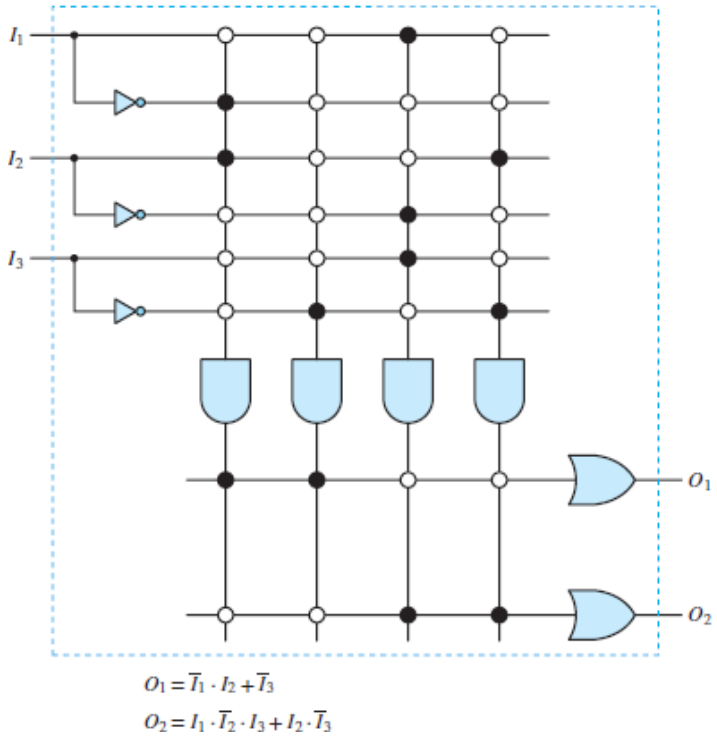


Figure 11.62 Internal structure of a simple PLA

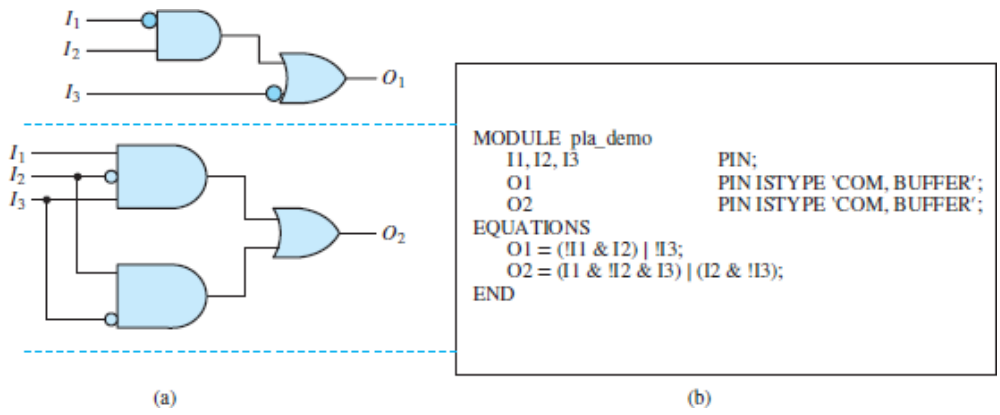


Figure 11.63 (a) Logic circuit; (b) corresponding HDL code

A second example of the use of a PLD introduces the concept of timing diagrams, which are covered in greater detail in Chapter 12. Figure 11.64 depicts a timing diagram related to an automotive fuel-injection system, in which multiple injections are to be performed. Three *pilot* injections and one *primary* injection are to be performed. The *master control* line enables the entire sequence. The resulting output sequence, shown at the bottom of the plot and labeled "injector fuel pulse,"Page 688 is the combination of the three pilot pulses and the primary pulse. Based on the timing plot of the signals shown in Figure 11.64(a), the following

inputs are used: I1=master control, I2=pilot inject #1, I3=pilot inject #2, I4=pilot inject #3, I5=primary inject, and the output O1=injector fuel pulse. The required logic function is

$$O_1 = I_1 \cdot (I_2 + I_3 + I_4 + I_5)$$

This function is realized by the code in Figure 11.64(b).

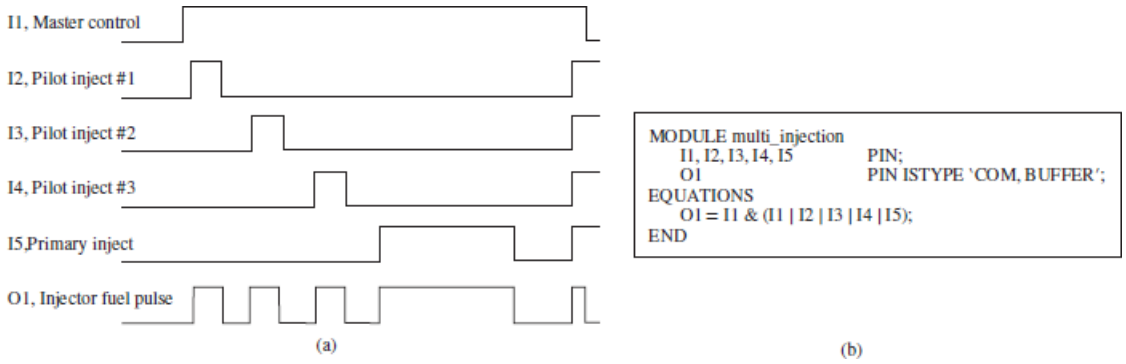
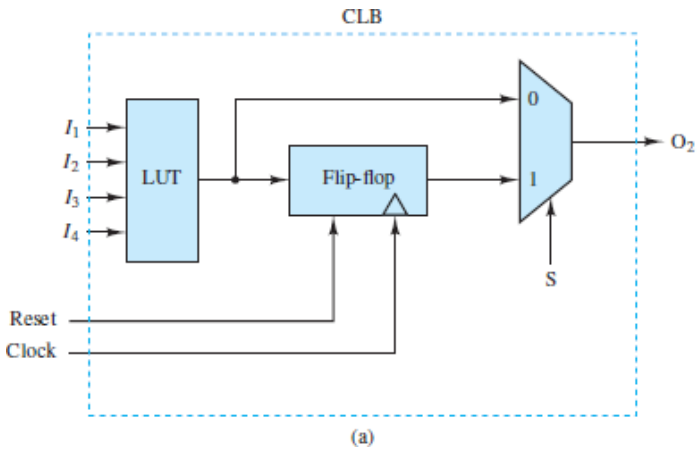


Figure 11.64 (a) Injector timing sequence; (b) sample code for multiple-injection sequence

A final example illustrates the use of a single FPGA programmable logic block, commonly referred to as a **configurable logic block** (**CLB**), to implement a logical function. Figure 11.65(a) shows the layout of a simple CLB consisting of a four-input LUT, a flip-flop, and a two-input MUX. To implement the logical function $O_2 = I_1 \cdot \overline{I}_2 \cdot I_3 + I_2 \cdot \overline{I}_3$, the LUT is programmed with the values shown in Figure 11.65(b). For a combinational logic implementation of the function, the CLB's flip-flop is bypassed by programming the *S* input of the MUX to be tied low. For a sequential logic implementation of the function, the *S* input of the MUX is tied high.



I_1	I_2	I_3	I_4	O_2
0	0	0	×	0
0	0	1	×	0
0	1	0	×	1
0	1	1	×	0
1	0	0	×	0
1	0	1	×	1
1	1	0	×	1
1	1	1	×	0

(b)

Figure 11.65 (a) A simple CLB; (b) sample contents of LUT

A typical FPGA contains an array of many thousands of CLBs. The contents of CLBs vary widely between FPGA product families and vendors.

FOCUS ON MEASUREMENTS

EPROM-based Lookup Table for Automotive Fuel-Injection System Control

One of the most common applications of EPROMs is an *arithmetic lookup table*, which is used to store computed values of certain functions, eliminating the need to compute the function. A practical application of this concept is present in every automobile manufactured in the United States since the early 1980s, as part of the exhaust emission control system. For the catalytic converter to minimize the emissions of exhaust gases (especially hydrocarbons, oxides of nitrogen, and carbon monoxide), it is necessary to maintain the *air-to-fuel ratio A/F* as close as possible to the stoichiometric ratio of 14.7 parts of air for each part of fuel. Most modern engines are equipped with fuel-injection systems that are capable of delivering accurate amounts of fuel to each individual cylinder—thus, the task of maintaining an accurate *A/F* amounts to measuring the mass of air that is aspirated into each cylinder and computing the corresponding mass of fuel. Many automobiles are equipped with a *mass airflow sensor*, capable of measuring the mass of air drawn into each cylinder during each engine cycle. Let the output of the mass airflow sensor be denoted by the variable M_A, and let this variable represent the mass of air (in

grams) actually entering a cylinder during a particular stroke. It is then desired to compute the mass of fuel M_F (also expressed in grams) required to achieve an A/F of 14.7. This computation is simply

$$M_F = \frac{M_A}{14.7}$$

Although this computation is a simple division, its actual calculation in a low-cost digital computer (such as would be used on an automobile) is rather complicated. It would be much simpler to tabulate a number of values of M_A, to precompute the variable M_F, and then to store the result of this computation in an EPROM. If the EPROM address were made to correspond to the tabulated values of air mass, and the content at each address to the corresponding fuel mass (according to the precomputed values of the expression $M_F = M_A/14.7$), it would not be necessary to perform the division by 14.7. For each measurement of air mass into one cylinder, an EPROM address is specified and the corresponding content is read. The content at the specific address is the mass of fuel required by that particular cylinder.

In practice, the fuel mass needs to be converted to a time interval corresponding to the duration of time during which the fuel injector is open. This final conversion factor can also be accounted for in the table. Suppose, for example, that the fuel injector is capable of injecting K_F g/s of fuel; then the time duration T_F during which the injector should be open to inject M_F g of fuel into the cylinder is given by

$$T_F = \frac{M_F}{K_F} \text{ s}$$

Therefore, the complete expression to be precomputed and stored in the EPROM is

$$T_F = \frac{M_A}{14.7 \times K_F} \text{ s}$$

Figure 11.66 illustrates this process graphically.

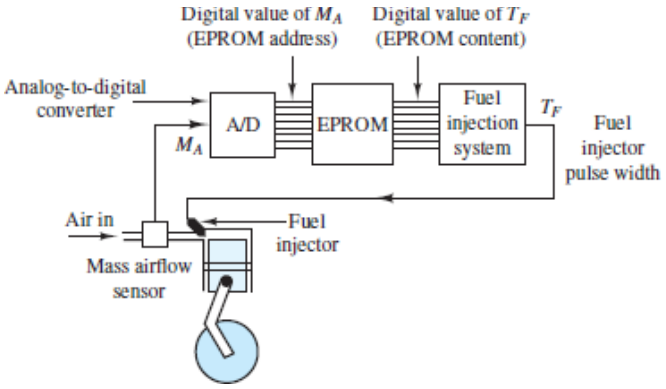


Figure 11.66 Use of EPROM lookup table in automotive fuel-injection system

To provide a numerical illustration, consider a hypothetical engine capable of aspirating air in the range $0 < M_A < 0.51$ g and equipped with fuel injectors capable of injecting at the rate of 1.36 g/s. Thus, the relationship between T_F and M_A is

$$T_F = 50 \times M_A, \quad \text{ms}$$

If the digital value of M_A is expressed in decigrams (dg, or tenths of a gram), the lookup table of Figure 11.67 can be implemented, illustrating the conversion capabilities provided by the EPROM. Note that to represent the quantities of interest in an appropriate binary format compatible with the 8-bit EPROM, the units of air mass and of time have been scaled.

M_A (g) $\times 10^1$	Address (digital value of M_A)	Content (digital value of T_F)	T_F (ms)
0	00000000	00000000	0
1	00000001	00000101	5
2	00000010	00001010	10
3	00000011	00001111	15
4	00000100	00010100	20
5	00000101	00011001	25
⋮	⋮	⋮	⋮
51	00110011	11111111	255

Figure 11.67 Lookup table for automotive fuel-injection application

CHECK YOUR UNDERSTANDING

Which combination of the control lines will select the data line D_3 for a 4-to-1 MUX?

Show that an 8-to-1 MUX with eight data inputs (D_0 through D_7) and three control lines (I_0 through I_2) can be used as a data selector. Which combination of the control lines will select the data line D_5?

Page 691

Which combination of the control lines will select the data line D_4 for an 8-to-1 MUX?

Answer: To select D_3 use $I_1I_0 = 11$; to select D_5 use $I_2I_1I_0 = 101$; to select D_4 use $I_2I_1I_0 = 100$

CHECK YOUR UNDERSTANDING

How many address inputs do you need if the number of words in a memory array is 16?

Answer: Four

Conclusion

This chapter contains an overview of digital logic circuits. These circuits form the basis of all digital computers, and of most electronic devices used in industrial and consumer applications. Upon completing this chapter, a student will have learned to:

1. *Apply concepts of analog and digital signals and of quantization.*
2. *Convert between decimal and binary number systems and use the hexadecimal system and BCD and Gray codes.* The binary and hexadecimal systems form the basis of numerical computing.
3. *Write truth tables, and realize logic functions from truth tables using logic gates.* Boolean algebra permits the analysis of digital circuits through a relatively simple set of rules. Digital logic gates are the means through which one can implement logic functions; truth tables permit the easy visualization of logic functions and can aid in the realization of these functions by using logic gates.

4. *Systematically design logic functions using Karnaugh maps.* The design of logic circuits can be systematically approached by using an extension of truth tables called the Karnaugh map. Karnaugh maps facilitate the simplification of logic expressions and their realization through logic gates in either sum-of-products or product-of-sums form.
5. *Apply various combinational logic modules, including multiplexers, memory and decoder elements, and programmable logic arrays.* Practical digital logic circuits rarely consist of individual logic gates; gates are usually integrated into combinational logic modules that include memory elements and gate arrays.

HOMEWORK PROBLEMS

Section 11.2: The Binary Number System

11.1 Convert the following base-10 numbers to hexadecimal and binary:

a. 303 b. 275 c. 18 d. 43 e. 87

11.2 Convert the following hexadecimal numbers to base-10 and binary:

a. *C* b. 44 c 28 d. 59 e. 14

11.3 Convert the following base-10 numbers to binary:

a. 231.45 b. 58.78 c. 21.22 e. 93.375

Page 692

11.4 Convert the following binary numbers to hexadecimal and base 10:

a. 1101
b. 1000100
c. 1111100
d. 1110000
e. 10000
f. 101010

11.5 Perform the following additions, all in the binary system:

a. 10101111 + 10100
b. 111100001 + 111000
c. 111001011 + 111001

11.6 Perform the following subtractions, all in the binary system:

a. 11010001 − 11100

b. 11111100 − 101010

c. 100110110 − 1001100

11.7 Assuming that the most significant bit is the sign bit, find the decimal value of the following sign-magnitude form 8-bit binary numbers:

a. 10100111 b. 01010110 c. 11111100

11.8 Find the sign-magnitude form binary representation of the following decimal numbers:

a. 122 b. −110 c. −87 d. 40

11.9 Find the twos complement of the following binary numbers:

a. 1110 b. 1100101 c. 1110000 d. 11100

11.10 Using 10 fingers, including thumbs:

a. How high can one count in a binary (base 2) number system?

b. How high can one count in base 6, using one hand to count units and the other hand for carries?

Section 11.3: Boolean Algebra and Logic Gates

11.11 Use a truth table to show that

$$\overline{A} + AB = \overline{A} + B.$$

11.12 Realize the logic function:

$$Y = (A + \overline{B}) \cdot (\overline{C} \cdot D) + A$$

using logic gates and construct its truth table.

11.13 Using the method of proof by perfect induction, show that

$$(X + Y) \cdot (\overline{X} + X \cdot Y) = Y$$

11.14 Simplify the expression

$$Y = \overline{A} \cdot \overline{B} \cdot C + \overline{A} \cdot B \cdot C + \overline{A} \cdot \overline{C}$$

using boolean algebra, and then draw the logic circuit using logic gates.

11.15 Simplify the expression

$$Y = A \cdot \overline{B} \cdot \overline{C} + A \cdot \overline{B} \cdot C + \overline{A} \cdot B \cdot C + A \cdot B \cdot C$$

using the boolean algebra.

11.16 Simplify the expression

$$Y = \overline{A} \cdot \overline{B} \cdot \overline{C} + \overline{A} \cdot \overline{B} \cdot C + \overline{A} \cdot B \cdot \overline{C}$$

using the boolean algebra.

11.17 Find a logic function equivalent to the truth table given in Figure P11.17.

A	*B*	*C*	*F*
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

Figure P11.17

11.18 Determine the boolean function describing the operation of the circuit shown in Figure P11.18 and simplify it using boolean algebra.

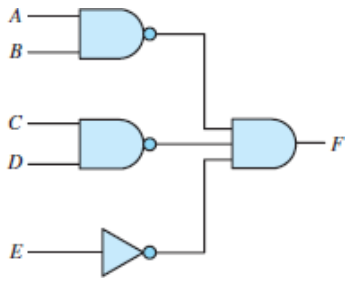


Figure P11.18

11.19 Use a truth table to show when the output of the circuit of Figure P11.19 is 1.

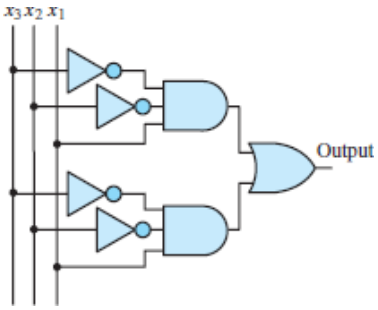


Figure P11.19

Page 693

11.20 Baseball is a complicated game, and often the manager has a difficult time keeping track of all the rules of thumb that guide decisions. To assist your favorite baseball team, you have been asked to design a logic circuit that will flash a light when the manager should give the steal sign. The rules have been laid out for you by a baseball fan with limited knowledge of the game as follows: Give the steal sign if there is a runner on first base and

a. There are no other runners, the pitcher is right-handed, and the runner is fast; or

b. There is one other runner on third base, and one of the runners is fast; or

c. There is one other runner on second base, the pitcher is left-handed, and both runners are fast.

Under no circumstances should the steal sign be given if all three bases have runners. Design a logic circuit that implements these rules to indicate when the steal sign should be given.

11.21 A small county board is composed of three commissioners. Each commissioner votes on measures presented to the board by pressing a button indicating whether the commissioner votes for or against a measure. If two or more commissioners vote for a measure, it passes. Design a logic circuit that takes the three votes as inputs and lights either a green or a red light to indicate whether a measure passed.

11.22 A water purification plant uses one tank for chemical sterilization and a second, larger tank for settling and aeration. Each tank is equipped with two sensors that measure the height of water in each tank and the flow rate of water into each tank. When the height of water or the flow rate is too high, the sensors produce a logic high output. Design a logic circuit that sounds an alarm whenever the height of water in both tanks is too high and either of the

flow rates is too high, or whenever both flow rates are too high and the height of water in either tank is also too high.

11.23 Many automobiles incorporate logic circuits to alert the driver to problems or potential problems. In one particular car, a buzzer is sounded whenever the ignition key is turned and either a door is open or a seat belt is not fastened. The buzzer also sounds when the key is not turned but the lights are on. In addition, the car will not start unless the key is in the ignition, the car is in park, and all doors are closed and seat belts fastened. Design a logic circuit that takes all the inputs listed and sounds the buzzer and starts the car when appropriate.

11.24 An on/off start-up signal governs the compressor motor of a large commercial air conditioning unit. In general, the start-up signal should be on whenever the output of a temperature sensor S exceeds a reference temperature. However, you are asked to limit the compressor start-ups to certain hours of the day and also enable service technicians to start up or shut down the compressor through a manual override. A time-of-day indicator D is available with on/off outputs, as is a manual override switch M. A separate timer T prohibits a compressor start-up within 10 min of a previous shutdown. Design a logic diagram that incorporates the state of all four devices (S, D, M, and T) and produces the correct on/off condition for the motor start-up.

11.25 NAND gates require one less transistor than AND gates. They are often used exclusively to construct logic circuits. One such logic circuit that uses three-input NAND gates is shown in Figure P11.25.

a. Determine the truth table for this circuit.

b. Find a logic function that represents the circuit.

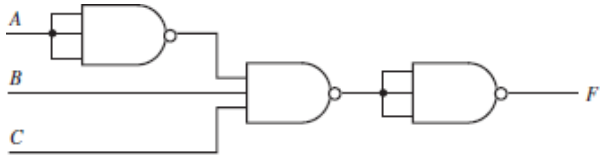


Figure P11.25

11.26 Draw a logic circuit that is equivalent to the function:

$$F = (A + \overline{B}) \cdot \overline{(C + \overline{A})} \cdot B$$

11.27 The circuit shown in Figure P11.27 is called a half adder for two single bit inputs, giving a two-bit sum as outputs. Build a truth table and verify that it indeed acts as a summer.

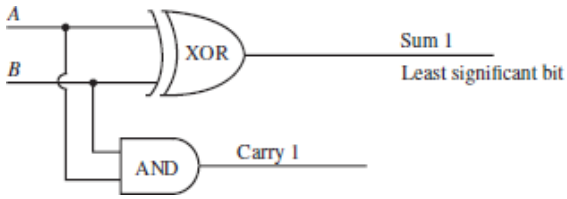


Figure P11.27

11.28 Draw a logic circuit that is equivalent to the function:

$$F = [(A + C \cdot \overline{B}) + A \cdot \overline{B} \cdot \overline{C}] \cdot \overline{(B + C)}$$

Page 694

11.29 Determine the truth table (F given A, B, C, and D) and a logical expression for the circuit of Figure P11.29.

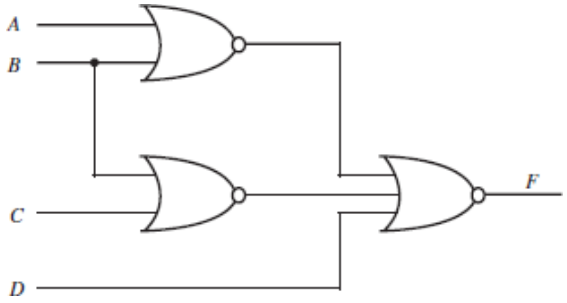


Figure P11.29

11.30 Determine the truth table (F given A, B, and C) and a logical expression for the circuit of Figure P11.30.

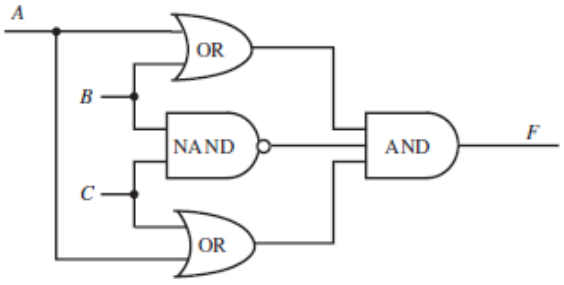


Figure P11.30

11.31 A "vote taker" logic circuit forces its output to agree with a majority of its inputs. Such a circuit is shown in Figure P11.31 for the three voters, A, B and

C. Write a logic expression for the output of this circuit in terms of its inputs. Also create the truth table for the output in terms of the inputs.

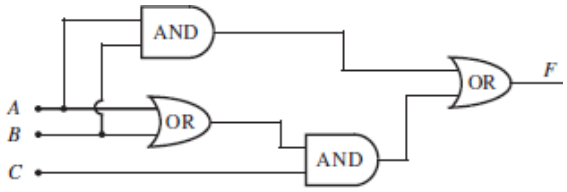


Figure P11.31

11.32 A "consensus indicator" logic circuit is shown in Figure P11.32. Write a logical expression for the output of this circuit in terms of its input. Also create the truth table for the output in terms of the inputs.

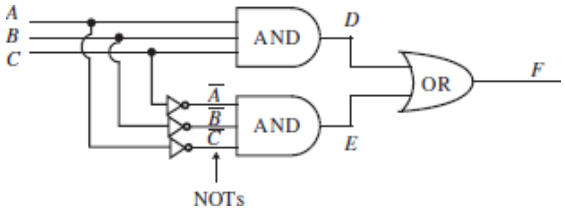


Figure P11.32

11.33 A half-adder circuit is shown in Figure P11.33. Write a logical expression for the outputs of this circuit in terms of its inputs. Also create the truth table for the outputs in terms of the inputs.

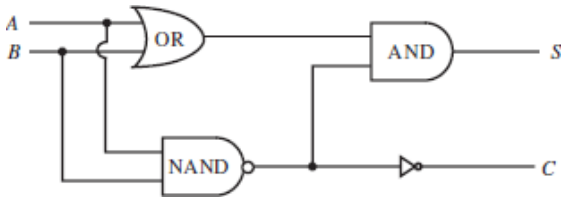


Figure P11.33

11.34 For the logic circuit shown in Figure P11.34, write a logical expression for the outputs of this circuit in terms of its inputs, and create the truth table for the outputs in terms of the inputs, including any required intermediate variables.

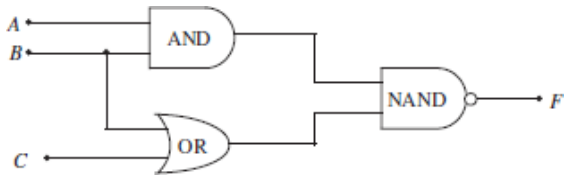


Figure P11.34

11.35 For the logic circuit in Figure P11.35, write a logical expression for the outputs of this circuit in terms of its inputs, and create the truth table for the Page 695outputs in terms of the inputs, including any required intermediate variables.

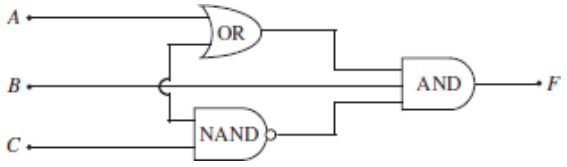


Figure P11.35

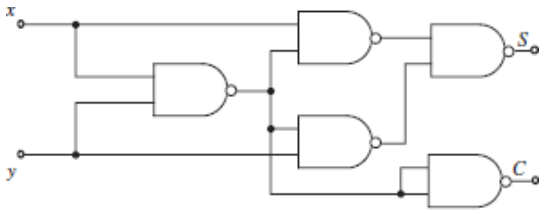
11.36 Simplify the following logic function:

$$f(A, B, C) = (A + B) \cdot A \cdot B + \overline{A} \cdot C + A \cdot \overline{B} \cdot C + \overline{B} \cdot \overline{C}$$

11.37 Complete the truth table for the circuit of Figure P11.37.

a. What mathematical function does this circuit perform, and what do the outputs signify?

b. How many standard 14-pin ICs would it take to construct this circuit?



x	y	C	S
0	0		
0	1		
1	0		
1	1		

Figure P11.37

Section 11.4: Karnaugh Maps and Logic Design

11.38 Find the logic function corresponding to the truth table of Figure P11.38 in the simplest SOP form.

A	B	C	F
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

Figure P11.38

11.39 Find the minimum expression for the output of the logic circuit shown in Figure P11.39.

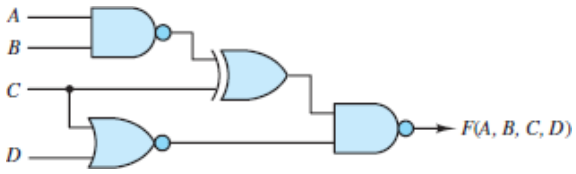


Figure P11.39

11.40 Build the Karnaugh map of the function $Y = \overline{A \cdot B \cdot C}$ and verify it using boolean algebra.

11.41 Use a Karnaugh map to minimize the function $Y = \overline{C} \cdot \overline{B} \cdot A + \overline{C} \cdot B \cdot \overline{A} + \overline{C} \cdot \overline{B} \cdot \overline{A}$

11.42 Fill in the Karnaugh map for the function $Y = f(A,B,C)$ defined by the truth table of Figure P11.42, and find the minimum expression for the function.

A	B	C	Y
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

Figure P11.42

11.43 A function F is defined such that it equals 1 when a 4-bit input code is equivalent to any of the decimal numbers 3, 6, 9, 12, or 15. Function F is 0 for input codes 0, 2, 8, and 10. Other input values cannot occur. Use a

Karnaugh map to determine a minimal expression for this function. Design and sketch a circuit to implement this function, using only AND and NOT gates.

11.44 Design the circuit of the function $Y = f(A,B,C)$ described in Figure P11.44.

A	B	C	Y
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	x

Figure P11.44

11.45 Design a logic circuit that will produce the ones complement of an 8-bit signed binary number.

Page 696

11.46 Construct the Karnaugh map for the logic function defined by the truth table of Figure P11.46, and find the minimum expression for the function.

A	B	C	D	F
0	0	0	0	1
0	0	0	1	0
0	0	1	0	1
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	1
1	0	0	0	1
1	0	0	1	0
1	0	1	0	1
1	0	1	1	0
1	1	0	0	1
1	1	0	1	1
1	1	1	0	1
1	1	1	1	0

Figure P11.46

11.47 Use a Karnaugh map to minimize the function $Y = (A + \overline{B}) \cdot [(\overline{C} \cdot D) + \overline{A}]$

11.48 Find the minimum output expression for the circuit of Figure P11.48.

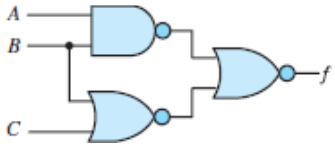


Figure P11.48

11.49 Design a combinational logic circuit that will add two 4-bit binary numbers.

11.50 Minimize the expression described in the truth table of Figure P11.50, and draw the circuit.

A	B	C	F
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

Figure P11.50

11.51 Find the minimum expression for the output of the logic circuit of Figure P11.51.

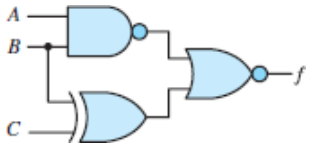


Figure P11.51

11.52 The objective of this problem is to design a combinational logic circuit that will aid in determination of the acceptability of emergency blood transfusions. It is known that human blood can be categorized into four types: A, B, AB, and O. Persons with type A blood can donate to both A and AB types and can receive blood from both A and O types. Persons with type B blood can donate to both B and AB types and can receive from both B and O types. Persons with type AB blood can donate only to type AB but can receive from any type. Persons with type O blood can donate to any type but can receive only from type O. Make appropriate variable assignments, and design a circuit that will approve or disapprove any particular transfusion based on these conditions.

11.53 Find the minimum expression for the logic function at the output of the logic circuit of Figure P11.53.

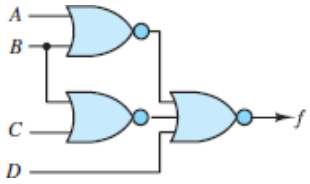


Figure P11.53

11.54 Determine the minimum boolean logic expression associated with the Karnaugh map in Figure P11.54 and create (realize) the logic circuit.

C·D \ A·B	00	01	11	10
00	1	0	0	1
01	1	1	1	1
11	0	0	1	0
10	0	1	0	0

Figure P11.54

11.55 a. Construct a Karnaugh map associated with the truth table of Figure P11.55.

b. What is the minimum expression for the function?

Page 697

c. Draw the logic circuit, using AND, OR, and NOT gates.

A	*B*	*C*	*f*(*A*, *B*, *C*)
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

Figure P11.55

11.56 Fill in the Karnaugh map for the logic function defined by the truth table of Figure P11.56. What is the minimum expression for the function?

A	B	C	D	F
0	0	0	0	1
0	0	0	1	0
0	0	1	0	1
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	1
1	0	0	0	1
1	0	0	1	0
1	0	1	0	1
1	0	1	1	0
1	1	0	0	1
1	1	0	1	1
1	1	1	0	1
1	1	1	1	0

Figure P11.56

11.57 Fill in the Karnaugh map for the logic function defined by the truth table of Figure P11.57. What is the minimum expression for the function? Realize the function using only NAND gates.

A	B	C	D	F
0	0	0	0	1
0	0	0	1	1
0	0	1	0	1
0	0	1	1	1
0	1	0	0	0
0	1	0	1	1
0	1	1	0	0
0	1	1	1	1
1	0	0	0	1
1	0	0	1	1
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	1
1	1	1	0	1
1	1	1	1	0

Figure P11.57

11.58 Design a circuit with a 4-bit input representing the binary number $A_3A_2A_1A_0$. The output should be 1 if the input value is divisible by 3. Assume that the circuit is to be used only for the digits 0 through 9 (thus, values for 10 to 15 can be don't-care conditions).

a. Draw the Karnaugh map and truth table for the function.

b. Determine the minimum expression for the function.

c. Draw the circuit, using only AND, OR, and NOT gates.

11.59 Find the simplest SOP representation of the function associated with the Karnaugh map shown in Figure P11.59.

C·D \ A·B	00	01	11	10
00	0	1	0	0
01	1	1	0	0
11	0	*x*	1	0
10	0	0	1	0

Figure P11.59

11.60 Can the circuit for Problem 11.54 be simplified if it is known that the input represents a BCD (binary-coded decimal) number, that is, if it can never be greater than 9_{10}? If not, explain why not. Otherwise, design the simplified circuit.

11.61 Find the simplest SOP representation of the function associated with the Karnaugh map shown in Figure P11.61.

C·D \ A·B	00	01	11	10
00	0	1	*x*	0
01	0	1	*x*	0
11	0	1	0	1
10	*x*	*x*	1	0

Figure P11.61

Page 698

11.62 One method of ensuring reliability in data transmission systems is to transmit a parity bit along with every nibble, byte, or word of binary data transmitted. The parity bit confirms whether an even or odd number of 1s were transmitted in the data. In even-parity systems, the parity bit is set to 1 when the number of 1s in the transmitted data is odd. Odd-parity systems set the parity bit to 1 when the number of 1s in the transmitted data is even. Assume that a parity bit is transmitted for every nibble of data. Design a logic circuit

that checks the nibble of data and transmits the proper parity bit for both even- and odd-parity systems.

11.63 Assume that a parity bit is transmitted for every nibble of data. Design two logic circuits that check a nibble of data and its parity bit to determine if there may have been a data transmission error. Assume first an even-parity system, then an odd-parity system.

11.64 Design a logic circuit that takes a 4-bit Gray code input from an optical encoder and translates it into two 4-bit nibbles of BCD.

11.65 Design a logic circuit that takes a 4-bit Gray code input from an optical encoder and determines if the input value is a multiple of 3.

11.66 The 4221 code is a base 10–oriented code that assigns the weights 4221 to each of 4 bits in a nibble of data. Design a logic circuit that takes a BCD nibble as input and converts it to its 4221 equivalent. The logic circuit should also report an error in the BCD input if its value exceeds 1001.

11.67 The 4-bit digital output of each of two sensors along an assembly line conveyor belt is proportional to the number of parts that pass by on the conveyor belt in a 30-s period. Design a logic circuit that reports an error if the outputs of the two sensors differ by more than one part per 30-s period.

Section 11.5: Combinational Logic Modules

11.68 A function F is defined such that it equals 1 when a 4-bit input code is equivalent to any of the decimal numbers 3, 6, 9, 12, or 15. F is 0 for input codes 0, 2, 8, and 10. Other input values cannot occur. Use a Karnaugh map to determine a minimal expression for this function. Design and sketch a circuit to implement this function using only AND and NOT gates.

11.69 Fill in the Karnaugh map for the logic function defined by the truth table of Figure P11.69. What is the minimum expression for the function? Realize the function using a 1-of-8 multiplexer.

A	B	C	D	f(A, B, C, D)
0	0	0	0	1
0	0	0	1	0
0	0	1	0	1
0	0	1	1	1
0	1	0	0	0
0	1	0	1	1
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	1
1	0	1	0	0
1	0	1	1	0
1	1	0	0	1
1	1	0	1	0
1	1	1	0	1
1	1	1	1	1

Figure P11.69

11.70 Fill in the truth table for the multiplexer circuit shown in Figure P11.70. What binary function is performed by these multiplexers?

x	y	C	S
0	0		
0	1		
1	0		
1	1		

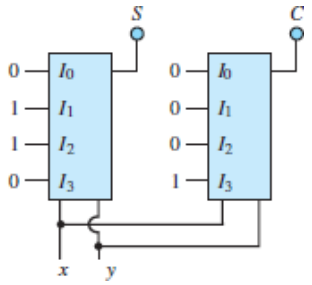


Figure P11.70

11.71 The circuit of Figure P11.71 operates as a 4:16 decoder. Terminal EN denotes the enable input. Describe its operation. What is the role of logic variable *A*?

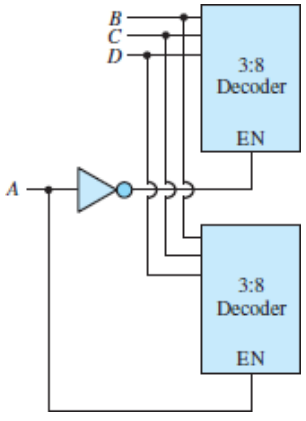


Figure P11.71

11.72 Show that the circuit given in Figure P11.72 converts 4-bit binary numbers to 4-bit Gray code.

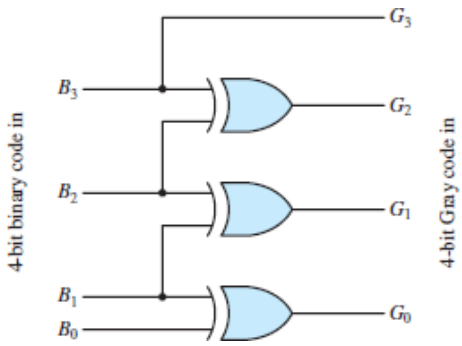


Figure P11.72

11.73 Suppose one of your classmates claims that the following boolean expressions represent the conversion from 4-bit Gray code to 4-bit binary numbers:

$$B_3 = G_3$$
$$B_2 = G_3 \oplus G_2$$
$$B_1 = G_3 \oplus G_2 \oplus G_1$$
$$B_0 = G_3 \oplus G_2 \oplus G_1 \oplus G_0$$

a. Show that your classmate's claim is correct.

b. Draw the circuit that implements the conversion.

11.74 Select the proper inputs for a four-input multiplexer to implement the function $f(A, B, C) = \overline{A}B\overline{C} + A\overline{B}\overline{C} + AC$. Assume inputs I_0, I_1, I_2, and I_3 correspond to $\overline{A}\overline{B}, \overline{A}B, A\overline{B}$, and AB, respectively, and that each input may be $0, 1, \overline{C}$, or C.

11.75 Select the proper inputs for an 8-bit multiplexer to implement the function $f(A, B, C, D) = \sum(2, 5, 6, 8, 9, 10, 11, 13, 14)_{10}$. Assume the inputs I_0 through I_7 correspond to $\overline{A}\overline{B}\overline{C}, A\overline{B}\overline{C}, \overline{A}B\overline{C}, AB\overline{C}, \overline{A}\overline{B}C, A\overline{B}C, \overline{A}BC$, and ABC, respectively, and that each input may be $0, 1, \overline{D}$, or D.

11.76 Use a 3:8 decoder and a three-input OR gate to implement the logic function $f(x, y, z) = xy + x\overline{y} + \overline{x}\overline{y}z$. Draw a logic diagram and create the associated truth table.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for “Focus on measurements” weighing scales: Media Bakery.

[1]American Standard Code for Information Interchange.

[2]A useful rule to remember is that in a two-variable map, there are two minterms adjacent to any given minterm; in a three-variable map, three minterms are adjacent to any given minterm; in a four-variable map, the number is four, and so on.

[3]The size of a word depends upon the particular system architecture.

C H A P T E R

12

DIGITAL SYSTEMS

In general, a digital system is able to perform computations on digital signals and data and store the results in memory. Some of these computations are performed by *combinational logic gates* (Chapter 11), which require no knowledge of prior *logic states*. However, other computations do require such knowledge. *Sequential logic gates*, which are built up from combinational logic gates, employ *feedback* from outputs to inputs to generate output logic states that depend upon the output logic states at earlier times. In effect, these sequential logic gates have *memory*. The first part of this chapter is focused on sequential logic gates and rudimentary devices, such as flip-flops, counters, and registers that are built up from them.

The second part of the chapter builds upon the first part by describing basic computer system architecture, including registers, which are the most fundamental units of digital memory. A discussion of *microcontrollers*, and details of the Atmel ATmega328P® microcontroller, in particular, follows the section on architecture.[1] The popular, open-source, microcontroller-based *Arduino* hardware and software project is also described in sufficient detail to motivate applications and student projects, such as motor control and data acquisition.

Learning Objectives

Students will learn to...

1. Analyze the operation of flip-flops and latches. *Section 12.1*.
2. Analyze and apply digital counters and registers. *Section 12.2*.
3. Design simple sequential circuits using state transition diagrams. *Section 12.3*.
4. Describe the basic architecture of computers. *Section 12.4*.
5. Identify the architecture of microprocessors, microcontrollers, and the ATmega328P®, in particular. *Section 12.5*.
6. Utilize the Arduino hardware and software project. *Section 12.6*.

12.1 LATCHES AND FLIP-FLOPS

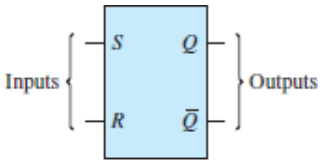
A **flip-flop** is an elementary **sequential logic** gate. Various types of flip-flops exist; however, all flip-flops share the following characteristics:

1. A flip-flop is a **bistable device**; that is, it can remain in one of two stable states (0 and 1) until appropriate conditions cause it to change state. Thus, a flip-flop can serve as a **memory element**.
2. A flip-flop has two outputs, one of which is the complement of the other.

RS Flip-Flop

It is customary to depict flip-flops by their block diagram and their output by a name, such as Q. Figure 12.1 represents the ***RS* flip-flop**, which has two inputs, denoted by S and R, and two outputs Q and $\overline{Q}$. The value at Q is called the binary output *state* of the flip-flop. The two inputs R and S are used to change the state of the flip-flop, according to the following rules:

1. When $R = S = 0$, Q remains unchanged from its present state.
2. When $S = 1$ and $R = 0$, the output is *set* such that $Q = 1$.
3. When $S = 0$ and $R = 1$, the output is *reset* such that $Q = 0$.
4. S and R are not permitted to be 1 simultaneously.



S	R	Q
0	0	Present state
0	1	Reset
1	0	Set
1	1	Disallowed

Figure 12.1 *RS* flip-flop symbol and truth table

A **timing diagram** is a convenient means of describing the transitions that occur in the output of a flip-flop due to changes in its inputs. Figure 12.2 depicts a table of transitions for an *RS* flip-flop Q as well as the corresponding timing diagram.

S	R	Q
1	0	1
0	0	1
0	0	1
0	1	0
0	0	0
0	0	0
0	1	0
0	0	0
1	0	1
0	0	1

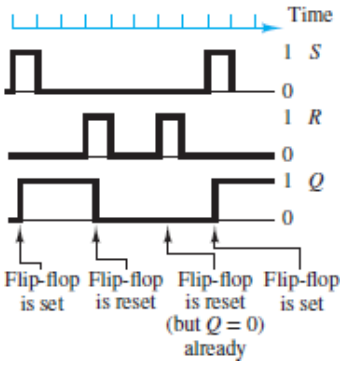


Figure 12.2 Timing diagram for the *RS* flip-flop

It is important to note that the *RS* flip-flop is **level-sensitive**. This means that the set and reset operations are completed only after the R and S inputs have reached the appropriate levels. Thus, in Figure 12.2 the transitions in the Q output occur with a small delay relative to the transitions in the R and S inputs.

Figure 12.3 illustrates how an *RS* flip-flop could be constructed from two inverters and two NAND gates. Consider the case when $S = R = 0$ such that $\bar{S} = \bar{R} = 1$. Then the result of each NAND gate is determined entirely by $\bar{Q}$ and Q.

That is, when one input to a NAND gate is set high to 1, the output of that NAND gate is the inversion of the other input (refer to the NAND gate truth table in Page 703Chapter 11). Thus, when $S = R = 0$, the outputs of the two NAND gates in Figure 12.3 are simply $\bar{\bar{Q}} = Q$ and $\bar{Q}$. In other words, the output states of the *RS* flip-flop remain unchanged from their prior states whenever *S* and *R* are both set low to 0.

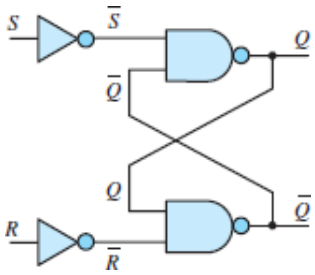


Figure 12.3 NAND gate implementation of the *RS* flip-flop

When *S* is set high to 1, the output of the upper NAND gate *Q* is also set high to 1. Why? Because when *S* is set high to 1, $\bar{S}$ is set low to 0, and when one input to a NAND gate is low, the output of the NAND gate is high *regardless* of the state of the other input. Likewise, when *R* is set high to 1, $\bar{Q}$ is set high to 1.

The only difficulty with the *RS* flip-flop occurs when both *S* and *R* are set high to 1. Clearly, it is an inherent contradiction to suppose that both *Q* and $\bar{Q}$ are both set high to 1 at any point in time. Why? Because $\bar{Q}$ is, by definition, the inversion of *Q*. Thus, $S = R = 1$ is not allowed. The *RS* flip-flop cannot be both set and reset at the same time. In practice, one could set $S = R = 1$, but the output will be unstable.

As is true for any logic network, it is possible to find alternate formulations of the *RS* flip-flop. One of DeMorgan's laws states that a NAND gate is equivalent to an OR gate with inverted inputs. Make this replacement in Figure 12.3, and note that all of the inputs are now inverted prior to the OR gates. When inverters are added to the outputs of the OR gates the result is NOR gates with *Q* and $\bar{Q}$ interchanged. The result is shown in Figure 12.4.

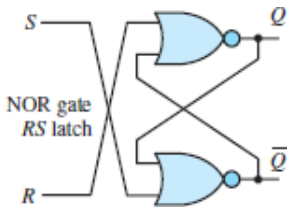


Figure 12.4 NOR gate implementation of the *RS* flip-flop

Figure 12.5 shows the same two-NOR-gate implementation of the *RS* flip-flop, but with an enable input *E* connected to two AND gates such that the *R* or *S* inputs will be effective only when $E = 1$. The enable input is often a **clock** signal used to synchronize other inputs.

Figure 12.5 also illustrates two additional features: the **preset** *P* and **clear** *C* functions. These features have no effect when set low to 0. However, when *P* is set high to 1, the output of the upper NOR gate $\bar{Q}$ is set low to 0 and, thus, *Q* is set high to 1. $P = 1$ always results in $Q = 1$. Likewise, when *C* is set high to 1, the output of the lower NOR gate *Q* is set low to 0. The preset and clear are not controlled by the enable input and, thus, are said to be **asynchronous**. It is important to realize that $P = C = 1$ is not allowed. The timing diagram of Figure 12.5 illustrates the role of the enable, preset, and clear inputs. Notice that transitions due to *S* and *R* occur only after *E* is set high. The flip-flop can be designed so Page 704that the *P* and *C* inputs are also controlled by *E*; in fact, many commercial flip-flops are designed this way so that all inputs are synchronized with *E*.

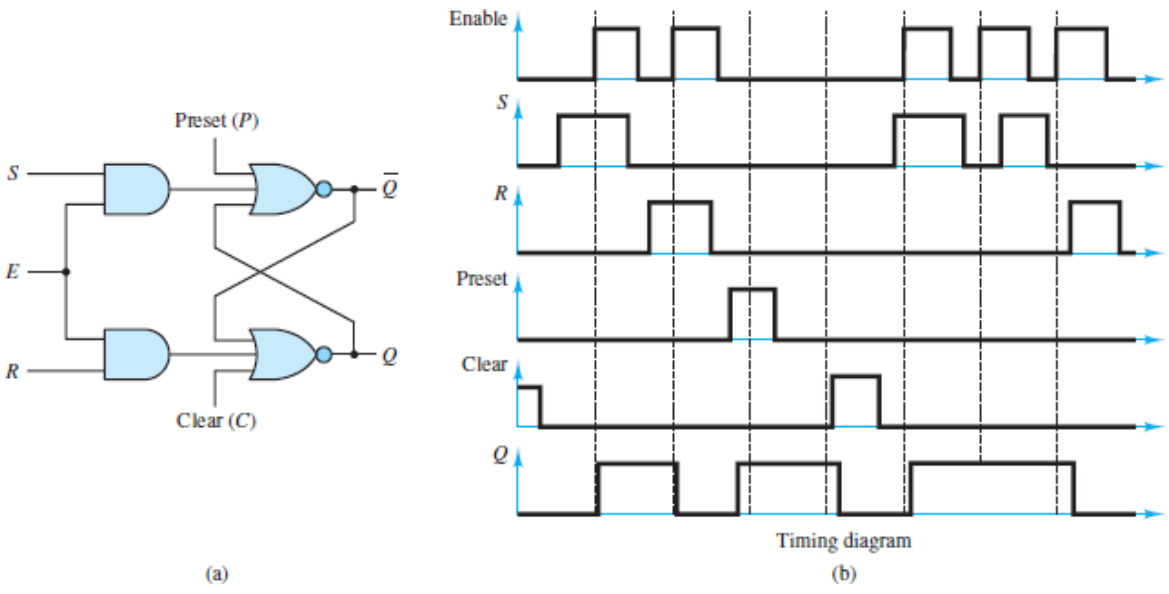


Figure 12.5 The *RS* flip-flop with enable, preset, and clear lines: (a) logic diagram, (b) example timing diagram

Another extension of the *RS* flip-flop, called a **data latch**, or **delay**, is shown in Figure 12.6. In this circuit, $R = \bar{S}$ such that when $E = 1$, $Q = D$. When *E* is set low to 0, the output *Q* does not change but retains its value until *E* is set high

again. In other words, Q is *latched* when E is set low and *unlatched* when E is set high. The timing diagram illustrates that this effect also delays the impact of D on Q until the *next* time E is set high.

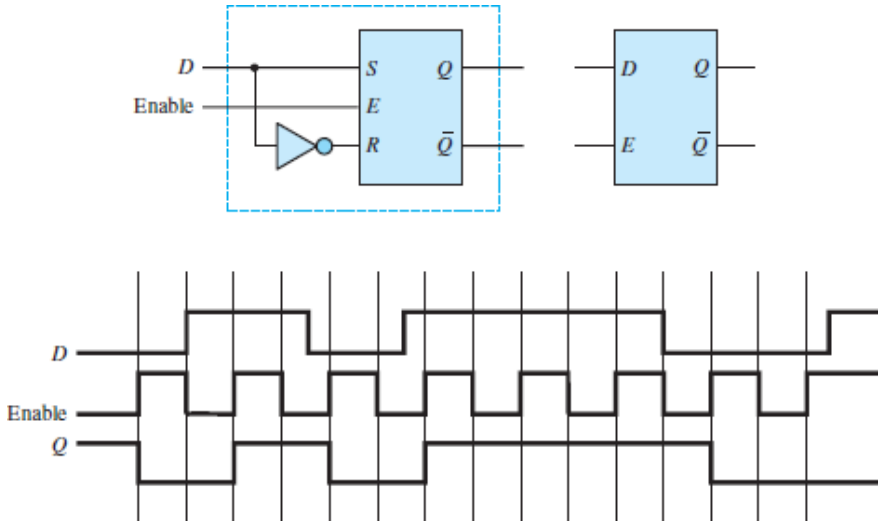


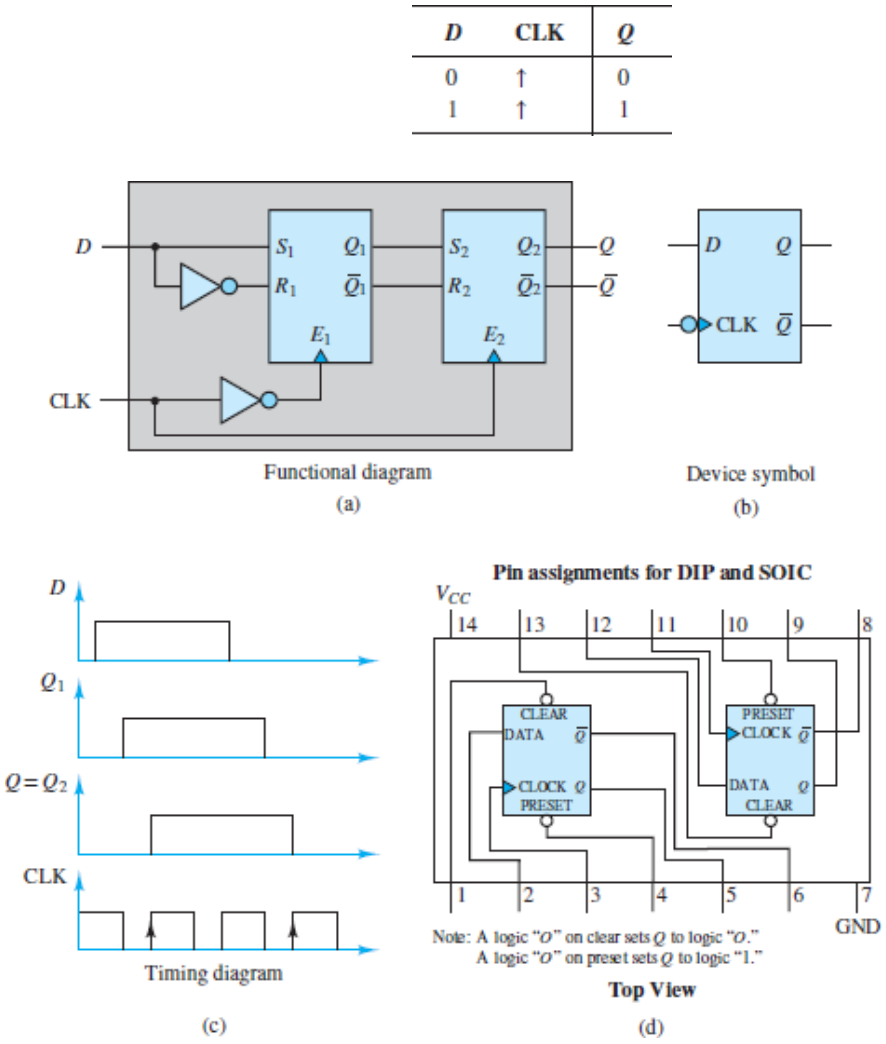
Figure 12.6 Data latch and associated timing diagram

Page 705

D Flip-Flop

The ***D* flip-flop** is an extension of the data latch that utilizes two *RS* flip-flops, as shown in Figure 12.7(a), and a clock signal to drive their enable inputs. Note that the clock is inverted prior to E_1 such that latch 1 is enabled when the clock goes low. However, latch 2 is disabled when the clock is low, such that its output will not change state until the clock subsequently goes high to enable the transfer of state from Q_1 to Q_2.

It is important to note the triangular "knife-edge" symbol shown at the CLK input in Figure 12.7(b). This symbol indicates that the *D* flip-flop changes state only on a positive clock *transition*; that is, a transition from low to high. Internally, Q_1 is set on a negative transition, whereas Q_2 (and therefore Q) is set on a positive transition, as shown in Figure 12.7(c). Thus, this particular *D* flip-flop is said to be positive edge–triggered, or **leading edge–triggered**, as indicated in the following truth table, where the symbol ↑ indicates a positive transition.



D	CLK	Q
0	↑	0
1	↑	1

Figure 12.7 The *D* flip-flop: (a) functional diagram, (b) device symbol, (c) timing waveforms, and (d) IC schematic

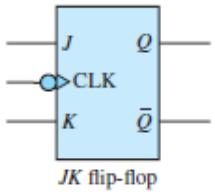
Page 706

JK Flip-Flop

The symbol and truth table of the ***JK* flip-flop** are shown in Figure 12.8. The *bubble* at the clock input signifies it is negative or **trailing edge–triggered**. Its operating rules are

- When J and K are both low, the output is unchanged.
- When $J = 0$ and $K = 1$, the output is reset to 0.
- When $J = 1$ and $K = 0$, the output is set to 1.

- When both *J* and *K* are high, the output will toggle between states at every negative transition of the clock input, denoted by the symbol ↓.

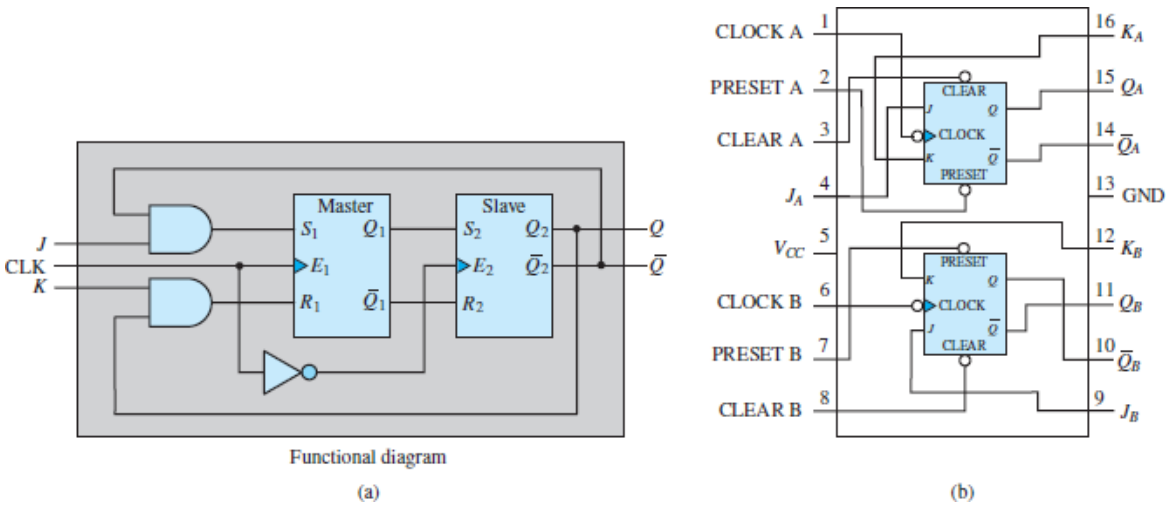


JK flip-flop

J_n	K_n	Q_{n+1}
0	0	Q_n
0	1	0 (reset)
1	0	1 (set)
1	1	$\bar{Q}_n$ (toggle)

Figure 12.8 Truth table for the *JK* flip-flop

The operation of the *JK* flip-flop can also be explained in terms of two *RS* flip-flops as shown in Figure 12.9(a). When the clock transitions from low to high, the *master* is enabled; however, the *slave* does not receive the master outputs until it is enabled during a negative clock transition. This behavior is similar to that of an *RS* flip-flop, except for the $J = 1$, $K = 1$ condition, which is allowed and results in the outputs being *toggled*.



Functional diagram (a) (b)

Figure 12.9 The *JK* flip-flop: (a) functional diagram, and (b) IC schematic of two independent *JK* flip-flops

The *JK* flip-flop is also known as the *universal flip-flop* because it can be configured to behave as an *RS* or *D* flip-flop. When both inputs are 0, the outputs remain in their previous state during a clock transition. With the assignments $J = S$ and $K = R$ (but avoiding $J = K = 1$) the *JK* flip-flop acts as an *RS* flip-flop. With the assignments $J = \overline{K} = D$, the *JK* flip-flop acts as a *D* flip-flop. Finally, when the inputs are set so that $K = J$, the outputs behave as a *T* flip-flop, which is described in Example 12.2.

EXAMPLE 12.1 *RS* Flip-Flop Timing Diagram

Problem

Determine the output of an *RS* flip-flop for the series of inputs given in the table below.

R	0	0	0	1	0	0	0
S	1	0	1	0	0	1	0

Solution

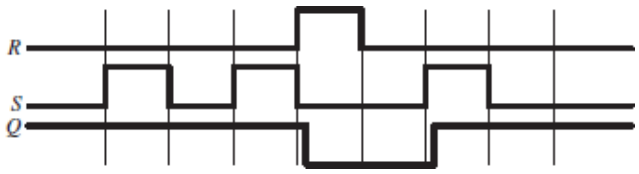
Known Quantities: *RS* flip-flop truth table (Figure 12.1).

Find: Output *Q* of *RS* flip-flop.

Analysis: The timing diagram for the *RS* flip-flop is completed, following the rules stated earlier to determine the output of the device; the result is summarized below.

R	0	0	0	1	0	0	0
S	1	0	1	0	0	1	0
Q	1	1	1	0	0	1	1

A sketch of the waveforms, shown below, can also be generated to visualize the transitions.



EXAMPLE 12.2 The *T* Flip-Flop

Problem

Determine the truth table of the ***T* flip-flop** of Figure 12.10. Note that the *T* flip-flop is a *JK* flip-flop with its inputs tied together.

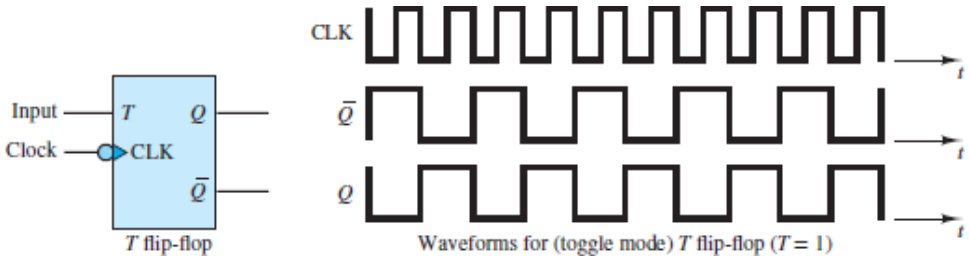


Figure 12.10 The *T* flip-flop symbol and timing waveforms

Page 708

Solution

Known Quantities: *JK* flip-flop truth table (see Figure 12.8).

Find: Truth table and timing diagram for *T* flip-flop.

Analysis: The *T* flip-flop is a *JK* flip-flop with $K = J$. Thus, the flip-flop will need only a two-element truth table to describe its operation, corresponding to the top and bottom entries in the truth table of Figure 12.8. The truth table is shown below. A timing diagram is also included in Figure 12.10.

T	CLK	Q_{k+1}
0	↓	Q_k
1	↓	$\overline{Q_k}$

Comments: The *T* flip-flop takes its name from the fact that it *toggles* between the high and low states. Note that the toggling frequency is one-half that of the clock. Thus the *T* flip-flop also acts as a *divide-by-2* counter.

EXAMPLE 12.3 The *JK* Flip-Flop Timing Diagram

Problem

Determine the output of a *JK* flip-flop for the series of inputs given in the table below. The initial state of the flip-flop is $Q_0 = 1$.

J	0	1	0	1	0	0	1
K	0	1	1	0	0	1	1

Solution

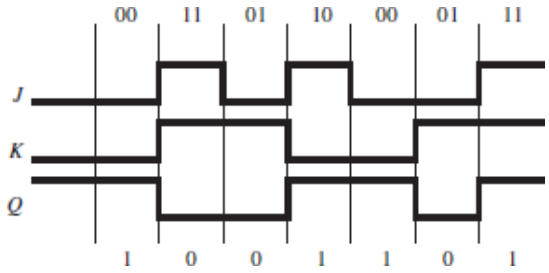
Known Quantities: *JK* flip-flop truth table (see Figure 12.8).

Find: Output of *JK* flip-flop as a function of the input transitions.

Analysis: Complete the timing diagram for the *JK* flip-flop, using the rules of Figure 12.8.

J	0	1	0	1	0	0	1
K	0	1	1	0	0	1	1
Q	1	0	0	1	1	0	1

A sketch of the waveforms, shown below, can also be generated to visualize the transitions. Each vertical line corresponds to a clock transition.



Page 709

Comments: How would the timing diagram change if the initial state of the flip-flop were $Q_0 = 0$?

CHECK YOUR UNDERSTANDING

Derive the detailed truth table and draw a timing diagram for the *JK* flip-flop, using the model of Figure 12.9 with two flip-flops. Include each unique internal input in the table and timing diagram.

12.2 DIGITAL COUNTERS AND REGISTERS

One of the more immediate applications of flip-flops is in the design of **counters**. A counter is a sequential logic device that can take one of *N* possible states, stepping through these states in a sequential fashion. When the counter has reached its last state, it resets to 0 and is ready to start counting again. For example, a 3-bit **binary up counter** would have $2^3 = 8$ possible states and might appear as shown in the functional block of Figure 12.11. The clock input steps the counter through the eight states, one transition per clock pulse. This particular counter also has a reset input, which can force the counter outputs low: $b_2b_1b_0 = 000$.

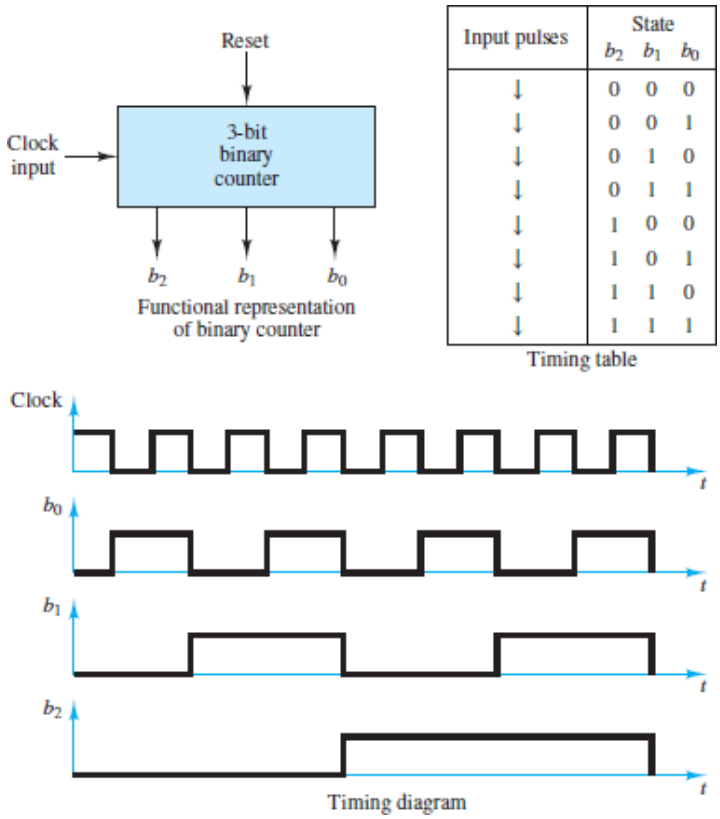


Figure 12.11 Binary up counter functional representation, state table, and timing waveforms

Page 710

Although binary counters are very useful in many applications, one is often interested in a **decade counter**, that is, a counter that counts from 0 to 9 and then resets. A 4-bit binary counter can easily be configured in principle to provide this function. As shown in Figure 12.12(b), if bits b_3 and b_0 are tied to a four-input AND gate, along with $\overline{b}_2$ and $\overline{b}_1$, the output of the AND gate will reset the counter immediately after the count reaches $1001^2 = 9_{10}$. Additional logic can provide a *carry* bit, which could be passed along to another decade counter, enabling counts up to 99. Decade counters can be cascaded to represent any series of decimal digits.

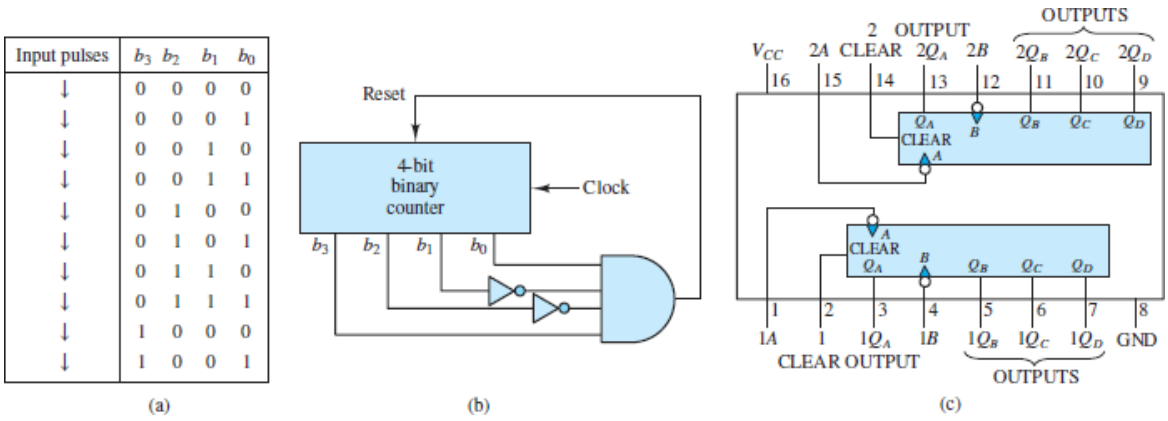
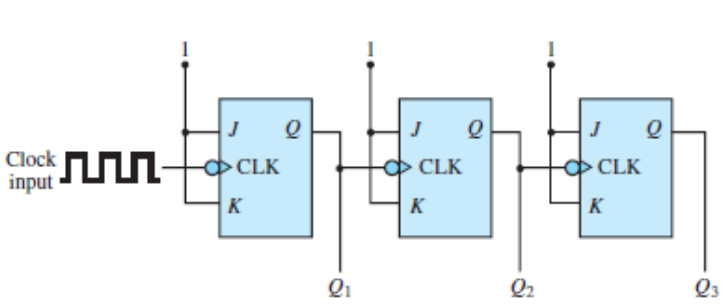


Figure 12.12 Decade counter: (a) counting sequence; (b) functional diagram; and (c) IC schematic of two independent counters

Although the decade counter of Figure 12.12 is attractive because of its simplicity, this configuration would never be used in practice, because of **propagation delays**, which are due to the finite response time of the internal transistors. In general, propagation delays are not the same for any two gates or flip-flops. Thus, if the reset signal—which is presumed to be applied at exactly the same time to each of the four *JK* flip-flops in the 4-bit binary counter—does not cause all four flip-flops to reset prior to the next clock trigger, then the binary word appearing at the output of the counter will change from 1001 to a number other than 0000, and the output of the four-input AND gate will no longer be high. Here, the CLEAR function is assumed to be active high, such that a reset occurs when the output of the AND gate goes high. This problem can be addressed with the aid of **state transition diagrams**, which are discussed in the next section.

An implementation of a 3-bit binary **ripple counter** is shown in Figure 12.13. Its transition table illustrates how the *Q* output of each stage becomes the clock input to the next stage, while each flip-flop is held in the toggle mode. The output transitions assume that the clock (CLK) is a simple square wave (with all three *JK*s negative edge–triggered).



Input	Q_3	Q_2	Q_1
↓	0	0	0
↓	0	0	1
↓	0	1	0
↓	0	1	1
↓	1	0	0
↓	1	0	1
↓	1	1	0
↓	1	1	1

Figure 12.13 Ripple counter

This 3-bit ripple counter can be used to provide a divide-by-8 counter by connecting the outputs to an AND gate, as shown in Figure 12.14. The result is one output pulse for every eight clock pulses. Note that the clock input signal is also connected to the AND gate to synchronize the output. This application of ripple counters is further illustrated in Example 12.4.

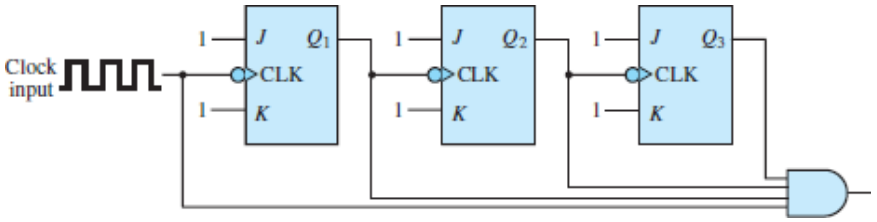


Figure 12.14 Ripple counter used to produce a divide-by-8 circuit

Page 711

A slightly more complex version of the binary counter is the **synchronous counter**, in which the input clock triggers all the flip-flops simultaneously. Figure 12.15 depicts a 3-bit synchronous counter constructed using *T* flip-flops, which are *JK* flip-flops with the *JK* inputs tied together (see Example 12.2). Q_0 toggles to 1 first and then Q_1 does the same. The AND gate ensures that Q_2 will toggle only after Q_0 and Q_1 have both reached the 1 state ($Q0 \cdot Q1 = 1$).

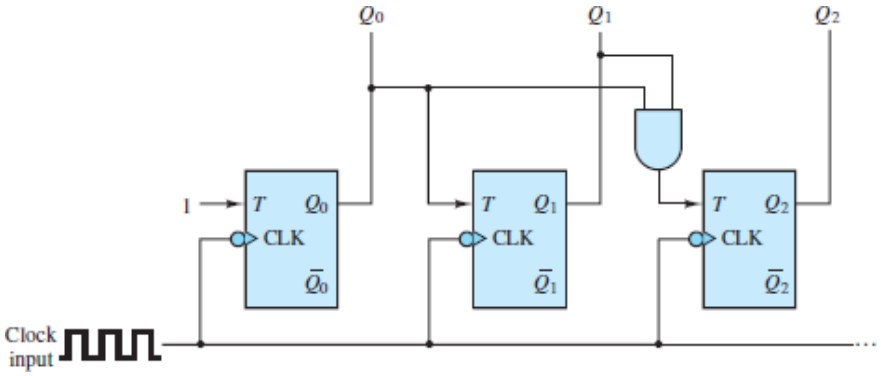


Figure 12.15 Three-bit synchronous counter

Other common counters are the **ring counter**, illustrated in Example 12.5, and the **up-down counter**, which has an additional select input that determines whether the counter counts up or down.

FOCUS ON MEASUREMENTS

Digital Measurement of Angular Position and Velocity

Another type of angular position encoder, besides the angular encoder discussed in Chapter 11, is the slotted encoder shown in Figure 12.16. This encoder can be used in conjunction with a pair of counters and a high-frequency clock to determine the speed of rotation of the slotted wheel. As shown in Figure 12.17, a clock of known frequency is connected to a counter while another counter records the number of slot pulses detected by an optical slot detector as the wheel rotates. Dividing the counter values, the speed of the rotating wheel can be obtained in radians per second. For example, assume a clocking frequency of 1.2 kHz. If both counters are started at zero and at some instant the timer counter reads 3,050 and the encoder counter reads 2,850, then the speed of the rotating encoder is found to be

$$1{,}200\ \frac{\text{cycles}}{\text{s}} \cdot \frac{2{,}850\ \text{slots}}{3{,}050\ \text{cycles}} = 1{,}121.3\ \frac{\text{slots}}{\text{s}}$$

and

$$1{,}121.3\ \frac{\text{slots}}{\text{s}} \times 1^\circ\ \text{per slot} \times \frac{2\pi}{360}\ \frac{\text{rad}}{\text{deg}} = 19.6\ \frac{\text{rad}}{\text{s}}$$

If this encoder is connected to a rotating shaft, it is possible to measure the angular position and velocity of the shaft. Such shaft encoders are used in measuring the speed of rotation of electric motors, machine tools, engines, and other rotating machinery.

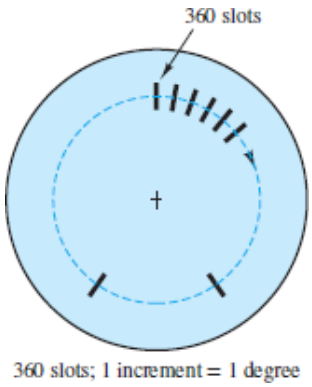


Figure 12.16

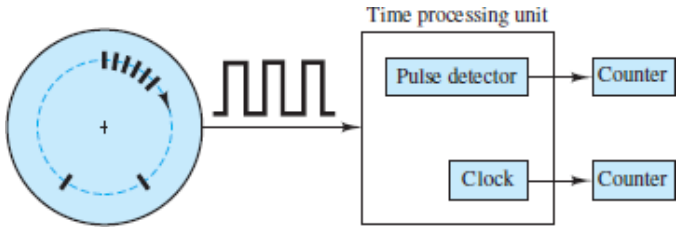


Figure 12.17 Calculating the speed of rotation of the slotted wheel

A typical application of the slotted encoder is to compute the ignition and injection timing in an automotive engine. In an automotive engine, information related to speed Page 713is obtained from the camshaft and the flywheel, which have known reference points. The reference points determine the timing for the ignition firing points and fuel-injection pulses and are identified by special slot patterns on the camshaft and crankshaft. Two methods are used to detect the special slots (reference points): *period measurement with additional transition detection (PMA)* and *period measurement with missing transition detection (PMM)*. In the PMA method, an additional slot (reference point) determines a known reference position on the crankshaft or camshaft. In the PMM method, the reference position is determined by the absence of a slot. Figure 12.18 illustrates a typical PMA pulse sequence, showing the presence of an additional pulse. The additional slot may be used to determine the timing for the ignition pulses relative to a known position of the crankshaft. Figure 12.19 depicts a typical PMM pulse sequence. Because the period of the pulses is known, the additional slot or the

missing slot can be easily detected and used as a reference position. These pulse sequences can be implemented using ring counters.

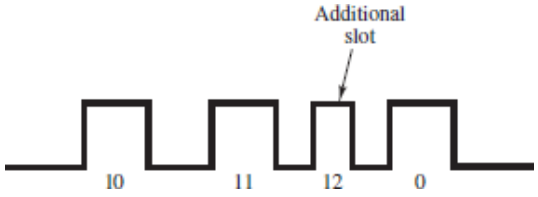


Figure 12.18 PMA pulse sequence

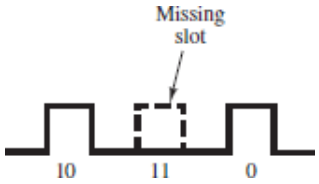


Figure 12.19 PMM pulse sequence

Registers

A register consists of a cascade of flip-flops that can store binary data, 1 bit in each flip-flop. The simplest type of register is the parallel input–parallel output register shown in Figure 12.20. In this register, the *load* input pulse, which acts on all clocks simultaneously, causes the parallel inputs $b_0b_1b_2b_3$ to be transferred to the respective flip-flops. The *D* flip-flop employed in this register allows the transfer from b_2 to Q_2 to occur very directly. Thus, *D* flip-flops are very commonly used in this type of application. The binary word $b_3b_2b_1b_0$ is now "stored," each bit being represented by the state of a flip-flop. Until the load input is applied again and a new word appears at the parallel inputs, the register will preserve the Page 714stored word. Note that a *D* flip-flop can be constructed using a *JK* flip-flop with $J = \overline{K} = D$.

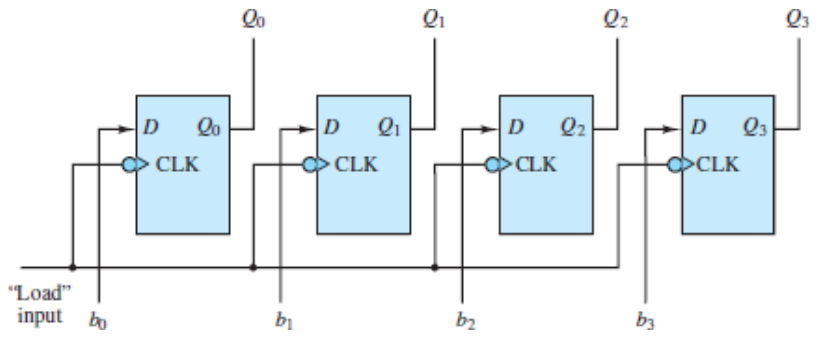


Figure 12.20 A 4-bit parallel register

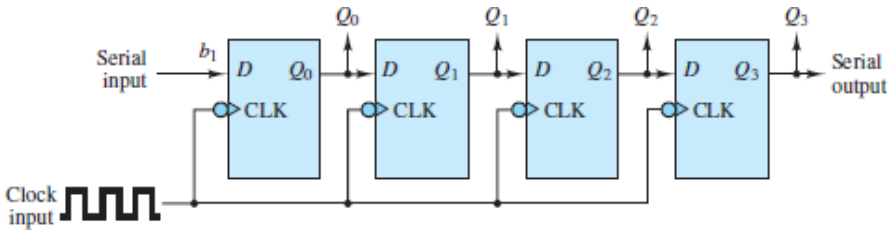


Figure 12.21 A 4-bit shift register

The construction of the parallel register presumes that the *N*-bit word to be stored is available in parallel form. However, often a binary word will arrive in serial form, that is, 1 bit at a time. A register that can accommodate this type of logic signal is called a **shift register**. Figure 12.21 illustrates how the same basic structure of the parallel register applies to the shift register. The input is now applied to the first flip-flop and the output of each flip-flop is shifted forward as the input of its succeeding flip-flop. Note that this type of register provides both a serial and a parallel output.

FOCUS ON MEASUREMENTS

Seven-Segment Display

A **seven-segment display** (Figure 12.22) is a very convenient device for displaying digital data. Operation of a seven-segment display requires a decoder circuit to light the proper combinations of segments corresponding to the desired decimal digit.

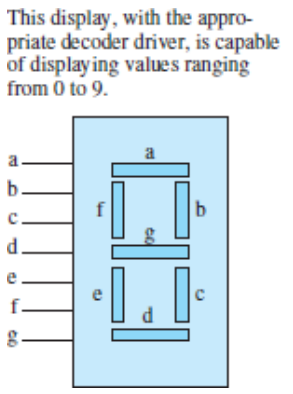


Figure 12.22 Sevensegment display

A typical *BCD* to seven-segment decoder function block is shown in Figure 12.23, where the lowercase letters correspond to the segments shown in Figure 12.22. The decoder features four data inputs (*A*, *B*, *C*, *D*), which are used to determine the state of the seven outputs. Each output is connected to one segment of the display. A *BCD* to seven-segment decoder is similar to the 2:4 decoder described in Chapter 11.

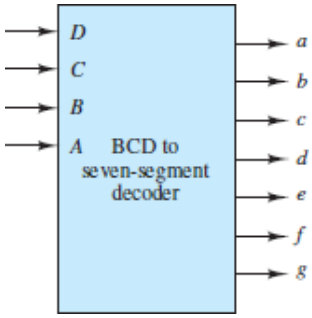


Figure 12.23

EXAMPLE 12.4 Divider Circuit

Problem

A binary ripple counter provides a means of dividing the fixed output rate of a clock by powers of 2. For example, the circuit of Figure 12.24 is a divide-by-2 or divide-by-4 counter. Draw the timing diagrams for the clock input, Q_0 and Q_1 to demonstrate these functions.

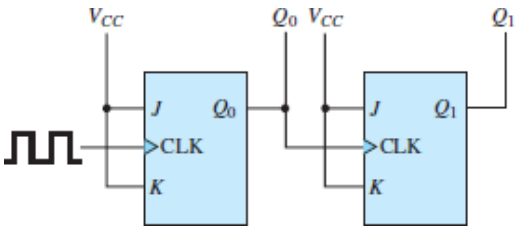


Figure 12.24

Solution

Known Quantities: *JK* flip-flop truth table (Figure 12.8).

Find: Output of each flip-flop Q as a function of the input clock transitions.

Assumptions: Assume positive edge–triggered devices. The DC supply voltage is V_{CC}. Outputs Q_0 and Q_1 begin low.

Analysis: With both inputs tied to V_{CC} (logic 1) the *JK* flip-flops act as toggle (T) flip-flops. Notice that the clock input is positive-edge triggered and that Q_0 serves as the clock input to flip-flop 1. Thus, Q_0 toggles from low to high when the clock first transitions from low to high. This positive Q_0 transition will also cause Q_1 to toggle from low to high. On the second positive clock transition Q_0 will toggle from high to low. However, this negative transition of Q_0 will leave Q_1 unchanged. On the third positive clock transition Q_0 will toggle from low to high and so Q_1 will also toggle but from high to low. Finally, on the fourth positive clock transition Q_0 will toggle from high to low such that the states of the clock, Q_0 and Q_1 are the same as when the entire sequence began. That sequence will continue to repeat from that time on. The overall result is that Q_0 switches at one-half the frequency of the clock and Q_1 switches at one-half the frequency of Q_0, hence the timing diagram shown in Figure 12.25.

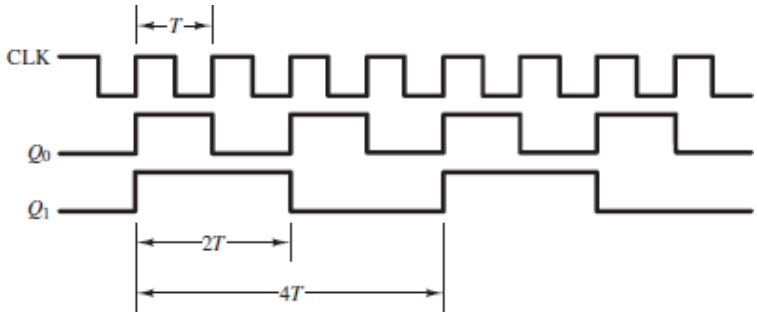


Figure 12.25 Divider circuit timing diagram

Page 716

EXAMPLE 12.5 Ring Counter

Problem

Draw the timing diagram for the ring counter of Figure 12.26.

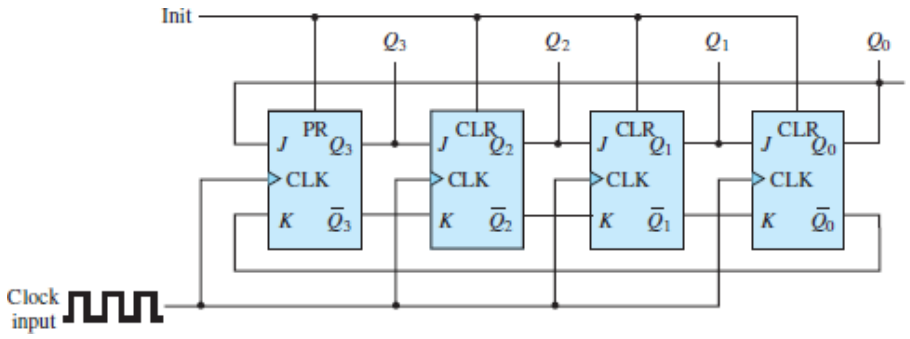


Figure 12.26

Solution

Known Quantities: *JK* flip-flop truth table (Figure 12.8).

Find: Output of each flip-flop *Q* as a function of the input clock transitions.

Assumptions: The *JK* flip-flops are positive edge–triggered. Also, the Init line is set high *after* the first positive edge transition of the clock and then is immediately set low to 0.

Analysis: After the first positive clock transition the Init line will *set* $Q_3 = 1$ and *reset* (clear) the other three flip-flops to $Q_2 = Q_1 = Q_0 = 0$. At the second positive clock transition, $Q_3 = 1$ such that the second flip-flop is *set* high to $Q_2 = 1$. Both Q_1 and Q_0 remain unchanged since their inputs were in the reset condition $J = 0$ and $K = 1$ at the clock transition. Likewise, the inputs to Q_3 were in the reset condition so it was also reset low to 0. The pattern continues, causing the 1 state to ripple from left to right over and over again as shown in the following transition table.

CLK	Q_3	Q_2	Q_1	Q_0
↑	1	0	0	0
↑	0	1	0	0
↑	0	0	1	0
↑	0	0	0	1
↑	1	0	0	0
↑	0	1	0	0
↑	0	0	1	0

Comments: The structure depicted in Figure 12.26 is known as a "ring" counter because the outputs of one flip-flop are the inputs to the succeeding flip-flop. That is, $Q_0 \rightarrow Q_3$, $Q_3 \rightarrow Q_2$, $Q_2 \rightarrow Q_1$ and $Q_1 \rightarrow Q_0$, and so on and so on.

CHECK YOUR UNDERSTANDING

The speed of the rotating encoder of the Focus on Measurements box, "Digital Measurement of Angular Position and Velocity," is found to be 9,425 rad/s. The encoder timer reads 10, and the clock counter reads 300. Assuming that both the timer counter and the encoder counter started at zero, find the clock frequency.

Answer: 45 kHz

12.3 SEQUENTIAL LOGIC DESIGN

The design of sequential circuits, just like the design of combinational circuits, can be carried out by means of a systematic procedure. A **state diagram** and its associated **state transition table** describe the logic states and their interrelationships required of the system design. Consider the 3-bit binary counter of Figure 12.27, which is made up of three T flip-flops. The input equations for this counter are $T_1 = 1$, $T_2 = q_1$, and $T_3 = q_1 \cdot q_2$. Knowing the inputs, the three outputs can be determined at any moment. The outputs Q_1, Q_2, and Q_3 form the **state** of the machine. It is straightforward to show that as the clock goes through a series of cycles, the counter will go through the transitions shown in Table 12.1, where the current state is indicated by lowercase q and the next state by an uppercase Q. Note that the state diagram of Figure 12.27 provides information regarding the sequence of states assumed by the counter in graphical form. In a state diagram, each state is denoted by a circle called a **node**, and the transition from one state to another is indicated by a **directed edge**, that is, a line with a directional arrow. The analysis of sequential circuits consists of determining either their state transition table or their state diagram.

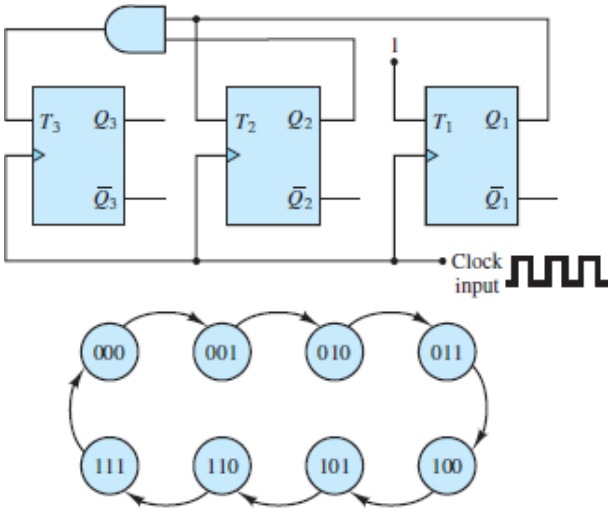


Figure 12.27 A 3-bit binary counter and state diagram

Page 718

Table 12.1 State transition table for 3-bit binary counter

Current state			Input			Next state		
q_3	q_2	q_1	T_3	T_2	T_1	Q_3	Q_2	Q_1
0	0	0	0	0	1	0	0	1
0	0	1	0	1	1	0	1	0
0	1	0	0	0	1	0	1	1
0	1	1	1	1	1	1	0	0
1	0	0	0	0	1	1	0	1
1	0	1	0	1	1	1	1	0
1	1	0	0	0	1	1	1	1
1	1	1	1	1	1	0	0	0

The reverse of this analysis process is the design process. That is, how can one systematically design a sequential circuit, such as a counter, by employing state transition tables and state diagrams?

The goal of the design process is to identify a logic circuit that matches the design specifications. There is no single unique implementation for a given set of output specifications. Therefore, the first step is to select a flip-flop and use its truth table characteristics to define its **excitation table**. The truth and excitation tables for the *RS*, *D*, and *JK* flip-flops are given in Tables 12.2, 12.3, and 12.4, respectively. Notice that each line in the excitation table represents one or more lines in the truth table that have the same pair of output states Q_t and Q_{t+1}. The don't-care entries indicate inputs that do not impact particular state transitions.

Table 12.2 Truth table and excitation table for *RS* flip-flop

Truth table for RS flip-flop				Excitation table for RS flip-flop			
S	R	Q_t	Q_{t+1}	Q_t	Q_{t+1}	S	R
0	0	0	0	0	0	0	$d^\dagger$
0	0	1	1	0	1	1	0
0	1	0	0	1	0	0	1
0	1	1	0	1	1	d	0
1	0	0	1				
1	0	1	1				
1	1	x*	x				
1	1	x	x				

*An x indicates that this combination of inputs is not allowed.
†A *d* denotes a don't-care entry.

Table 12.3 Truth table and excitation table for *D* flip-flop

Truth table for D flip-flop			*Excitation table for D flip-flop*		
D	Q_t	Q_{t+1}	Q_t	Q_{t+1}	D
0	0	0	0	0	0
0	1	0	0	1	1
1	0	1	1	0	0
1	1	1	1	1	1

Table 12.4 Truth table and excitation table for *JK* flip-flop

Truth table for JK flip-flop				*Excitation table for JK flip-flop*			
J	K	Q_t	Q_{t+1}	Q_t	Q_{t+1}	J	K
0	0	0	0	0	0	0	$d^\dagger$
0	0	1	1	0	1	1	d
0	1	0	0	1	0	d	1
0	1	1	0	1	1	d	0
1	0	0	1				
1	0	1	1				
1	1	0	1				
1	1	1	0				

†A *d* denotes a don't-care entry.

Page 719

The use of excitation tables will now be demonstrated in the design of a **modulo-4 binary up-down counter**. The phrase "modulo-4 binary" indicates that the counter output is limited to the integers 0 to 3 represented in binary form; that is, in bits. Of course, these four integers can be completely represented by 2 bits. The phrase "up-down" indicates that the counter will increment or decrement its output depending upon the value of a single bit input, which will be high or low (1 or 0) to increment or decrement the output, respectively. Figure 12.28 shows the state diagram for this counter, where a clockwise or counterclockwise progression is an increment or decrement, respectively. One flip-flop is required to produce the two states ($Q = 0$ and $Q = 1$) of each of the two output bits. For this example design, two *RS* flip-flops are used to construct the state transition table shown in Table 12.5. Note immediately that for a device with a single bit input and a double bit output there are eight distinct combinations of inputs and outputs. The first five columns of Table 12.5 specify the desired *next state* Q_1Q_2 for each possible input x and *current state* q_1q_2. This information matches the information presented in Figure 12.28.

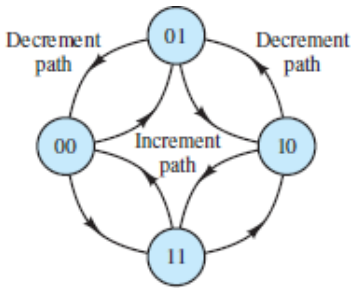


Figure 12.28 State diagram of a modulo-4 binary up-down counter

LO

Table 12.5 State transition table for modulo-4 binary up-down counter

Input x	Current state q_1	Current state q_2	Next state Q_1	Next state Q_2	S_1	R_1	S_2	R_2	Output y
0	0	0	1	1	1	0	1	0	1
0	0	1	0	0	0	*d*	0	1	0
0	1	0	0	1	0	1	1	0	1
0	1	1	1	0	*d*	0	0	1	0
1	0	0	0	1	0	*d*	1	0	1
1	0	1	1	0	1	0	0	1	0
1	1	0	1	1	*d*	0	1	0	1
1	1	1	0	0	0	1	0	1	0

Next, match the values of each output pair (Q_2, Q_{2+1}) found in the *RS* flip-flop excitation table to each of the two pairs of counter outputs (q_1, Q_1) and (q_2, Q_2) to determine the *RS* input pairs (S_1, R_1) and (S_2, R_2). For example, the first row of the counter's state transition table is developed by matching ($q_1 = 0$, $Q_1 = 1$) to the second row of the *RS* excitation table where ($Q_2 = 0$, $Q_{2+1} = 1$). Thus, the *RS* input pair ($S_1 = 1$, $R_1 = 0$) will produce the desired relationship between the *current state* variable q_1 and the *next state* variable Q_1. For the same first row of the state transition table, since $q_2 = q_1 = 0$ and $Q_2 = Q_1 = 1$, the other *RS* input pair must also be ($S_2 = 1$, $R_2 = 0$.) The other rows of the state transition table are filled out in exactly the same manner. A *d* in the table represents a don't-care condition. Remember that for this counter $x = 0$ indicates a decrement and $x = 1$ indicates an increment.

At this point, the required logic circuit can be determined using combinational logic tools, such as the Karnaugh maps of Figure 12.29. Verify that the following

expressions can be obtained from those maps.

$$S_1 = \bar{x}\bar{q}_1\bar{q}_2 + x\bar{q}_1 q_2 = (\bar{x}\bar{q}_2 + x q_2)\bar{q}_1$$
$$R_1 = \bar{x} q_1\bar{q}_2 + x q_1 q_2 = (\bar{x}\bar{q}_2 + x q_2) q_1$$
$$S_2 = \bar{q}_2$$
$$R_2 = q_2$$

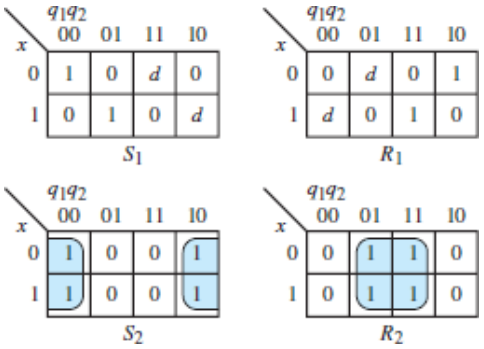


Figure 12.29 Karnaugh maps for flip-flop inputs in modulo-4 counter

The complete design is shown in Figure 12.30.

Page 720

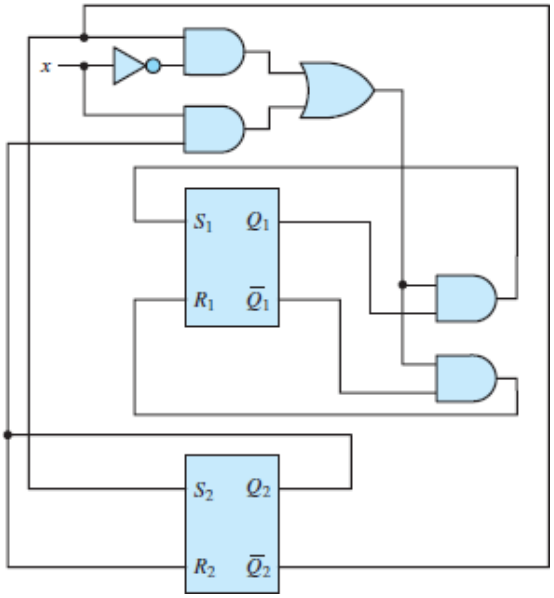


Figure 12.30 Implementation of modulo-4 counter

Programmable Logic Controllers

Sequential logic designs and state machines are found in programmable logic controllers, or PLCs, which are finite-state machines that are used in a variety of industrial applications to implement logic functions. For example, machining, packaging, material handling, and automated assembly are some of the example applications in which these systems are encountered. PLCs are specialized computers that are very effective at executing a series of complex logical decisions. Not long ago, microcontrollers began to replace PLCs in many industrial applications. The basic architecture of microcontrollers and a specific example of one are discussed in the remaining sections of this chapter.

12.4 COMPUTER SYSTEM ARCHITECTURE

The general structure of a computer system is shown in Figure 12.31. At the far left is the **central processing unit** (CPU), which receives data over the **CPU bus** from the memory and input data blocks and transmits data over the same Page 721bus to the output blocks. The CPU bus is simply a very low-resistance conducting pathway over which electrical signals (pulses) are transferred. Generally, only one set of data signals can travel over the bus at any one time; thus, it is important that transmitted data be managed properly to prevent interference. A CPU bus is usually characterized by the number of bits that are transferred during one *clock* cycle. Program instructions and addresses associated with memory locations are transferred along a separate bus. The CPU manages the flow of data and coordinates the different functions of the computer, in addition to processing data—in effect, the CPU is the heart and brains of the computer. A single CPU located on a single integrated circuit is known as a **single-core microprocessor**.

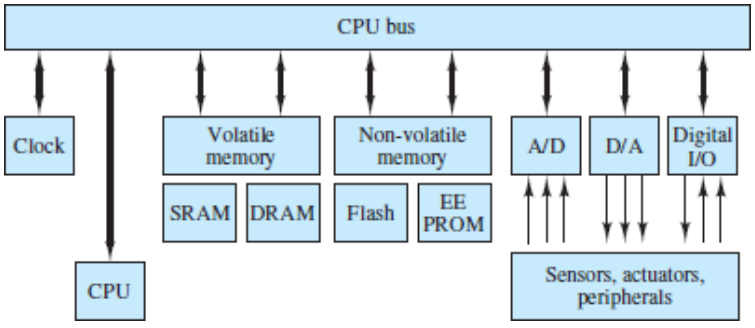


Figure 12.31 Generic computer architecture

One of the important features of a digital computer is its ability to store data. This is made possible by two general types of **memory**: *volatile and nonvolatile*.

The former type requires a steady source of power to retain its data, while the latter does not. Examples of volatile memory are **random-access memory**, such as **SRAM** and **DRAM**, which can be directly and quickly read and written to by the CPU and are therefore used primarily during the execution of programs to store partial or permanent results, and generally to store all the software currently in use by the computer. Examples of nonvolatile memory are **read-only memory**, such as **EEPROM** and **Flash memory**, which can be overwritten in smaller blocks than EEPROM. There are two types of flash memory: NAND and NOR. The former is found in many portable commercial devices. Other familiar types of nonvolatile memory are the **mass storage devices**, such as hard drives, solid-state drives, optical drives, and tape drives.

These various types of memory used within a computer offer tradeoffs between fixed cost, speed, reliability, durability, and power consumption. The main advantage of RAM is its access speed, with a latency period on the order of nanoseconds, whereas a typical latency period for a hard drive is on the order of microseconds. However, mass storage device memory is far less expensive per unit of memory.

Analog-to-digital converters (ADCs) and **digital-to-analog converters** (DACs) enable a computer to retrieve data from external sensors and transmit data to external actuators, respectively. Details of their operation and specifications can be found in Chapter 7.

Typical peripheral devices include a keyboard, a mouse, audio speakers and earphones, printers, and displays. There are many other peripheral devices available for a modern computer. These devices are commonly connected via USB ports and cables as well as many other types of network communication ports and cables.

The Clock

The **clock** represents the heartbeat of the CPU. The clock function is typically implemented by a **crystal oscillator** that determines the rate at which instructions are executed.

Memory

The CPU needs to have access to different kinds of memory to execute programs. Nonvolatile read-only memory (ROM) is used for permanent programs and data tPage 722hat are necessary, for example, to boot and initialize the system. Information stored in ROM remains unchanged even when power to the computer is turned off. Volatile random access read/write memory (RAM) is used to

temporarily store data and instructions. For example, the program that is executed by the CPU and the intermediate results of the calculations are stored in RAM. Many microcontrollers also employ electrically erasable programmable read-only (EEPROM) and **Flash** memory, which enable small changes in memory without overwriting the entire ROM.

Computer memory is arranged on the basis of **bits**, that is, a single digital variable with a value of 0 or 1. Bits are grouped in **bytes**, consisting of 8 bits, and in **words**, consisting of 16 or 32 bits. While the size of a word can vary, 1 byte always consists of 8 bits.

Mass storage devices can also be used to increase a computer's data storage capacity. However, access times to data stored in such devices are several orders of magnitude larger than access times for ROM and RAM.

Computer Programs

A computer program is a listing of instructions to be executed by the CPU. The instructions are coded in a special **machine language** that consists of combinations of bytes. To assist the programmer, each CPU instruction is encoded as a **mnemonic** for the actual operational instruction codes (**op codes**).

More commonly, a computer is programmed in a higher-level language such as C, C++, C#, or Java; the high-level language program is then translated to machine code by a **compiler**. High-level programming languages use various codes as well. A good example of such a code is the **ASCII**[2] character code, which represents all alphanumeric characters, and others, commonly used in printed documents and on computer displays, as hexadecimal values. This code is used to define the visual output associated with **char** type variables found in all high-level programming languages. It is important to realize that a **char** type variable stores an integer; however, the integer is interpreted as a reference to the ASCII character code. Thus, **char** type variables can be manipulated as integers (e.g., added and subtracted) to produce various results, such as the conversion of letters from upper- to lower-case, or vice-versa, by adding, or subtracting, 0*x*20 (32 decimal), respectively. The 128 members of the standard ASCII character set are listed in Appendix D along with their hexadecimal equivalents.

The ASCII code could be used, for example, to relate the numeric output of a seven-segment display (see the Focus on Measurements box "Seven-Segment Display" earlier in this chapter) to a computer program designed to produce numeric characters on such a display. For example, the ASCII code defines the hexadecimal equivalents of the digits 0 to 9 as 0*x*30 to 0*x*39. These hexadecimal

values are then readily converted to the BCD inputs commonly used by a seven-segment display driver chip.

CPU Registers

A CPU directly accesses volatile memory cells called **registers** to retrieve data and to store the results of computations, particularly those of an **arithmetic logic unit** (ALU), which is often incorporated into a CPU. A **memory map** defines the names Page 723and types of the memory locations that are accessible to the CPU in addition to the registers. Several of the typical uses of these registers are listed below.

An ***accumulator*** may be used to hold the results of arithmetic operations performed by the CPU.

An ***index*** is used to point to an address in memory where the CPU will read or write information.

A ***program counter (PC) register*** keeps track of the address of the next instruction to be executed by the CPU.

The ***condition code register (CCR)*** holds information that reflects the status of prior CPU operations. For example, *branch instructions* look at the CCR to make either/or decisions.

The ***stack pointer (SP) register*** contains return address information and the previous content of all CPU registers. When the CPU is interrupted or a subroutine is initiated, the status of the program is retained prior to the **interrupt** or subroutine branching. After the CPU has *serviced* the interrupt or has completed the subroutine, it can resume its previous operations by loading the contents of the SP register.

Interrupts

Interrupts perform an important function by allowing the CPU to interrupt its normal flow of operations to respond to an external event. For example, an interrupt request may occur when an ADC makes available to the CPU the digital value of a sensor reading.

FOCUS ON MEASUREMENTS

Reading Sensor Data Using Interrupts

In modern automotive instrumentation, a microcontroller performs all the signal processing operations for several measurements. A block diagram for such instrumentation is given in Figure 12.32. Depending on the technology used, the sensors' outputs can be either digital or analog. If the sensor signals are analog, they must be converted to digital format by means of an analog-to-digital converter, as shown in Figure 12.33. The analog-to-digital conversion process requires an amount of time that depends on Page 724the individual ADC, as is explained in Chapter 7. After the conversion is completed, the ADC then signals the computer by changing the logic state on a separate line that sets its *interrupt request* flip-flop. This flip-flop stores the ADC's interrupt request until it is acknowledged (see Figure 12.34).

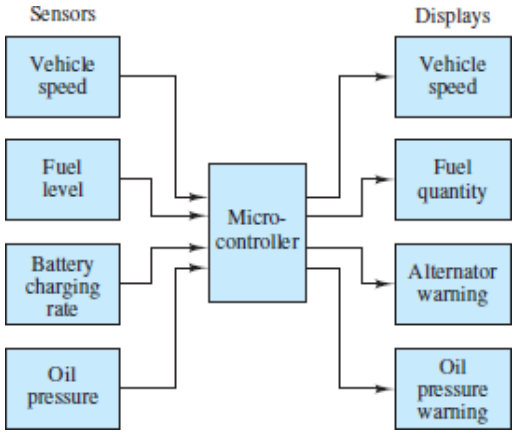


Figure 12.32 Automotive instrumentation

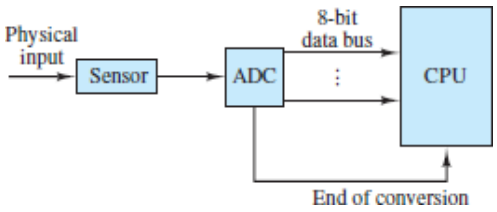


Figure 12.33 Sensor interface

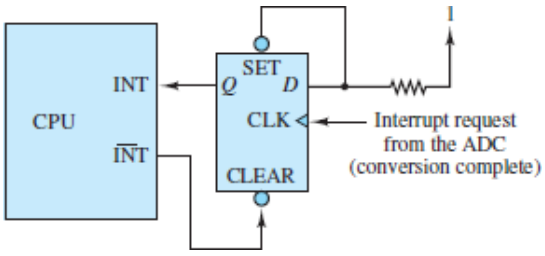


Figure 12.34 Interrupt request

When an interrupt occurs, the processor automatically jumps to a designated program location and executes the interrupt service subroutine. For the ADC, this would be a subroutine to read the conversion results and store them in some appropriate location, or to perform an operation on them. When the processor responds to the interrupt, the interrupt request flip-flop is cleared by a direct signal from the processor. To resume execution of the program at the proper point upon completion of the ADC service subroutine, the program counter content is automatically saved before control is transferred to the service subroutine. The service subroutine saves in a stack the content of any registers it uses and restores the registers' content before returning.

The interrupt may occur at any point in a program's execution, independent of the internal clock; it is therefore referred to as an *asynchronous* event.

Of course, a wide variety of computer systems are manufactured. Large mainframes are used by industry to store and analyze vast amounts of data. Such mainframes employ **parallel processing** wherein numerous CPUs work in concert to achieve very high rates of computation. Smaller computer systems, such as a laptop or desktop computer, contain a **microprocessor** that typically has multiple **cores**, which act as a set of CPUs working in parallel. On a smaller scale still are **microcontrollers**, which consist of a CPU, memory, and input and output ports, all on a single integrated-circuit chip mounted on a single **printed circuit board** (PCB). Some microcontrollers are designed to be versatile with no

particular dedicated application in mind. Others are designed for specific applications and can be embedded directly in devices and systems.

12.5 THE ATMEGA328P MICROCONTROLLER

Microcontrollers have become an essential part of many engineering products, processes, and systems and are often embedded in large and small products and systems, such as automobiles, home appliances, and portable devices. The Page 725ATmega328P®, made by Atmel® Corporation, is a versatile **reduced instruction set computing** (RISC) microcontroller that is used in automobiles and the well-known Arduino prototyping platform. The ATmega328P® has many of the capabilities commonly found in microcontrollers. Of course, detailed specifications vary across microcontrollers; nonetheless, a solid understanding of the ATmega328P® provides an excellent introduction to microcontrollers, in general. A block diagram of the ATmega328P® microcontroller is shown in Figure 12.35.

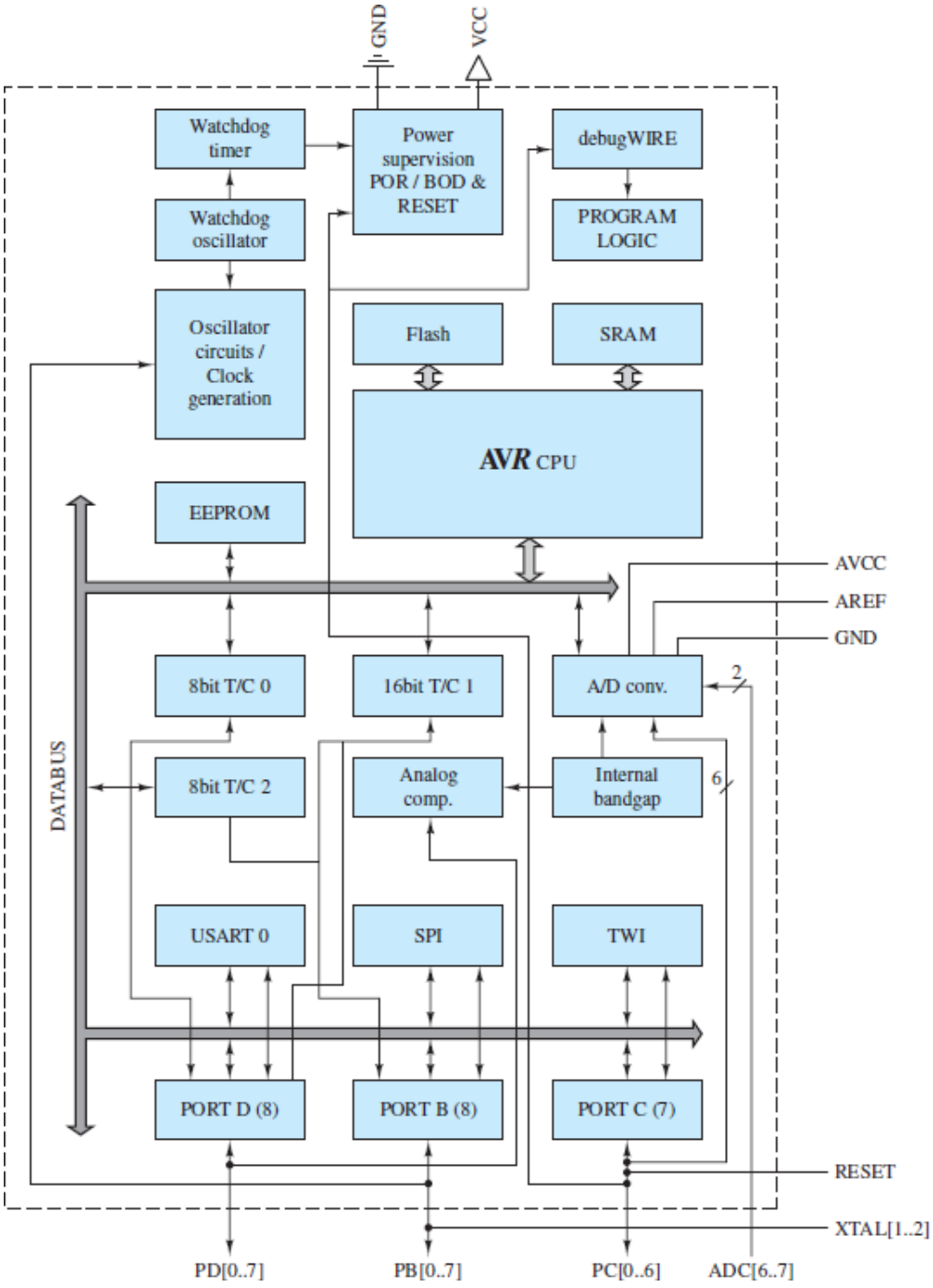


Figure 12.35 Block diagram of the ATmega328P® microcontroller *(Courtesy: Atmel® Corporation, which owns the intellectual property. Copyright Atmel Corporation 2013)*

Capabilities and Pin Configurations

The external pins of the ATmega328P® access two 8-bit bidirectional I/O ports, one 7-bit bidirectional I/O port, an analog voltage reference for analog-to-digital (A/D) conversion, and a reset. Additional pins are for connecting ground and power supplies. The ports are designed to provide specialized functions, such as six 10-bit ADCs, six **pulse-width modulation** (PWM) output channels, and two 8-bit and one 16-bit timer/counters. Portions of these ports can also be tasked to provide communication interfaces, such as a programmable **universal serial asynchronous receive and transmit** (USART) interface, a serial master/slave **serial peripheral interface** (SPI), and a **two-wire serial interface** (TWI) compatible with the Philips I^2C standard.

The equivalent network of each individual I/O pin is shown in Figure 12.36. Regardless of the configuration of other pins, each pin can be configured to provide general digital I/O. The configuration is determined by three single-bit registers: DD*xn*, PORT*xn*, and PIN*xn*, where *x* refers to the port (*B*, *C*, or *D*) and *n* refers to a specific pin on a port. By default, each pin is configured as a tristate input with [DD*xn* PORT*xn*] set to [00]. Setting PORT*xn* to logic 1 in input mode [01] enables an internal 20K pull-up resistor such that the pin takes on an *open-collector* type behavior. Setting DD*xn* to logic 1 configures a pin for output. Table 12.6 shows the complete list of possible states for each I/O pin.

Table 12.6 ATmega328P® I/O pin configurations

DDxn	*PORTxn*	*I/O*	*Pull-up*	*Configuration*
0	0	Input	No	Tri-state (Hi-Z)
0	1	Input	Yes	Open-collector (will source current when low)
1	0	Output	No	Output low (will sink current)
1	1	Output	No	Output high (will source current)

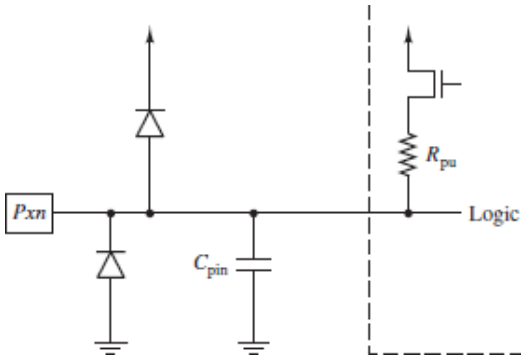


Figure 12.36 General I/O pin equivalent network *(Courtesy: Atmel® Corporation, which owns the intellectual property. Copyright Atmel Corporation 2013)*

Each I/O pin can source/sink up to 40 mA although the total current through any I/O port should not exceed 100 mA. It is important to keep in mind that exceeding the current limits will almost certainly damage a pin and may damage other parts of the microcontroller as well. It is also recommended that the state of all unused pins be defined by enabling the internal pull-up resistor or some other external method.

Page 727

Power Requirements

The ATmega328P® also provides six low-power modes to conserve battery life and generally reduce power consumption. When running at 1 MHz and 1.8 V, the ATmega328P® sinks 100 nA in power-down mode and 200 μA in active mode. Generally, the ATmega328P® operates at 4.5 to 5.5 V with a clock speed as high as 20 MHz; however, when the clock is limited to 4 MHz, it can operate at voltages as low as 1.8 V.

AVR® Architecture

The heart of the ATmega328P® microcontroller is an AVR® CPU core, which has a modified Harvard architecture, as shown in Figure 12.37. There is a dedicated bus and memory for data separate from those used for program instructions. An **arithmetic logic unit** (ALU) reads instructions and data from the thirty-two 8-bit general-purpose registers (R0–R31); performs various logical, arithmetic operations, and bit-functions on the data; and writes the results back to the registers, in most cases in one clock cycle. While the AVR® CPU is executing

these three Page 728steps, the modified Harvard architecture enables it to pre-fetch the next upcoming instruction from the Flash program memory. Three pairs of registers (R26–R31) are referred to as *X*, *Y*, and *Z* and may be used to store three 16-bit indirect address pointers; *X* may also point to look-up tables stored in Flash memory. All 32 registers are also assigned direct addresses 0*x*00 through 0*x*1F (hexadecimal) such that they are the first 32 bytes of the overall memory map. The individual bits in these registers can be accessed using the SBIS, SBIC, SBI, and CBI instructions (see below).

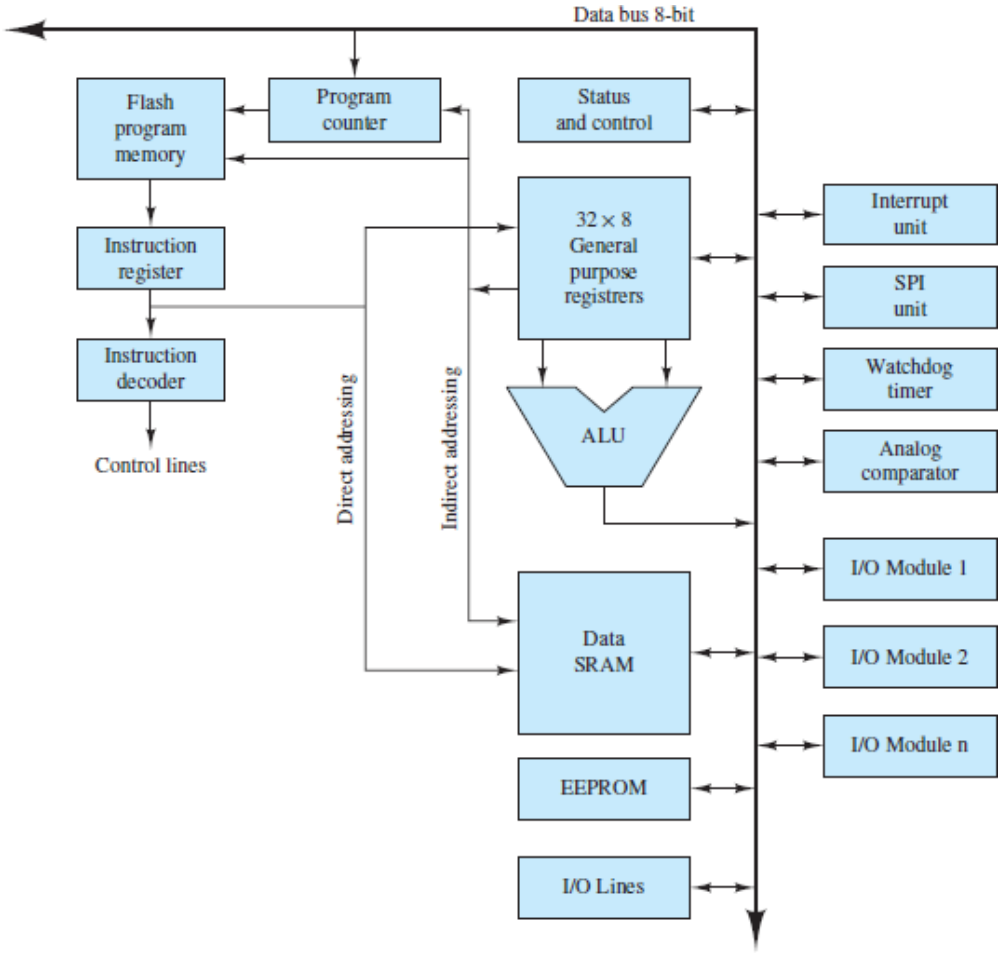


Figure 12.37 Block diagram of the ATmega328P® AVR® CPU architecture *(Courtesy: Atmel® Corporation, which owns the intellectual property. Copyright Atmel Corporation 2013)*

Memory

The ATmega328P® possesses nonvolatile memory of up to 32K bytes of in-system programmable Flash memory with read-while-write capability and 1K byte of EEPROM. These two memory banks can reliably undergo 10^4 and 10^5 write/erase cycles, respectively, over a lifetime. The Flash memory is organized into 16K 2-byte words because AVR® instructions are either 2 or 4 bytes long. These words are addressed by a 14-bit **program counter**. (*Note:* 14 bits can represent 2^{14} = 16,386 different items, or in this case, addresses.) The EEPROM can be accessed via an SPI interface. The read/write processes to EEPROM have a typical write time of 3.3 ms.

Up to 2K bytes of SRAM are also available for storing data sent to and from the AVR® registers. The SRAM is also used to store the **stack**, which includes the program counter.

AVR® Instructions

The AVR® operates on 131 different instructions, examples of which are listed in the Table 12.7. Most of these instructions are executed in a single clock cycle, which can be as short as 50 ns resulting in a throughput of up to 20 million instructions per second (MIPS).

Table 12.7 Example ATmega328P® Instructions

Type	*Mnemonic*	*Operands*	*Description*	*Operation*
Arithmetic	ADD	Rd, Rr	Add without Carry	Rd ← Rd + Rr
Arithmetic	ADC	Rd, Rr	Add with Carry two Registers	Rd ← Rd + Rr + C
Logic	AND	Rd, Rr	Logical AND registers	Rd ← Rd • Rr
Logic	NEG	Rd	Two's Complement	Rd ← 0x00 - Rd
Logic	CLR	Rd	Clear Register	Rd ← Rd ⊕ Rd
Branching	RJMP	k	Relative Jump	PC ← PC + k + 1
Branching	SBIS	P, B	Skip if Bit in I/O Register Cleared	if P(B) = 1 then PC ← PC + 2
Branching	SBIC	P, B	Skip if Bit in I/O Register Cleared	if P(B) = 0 then PC ← PC + 2
Bitwise	SBI	P, B	Set Bit in I/O Register	I/O(P,B) ← 1
Bitwise	CBI	P, B	Clear Bit in I/O Register	I/O(P,B) ← 0
Bitwise	LSL	Rd	Logical Shift Left	Rd(n+1) ← Rd(n), Rd(0) ← 0
Bitwise	LSR	Rd	Logical Shift Right	Rd(n) ← Rd(n+1), Rd(7) ← 0
Data Transfer	MOV	Rd, Rr	Move Between Registers	Rd ← Rr
Data Transfer	LD	Rd, X	Load Indirect	Rd ← (X)
Data Transfer	ST	X, Rr	Store Indirect	(X) ← Rr

Each mnemonic represents a 16-bit **op code**, such as [1001 1001 *pppp pbbb*] for SBIC P, B. In the op code, the *p* sequence indicates the specific general-purpose register *P* of the 32 available in the AVR® and the *b* sequence indicates the specific bit *B* within that register. Whereas many of the AVR® instructions can access various portions of the overall memory map, the SBI and CBI instructions only interact with the 32 general-purpose registers located at 0*x*00 through 0*x*1F. Many more details of the AVR® instruction set are available online.

A short sample code of AVR® instructions for configuring the external port *B* pins of the ATmega328P® is shown below. Also shown is an equivalent code written in C for comparison.[3]

```
; Define pull-ups and set outputs high
; Define directions for port pins
ldi  r16,(1<<PB7)|(1<<PB6)|(1<<PB1)|(1<<PB0)
ldi  r17,(1<<DDB3)|(1<<DDB2)|(1<<DDB1)|(1<<DDB0)
out  PORTB,r16
out  DDRB,r17
; Insert nop for synchronization nop
; Read port pins
in  r16,PINB
```

```
unsigned char i
/* Define pull-ups and set outputs high */
 /* Define directions for port pins */
 PORTB = (1<<PB7)|(1<<PB6)|(1<<PB1)|(1<<PB0);
 DDRB = (1<<DDB3)|(1<<DDB2)|(1<<DDB1)|(1<<DDB0);
 /* Insert no-op for synchronization*/
  __no_operation();
/* Read port pins */
 i = PINB;
```

Another sample code for reading and writing to EEPROM is shown below. Again the equivalent C code is presented for comparison. Both codes are presented as functions to be called.[4]

```
EEPROM_write:
   ; Wait for completion of previous write
   sbic EECR,EEPE
   rjmp EEPROM_write
   ; Set up address (r18:r17) in address register
   out EEARH, r18
   out EEARL, r17
   ; Write data (r16) to Data Register
   out EEDR,r16
   ; Write logical one to EEMPE
   sbi EECR,EEMPE
   ; Start eeprom write by setting EEPE
   sbi EECR,EEPE
   ret
```

```
void EEPROM_write(unsigned int uiAddress, unsigned
char ucData)
{
  /* Wait for completion of previous write */
  while(EECR & (1<<EEPE));
  /* Set up address and Data Registers */
  EEAR = uiAddress;
  EEDR = ucData;
  /* Write logical one to EEMPE */
  EECR |= (1<<EEMPE);
  /* Start eeprom write by setting EEPE */
  EECR |= (1<<EEPE);
}
```

Additional detailed information and development tools, including an integrated software environment, are freely available from the Atmel® website *www.atmel.com*.

Mechatronics and Embedded Systems

Industry and consumers demand engineering processes and products that are more reliable, more efficient, smaller, faster, and less expensive. The production and development of these devices require engineers who have integrated perspectives on system design. The discipline of *mechatronic design* involves the integration of mechanical, electrical, and computer science engineering (Figure 12.38). Design elements from these traditional disciplines don't simply exist side by side but are deeply integrated in the design process. Whether a given functionality should be achieved electronically, by software, or by elements from electrical or mechanical engineering domains requires mastery of analysis and synthesis techniques from the different areas. Being a successful mechatronics design engineer requires an in-depth understanding of many of, if not all, its constituent disciplines. Most major programs in the United States don't emphasize mechatronics as a primary curriculum component, but industry continues to motivate its development. The automotive, aerospace, manufacturing, power systems, test and instrumentation, consumer, and industrial electronics industries make use of and contribute to mechatronics.

One of the distinguishing features of the mechatronic approach to the design of products and processes is the use of *embedded microcontrollers*, which replace many mechanical functions with electronic ones, resulting in

much greater flexibility, ease of redesign or reprogramming, the ability to implement distributed control in complex systems, and the ability to conduct automated data collection and reporting. Mechatronic design represents the fusion of traditional mechanical, electrical, and software engineering design methods with sensors and instrumentation technology, electric drive and actuator technology, and embedded real-time microcontrollers and real-time software. Mechatronic systems range from heavy industrial machinery, to vehicle propulsion systems, to precision electromechanical motion control devices.

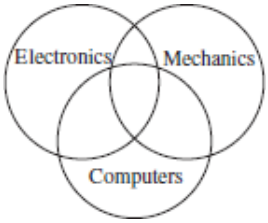


Figure 12.38 Mechatronics as the intersection of three engineering disciplines

Page 731

12.6 THE ARDUINO™ PROJECT

The **Arduino** project is a collection of open-source prototyping hardware and software resources aimed at casual users and professionals. In fact, the Arduino webpage states that the resources are "intended for artists, designers, hobbyists, and anyone interested in creating interactive objects or environments."[5] There are many incarnations of the Arduino hardware prototyping platform. The **Arduino Uno** hardware was developed around the ATmega328P® microcontroller.

The Arduino is well suited for academic projects, particularly those involving undergraduate engineering students because it was designed to handle and hide many of the complex details of working directly with a microprocessor. Commonly desired communication capabilities, such as a USB interface and serial output to a computer display, are built in. The Arduino's USB interface has a resettable polyfuse to protect against overcurrent damage to the host computer.[6] The Arduino also does not require the user to **burn** firmware onto the ATmega328P® chip. Finally, many high-level programming functions are

provided by the Arduino Team and third parties that integrate lower-level AVR® instructions to provide easy access to powerful and versatile capabilities, such as ADC.

Features

Figure 12.39 shows the Arduino Uno R3 board. It has fourteen digital I/O pins, six 10-bit analog input pins, a USB connector, a 6-pin *in-circuit serial programming* (ICSP) header, and a reset button. The Arduino prefers an external 7 to 12 VDC source, either from a battery, or through the DC power jack or USB cable. The board provides 5-V and 3.3-V regulated DC power pins with current limits of 40 mA and 50 mA, respectively. Internally, the Arduino operates on 5 VDC and each of its pins has the same 40-mA DC current limit as the ATmega328P®. Likewise, the available SRAM, EEPROM, and Flash memory are 2, 1, and 32 KB, respectively. The on-board clock speed is 16 MHz.

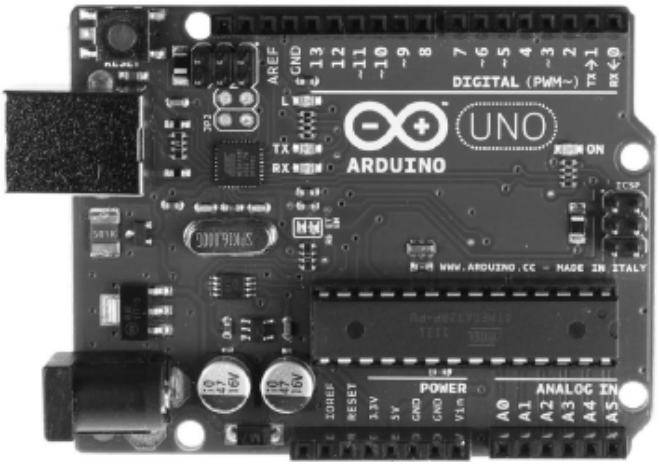


Figure 12.39 The Arduino Uno R3 PCB *(Courtesy of Arduino)*

Page 732

As with the ATmega328P®, the Arduino's 14 digital I/O pins may be individually configured as either inputs or outputs, with high and low values of 5 and 0 V. Each pin can source or sink up to 40 mA DC and may be enabled with an internal 20K pull-up resistor. Many of these pins can be individually tasked to provide specific useful capabilities, including six 8-bit PWM outputs (pins 3, 5, 6,

and 9 to 11), SPI communication (pins 10 to 13), UART serial communication (pins 0 and 1), and two external interrupts (pins 2 and 3).

Elsewhere on the board, pins A4 and A5 can be tasked as serial data (SDA) and serial clock (SCL) pins, respectively, to provide TWI (I^2C) communication. Additionally, the Arduino provides a regulated DC reference voltage (AREF) pin for use in conjunction with its analog input capability. A dedicated reset input pin can be driven low to reset the Arduino from an external switch or controlled source. Figure 12.40 shows the pin mapping between the ATmeg328 and the Arduino Uno.

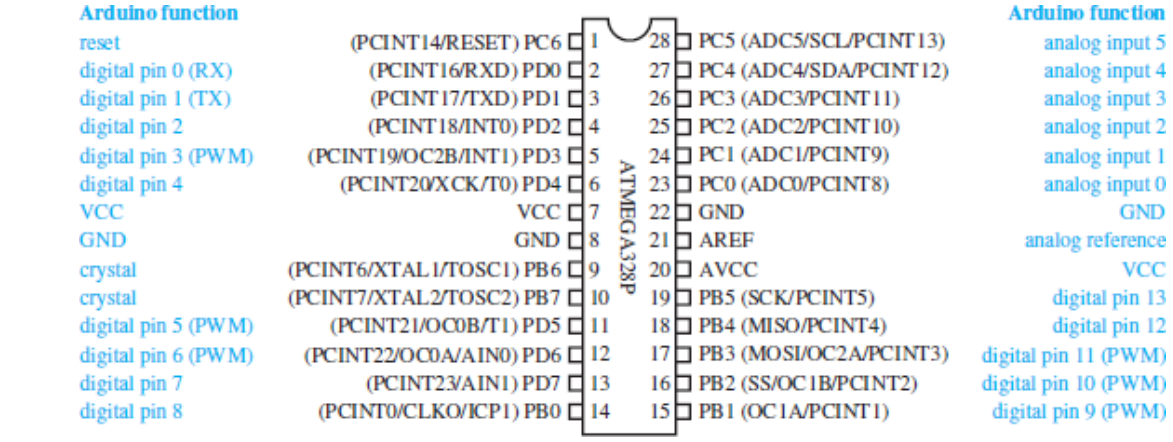


Figure 12.40 Pin mapping between ATmega328P® and Arduino Uno (*Courtesy: Atmel® Corporation, which owns the intellectual property. Copyright Atmel Corporation 2013*)

The ATmega328P® microcontroller chip on the Arduino Uno is preloaded with a **bootloader** that enables a user-written **sketch** to be uploaded without any additional burning hardware and software. The Arduino also provides its own built-in burner in case a user wishes to create his or her own Arduino board or install the bootloader on a stand-alone ATmega328P® chip.

User Programs: Sketches

In the parlance of the Arduino project, a sketch is a user-written C/C++ program. Variables and functions obey the same rules for *type*, *scope*, and *storage class*. Every sketch must include two particular functions: *setup()* and *loop()*. The

Arduino project provides a large library of high-level functions, all of which are well organized and documented on the Arduino website.

Shown below is a sketch that illustrates the use of the Arduino's I/O pins' pull-up resistors.[7]

Page 733

```
/* This program illustrates the use of Arduino I/O pin functions */

// Pins 2 and 4 are standard digital I/O pins

int noPullUpPin = 2;
int pullUpPin = 4;
int ledPin = 13;             // Pin 13 has an LED connected on most Arduino boards.
int logicDevice, pushButton;

// the setup routine runs once when you press reset:
void setup() {
   Serial.begin(9600);             // Serial output to computer enabled at 9600 baud
   pinMode(ledPin, OUTPUT);        // initialize pin 13 as an output.
   pinMode(noPullUpPin, INPUT);           // pin 2 as an input w/ pull-up resistor disabled
   pinMode(pullUpPin, INPUT_PULLUP);      // pin 4 as an input w/ pull-up resistor enabled
}

// the loop routine runs over and over again forever:
void loop() {
   digitalWrite(led, HIGH); // turn the LED on (HIGH is the voltage level)
   delay(1000);                  // wait for one second
   digitalWrite(led, LOW);  // turn the LED off by making the voltage LOW
   delay(1000);                  // wait for one second

   logicDevice = digitalRead(noPullUpPin); // read the value of a device connected to pin 2
   delay(50);                              // a small delay
   pushButton = digitalRead(pullUpPin);        // read the value of a pushbutton at pin 4
   delay(50);

   int sensorInt = analogRead(A0); // sensorInt is a 10-bit integer 0 to 1023
   char sensorChar = sensorValue/4;    // sensorChar is sensorValue scaled to 8-bits 0 to 255

   Serial.print(''The data is: ''); Serial.print(logicDevice);
   Serial.print(''\ t''); Serial.print(pushButton);
   Serial.print(''\ t''); Serial.print(sensorInt);
   Serial.print(''\ t''); Serial.println(sensorChar, DEC);
}
```

Notice that variables are declared and initialized as in a typical C/C^{++} program. The variables declared outside of *setup()* and *loop()* have global scope. On the other hand, the scope of variables *sensorInt* and *sensorChar* is limited to the *loop* function. Recall that the Arduino provides six 10-bit analog input pins A0 to A5; thus, the 10-bit integer value is assigned to *sensorInt*. By default, integer variables are allocated 2 bytes, whereas character variables are allocated 1 byte.

The *Serial.print* function sends data to the host computer terminal via the USB port. Notice the use of tab spaces sent as single character strings ("*t*").

Serial.print does not include a hidden carriage return, whereas *Serial.println* does.

Finally, when using a digital pin for input it is good practice to always ensure the pin value is determined, and not *floating* at any time. In the above Page 734example, the *logicDevice* value being read is assumed to be determined as high or low by an external logic device. In such cases, an internal pull-up resistor on an input pin does not need to be enabled. However, for many other common circumstances, such as when a switch is located between the pin and V_2 or ground, an internal pull-up resistor should be enabled. Otherwise, when the switch is open, the state of the pin will be indeterminate; that is, determined in a manner not set by the user.

An extensive list of useful sketches can be found on the Arduino website.

Conclusion

This chapter presents an overview of digital logic circuits. These circuits form the basis of all digital computers and of most electronic devices used in industrial and consumer applications. Upon completing this chapter, a student will have learned to:

1. *Analyze the operation of flip-flops and latches*, which are the building blocks of a sequential logic circuits. *Feedback* from outputs to inputs creates outputs whose future values depend upon their present values. In other words, these circuits possess *memory*. The operation flip-flops and latches are described by state transition tables and state diagrams.
2. *Analyze and apply digital counters and registers*. Counters are a very important class of digital circuits and are based on sequential logic elements. Registers are the most fundamental form of *random-access memory* (RAM).
3. *Design simple sequential circuits using state transition diagrams.* Sequential circuits can be designed using formal design procedures employing state diagrams.
4. *Describe the basic architecture of computers* and how aspects of that architecture are utilized to provide computers with their various capabilities.
5. *Identify the basic architecture of microcontrollers*, and the ATmega328P®, in particular. A microcontroller is a system that includes a *central processing unit* (CPU) and various I/O capabilities.
6. *Utilize the Arduino project*, including its hardware specifications, its software capabilities, and some of its many practical applications.

HOMEWORK PROBLEMS

Section 12.1: Latches and Flip-Flops

12.1 The input to the circuit of Figure P12.1 is a square wave having a period of 2 s, maximum value of 5 V, and minimum value of 0 V. Assume all flip-flops are initially in the reset state.

a. Explain what the circuit does.

b. Sketch the timing diagram, including the input and all four outputs.

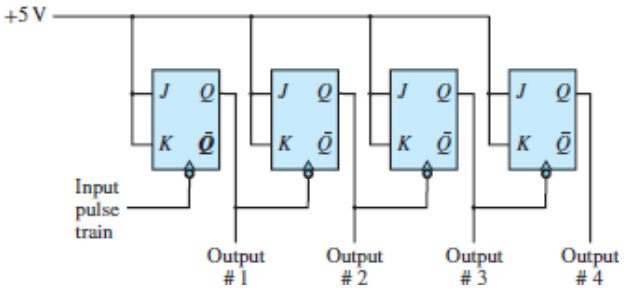


Figure P12.1

Page 735

12.2 Suppose a circuit is constructed from three *D*-type flip-flops, one input *I*, with

$$D_1 = Q_2, \qquad D_0 = Q_1$$
$$D_2 = I \cdot (Q_1 \cdot \overline{Q_0} + \overline{Q_1} \cdot Q_0) + \overline{I} \cdot (Q_1 \cdot Q_0 + \overline{Q_1} \cdot \overline{Q_0}) = \alpha \oplus I \qquad \text{with } \alpha = Q_1 \oplus Q_0$$

a. Draw the circuit diagram.

b. Assume the circuit starts with all flip-flops set. Sketch a table that shows the outputs of all three flip-flops.

12.3 Suppose that you want to use a *JK* flip-flop for a laboratory experiment. However, you have only *D* flip-flops. Assuming that you have all the logic gates available, make a *JK* flip-flop using a *D* flip-flop and some logic gate(s).

12.4 Draw a timing diagram (four complete clock cycles) for A_0, A_1, and A_2 for the circuit of Figure P12.4. Assume that all initial values are 0. Note that all flip-flops are negative edge–triggered.

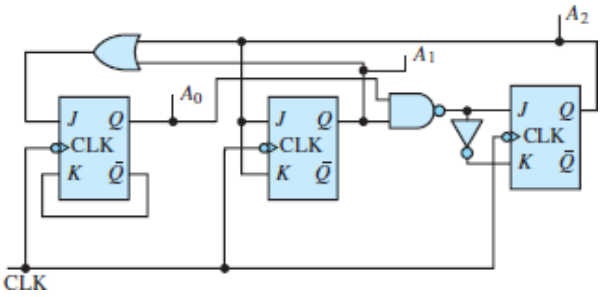


Figure P12.4

12.5 Given the sequential circuit of Figure P12.5, determine the output Y when input A is [1 0 1 1].

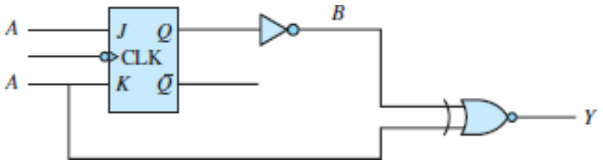


Figure P12.5

12.6 Write the truth table for an *RS* flip-flop with enable (E), preset (P), and clear (C) lines.

12.7 A *JK* flip-flop is wired as shown in Figure P12.7 with a given input signal. Assuming that Q is at logic 0 initially and the negative trailing-edge triggering is effective, sketch the output Q.

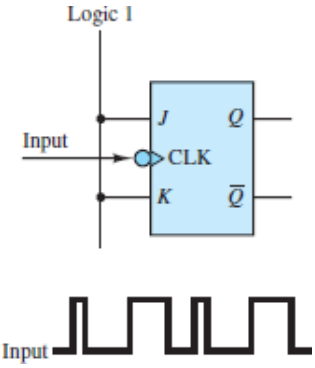


Figure P12.7

12.8 With reference to the *JK* flip-flop of Problem 12.7, assume that the output at the Q terminal is made to serve as the input to a second *JK* flip-flop wired exactly as the first. Sketch the Q output of the second flip-flop.

12.9 Figure P12.9 shows an *RS* flip-flop acting as a debouncing circuit for a single-pole, double-throw (SPDT) switch. Fill in the table to indicate the state of *Q* for each of the two switch positions *A* and *B*. What is the purpose of the two 10 K resistors?

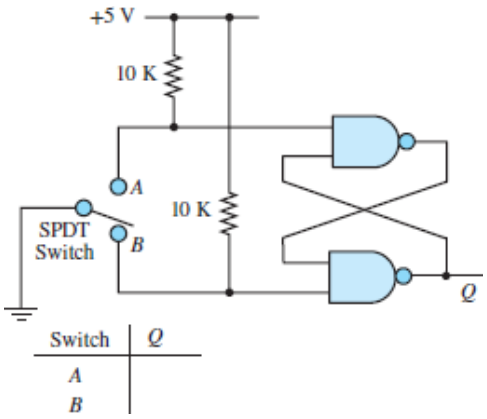


Figure P12.9

12.10 Figure P12.10 shows a *D* flip-flop with preset and clear acting as a debouncing circuit for a single-pole, double-throw (SPDT) switch. Fill in the table to indicate the state of *Q* for each of the two switch positions *A* and *B*. What is the purpose of the two 10 K resistors?

Page 736

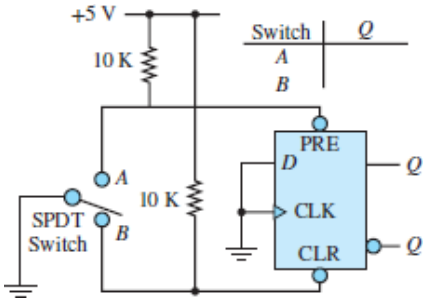


Figure P12.10

Section 12.2: Digital Counters and Registers

12.11 Assume that the slotted encoder shown in Figure P12.11 has a length of 1 m and a total of 1,000 slots (i.e., there is one slot per millimeter). If a counter is incremented by 1 each time a slot goes past a sensor, design a

digital counting system that determines the speed of the moving encoder (in meters per second).

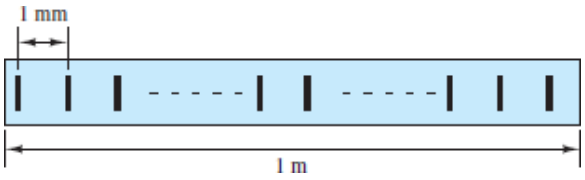


Figure P12.11

12.12 A binary pulse counter can be constructed by interconnecting *T*-type flip-flops in an appropriate manner. Assume it is desired to construct a counter that can count up to 100_{10}.

a. How many flip-flops would be required?

b. Sketch the circuit needed to implement this counter.

12.13 Explain what the circuit of Figure P12.13 does and how it works. (*Hint:* This circuit is called a 2-bit synchronous binary up-down counter.)

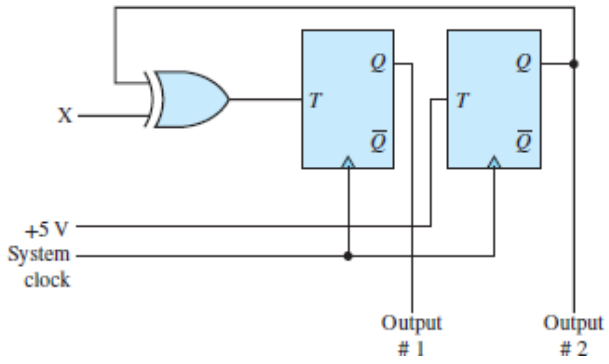


Figure P12.13

12.14 Figure P12.14 shows a simple divide-by-2 circuit using a leading (positive) edge–triggered *JK* flip-flop. Assume the clock pulse train repeats equal length low and high intervals. Draw the corresponding timing diagram for the output *Q*.

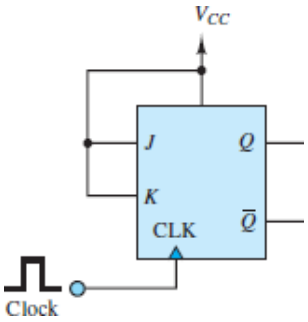


Figure P12.14

12.15 Figure P12.15 shows a simple divide-by-3 circuit using two leading (positive) edge–triggered *JK* flip-flops. Assume the clock pulse train repeats equal length low and high intervals. Draw the corresponding timing diagrams for the outputs *A* and *B* until they both repeat.

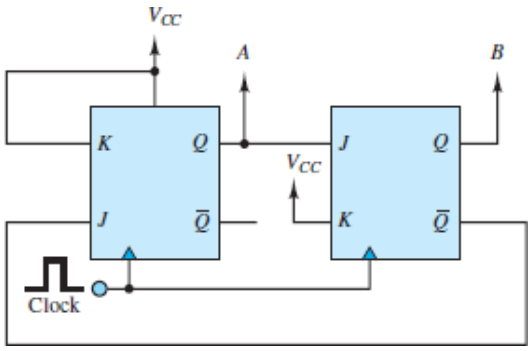


Figure P12.15

12.16 Figure P12.16 shows a simple divide-by-4 circuit using two leading (positive) edge–triggered *JK* flip-flops. Assume the clock pulse train repeats equal length low and high intervals. Draw the corresponding timing diagrams for the outputs *A* and *B* until they both repeat.

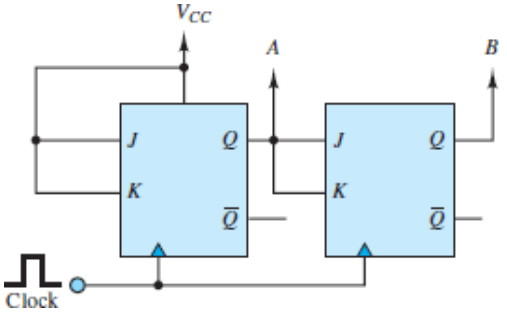


Figure P12.16

12.17 Figure P12.17 shows a Johnson counter using four leading (positive) edge–triggered *D* flip-flops, each with preset and clear. Draw the timing diagrams for the outputs Q_0, Q_1, Q_2 and Q_3 until they all repeat.

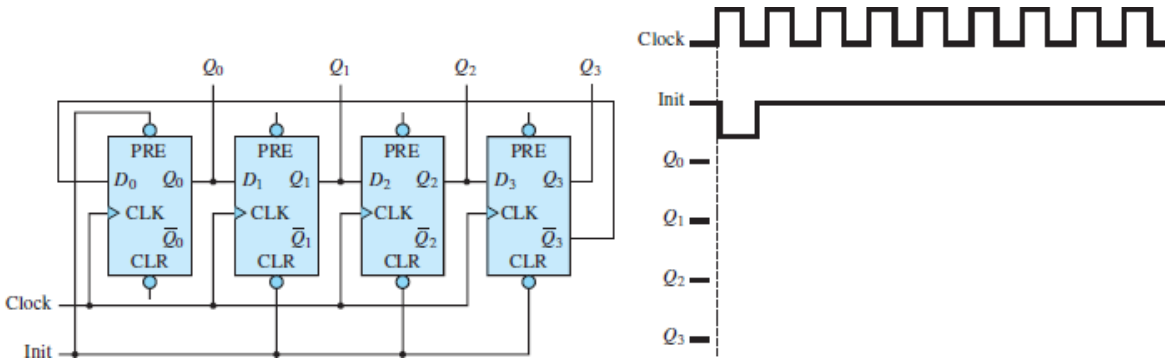


Figure P12.17

Section 12.3: Sequential Logic Design

12.18 Using necessary logic gates and *D*-type flip-flops, create a sequential circuit (one input–one output) from the state table given below.

Current state	Next state $D = Q'_{n+1}$		Output Q	
Q'_n	$I = 0$	$I = 1$	$I = 0$	$I = 1$
A	*A*	*B*	0	0
B	*B*	*A*	0	1
C	*C*	*B*	0	0
D	*D*	*A*	0	1

12.19 Use *JK* flip-flops to construct a sequential circuit with the state diagram shown in Figure P12.19.

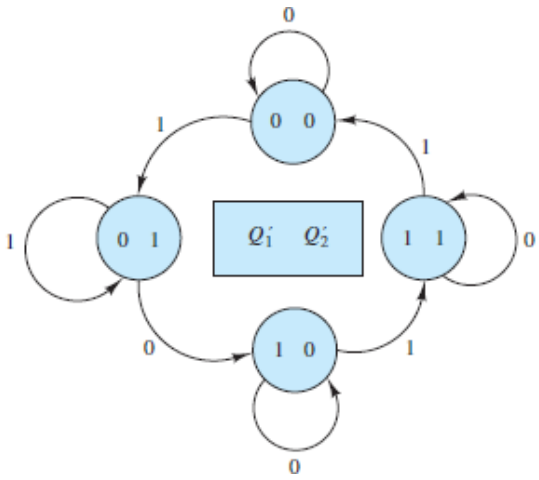


Figure P12.19

Section 12.4: Computer System Architecture

12.20 Explain the purpose of the ALU.

12.21 Name the internal registers of a microprocessor, and explain their functions.

12.22 Name the three different systems buses, and explain their functions.

12.23 Suppose a microprocessor has *n* registers.

a. How many control lines do you need to connect each register to all other registers?

b. How many control lines do you need if a bus is used?

12.24 Explain the function of the status register (flag register), and give an example.

12.25 What is the distinction between volatile and nonvolatile memory?

12.26 A typical PC has 8 GB of RAM.

a. How many 16-bit words is this?

b. How many nibbles is this?

c. How many bits is this?

12.27 Suppose it is desired to implement a 4K byte 16-bit memory.

a. How many bits are required for the memory address register?

b. How many bits are required for the memory data register?

12.28 Suppose a particular magnetic tape can be formatted with eight tracks per centimeter of tape width. The recording density is 200 bits/cm, and the transport mechanism moves the tape past the read heads at a velocity of 25 cm/s. How many bytes per second can be read from a 2-cm-wide tape?

12.29 Draw a block diagram of a circuit that will interface two interrupts, INT0 and INT1, to the INT input of a CPU so that INT1 has the higher priority and INT0 has the lower. In other words, a signal on INT1 is to be able to interrupt the CPU even when the CPU is currently handling an interrupt generated by INT0, but not vice versa.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]Atmel®, the Atmel logo and combinations thereof, AVR®, and others are registered trademarks, or trademarks of Atmel Corporation or its subsidiaries.

[2]American Standard Code for Information Interchange.

[3]These code segments are used courtesy of Atmel® Corporation and were taken directly from the document 7810C-AVR-10/12, p. 70.

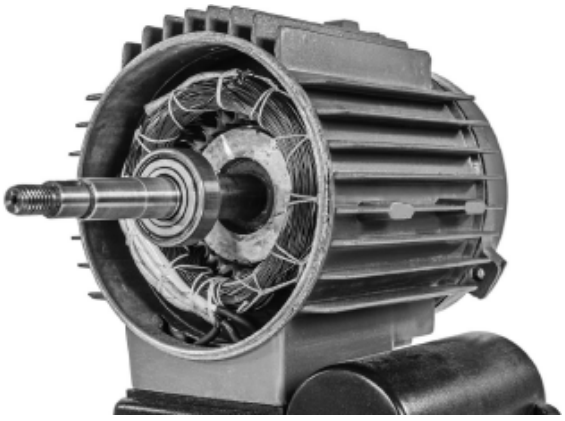
[4]Ibid., p. 23.

[5]See www.arduino.cc.

[6]The Arduino's digital I/O pins are not protected against overcurrent damage. The user must take precautions against exceeding their current limits when used as outputs.

[7]Portions are taken from example sketches found on the Arduino website and used under their Creative Commons Attribution-ShareAlike 3.0 license.

PART V ELECTRIC POWER AND MACHINES



Oleksandr Kostiuchenko/123RF

C H A P T E R
13

ELECTRIC POWER SYSTEMS

The basic concepts underlying simple AC power and the generation and distribution of electric power are extensions of those previously developed in Chapter 3, namely, phasors and impedance. Together, they pave the way for the material on electric machines in Chapters 14 to 16. The principal new concepts introduced in this chapter are average and complex power, and how they are computed for complex loads. The concept of the power factor is introduced as is the method for correcting (adjusting) it. A brief discussion of ideal transformers and maximum power transfer is provided, followed by an introduction to three-phase power, electrical safety, and finally a discussion of electric power generation and distribution.

In this chapter, quantities often involve angles. Unless indicated otherwise, angles are given in units of radians.

Learning Objectives

Students will learn to...

1. Understand the meaning of instantaneous and average power, use AC power notation, compute average power, and compute the power factor of a complex load. *Section 13.1*.

2. Use complex power notation; compute apparent, real, and reactive power for complex loads; and draw a power triangle. *Section 13.2*.
3. Compute the capacitance required to correct the power factor of a complex load *Section 13.3*.
4. Analyze an ideal transformer; compute primary and secondary currents, voltages, and turns ratios; calculate reflected sources and impedances across ideal transformers; and understand maximum power transfer. *Section 13.4*.
5. Use three-phase AC power notation; and compute load currents and voltages for balanced wye and delta loads. *Section 13.5*.
6. Understand the basic principles of residential electrical wiring and of electrical safety. *Sections 13.6 and 13.7*.

13.1 INSTANTANEOUS AND AVERAGE POWER

When a linear electric circuit is excited by a sinusoidal source, all voltages and currents in the circuit are also sinusoids of the same frequency as the source. Figure 13.1 depicts the general form of a linear AC circuit. The most general expressions for the voltage and current delivered to an arbitrary load are as follows:

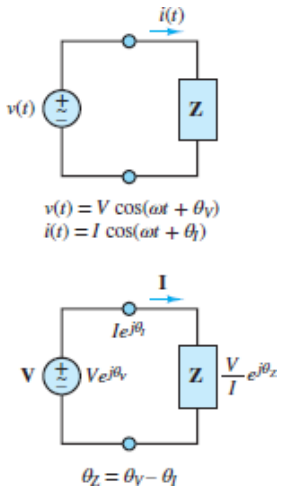


Figure 13.1 Time and frequency domain representations of an AC circuit. The phase angle of the load is $\theta_Z = \theta_V - \theta_I$.

$$\begin{aligned} v(t) &= V\cos(\omega t + \theta_V) \\ i(t) &= I\cos(\omega t + \theta_I) \end{aligned} \tag{13.1}$$

where V and I are the peak amplitudes of the sinusoidal voltage and current, respectively, and θ_V and θ_t are their phase angles. Two such waveforms are plotted in

Figure 13.2, with unit amplitude, angular frequency 150 rad/s, and phase angles $\theta_V = 0$ and $\theta_I = \pi/3$. Notice that the current *leads* the voltage; or equivalently, the voltage *lags* the current. Keep in mind that all phase angles are relative to some reference, which is usually chosen to be the phase angle of a source. The reference phase angle is freely chosen and therefore usually set to zero for simplicity. Also keep in mind that a phase angle represents a *time delay* of one sinusoid relative to its reference sinusoid.

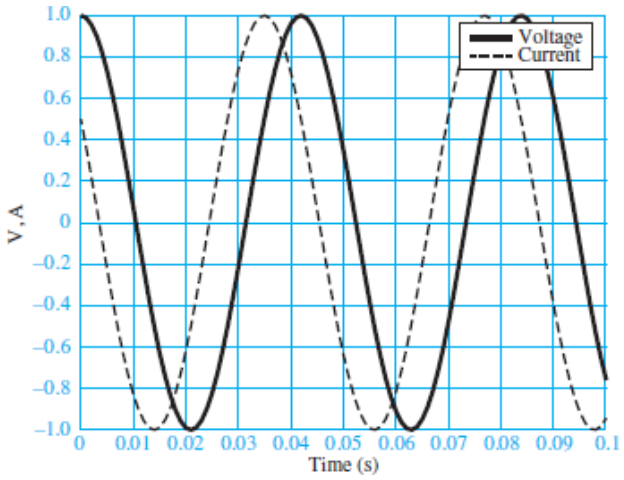


Figure 13.2 Current and voltage waveforms with unit amplitude and a phase shift of 60°

The **instantaneous power** dissipated by any element is the product of its instantaneous voltage and current.

$$p(t) = v(t)i(t) = VI\cos(\omega t + \theta_V)\cos(\omega t + \theta_I) \tag{13.2}$$

This expression is further simplified with the aid of the trigonometric identity:

$$2\cos(x)\cos(y) = \cos(x + y) + \cos(x - y) \tag{13.3}$$

Let $x = \omega t + \theta_V$ and $y = \omega t + \theta_I$ to yield:

$$\begin{aligned} p(t) &= \frac{VI}{2}[\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_V - \theta_I)] \\ &= \frac{VI}{2}[\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_Z)] \end{aligned} \tag{13.4}$$

Page 743

Equation 13.4 illustrates that the total instantaneous power dissipated by an element is equal to the sum of a constant $\frac{1}{2}VI\cos(\theta_Z)$ and a sinusoidal $\frac{1}{2}VI\cos(2\omega t + \theta_V + \theta_I)$, which

oscillates at twice the frequency of the source. Since the time average of a sinusoid is zero over one period or over a sufficiently long interval, the constant $\frac{1}{2}VI\cos(\theta_Z)$ is the time averaged power dissipated by a complex load **Z**, where θ_Z is the phase angle of that load.

Figure 13.3 shows the instantaneous and average power corresponding to the voltage and current signals of Figure 13.2. These observations can be confirmed mathematically by noting that the time average of the instantaneous power is defined by:

$$P_{avg} \equiv \frac{1}{T}\int_{t_0}^{t_0+T} p(t)\,dt \tag{13.5}$$

where T is one period of $p(t)$. Use equation 13.4 to substitute for $p(t)$ and yield:

$$\begin{aligned} P_{avg} &= \frac{1}{T}\int_{t_0}^{t_0+T} \frac{VI}{2}[\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_Z)]\,dt \\ &= \frac{VI}{2T}\int_{t_0}^{t_0+T} [\cos(2\omega t + \theta_V + \theta_I) + \cos(\theta_Z)]\,dt \end{aligned} \tag{13.6}$$

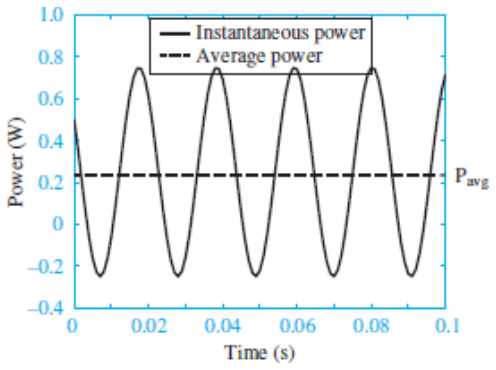


Figure 13.3 Instantaneous and average power corresponding to the signals in Figure 13.2

Page 744

The integral of the first part $\cos(2\omega t + \theta_V + \theta_I)$ is zero while the integral of the second part (a constant) is $T\cos(\theta_Z)$. Thus, the time averaged power Pavg is:

$$P_{avg} = \frac{VI}{2}\cos(\theta_Z) = \frac{1}{2}\frac{V^2}{|\mathbf{Z}|}\cos(\theta_Z) = \frac{1}{2}I^2|\mathbf{Z}|\cos(\theta_Z) \tag{13.7}$$

where

$$|\mathbf{Z}| = \frac{|\mathbf{V}|}{|\mathbf{I}|} = \frac{V}{I} \quad \text{and} \quad \theta_Z = \theta_V - \theta_I \tag{13.8}$$

Effective Values

In North America, AC power systems operate at a fixed frequency of 60 cycles per second, or hertz (Hz), which corresponds to an angular (radian) frequency ω given by:

$$\omega = 2\pi \cdot 60 = 377 \text{ rad/s} \quad \text{AC power frequency} \tag{13.9}$$

In Europe and most other parts of the world, the AC power frequency is 50 Hz.

Unless indicated otherwise, the angular (radian) frequency ω is assumed to be 377 rad/s throughout this chapter.

It is customary in AC power analysis to employ the *effective* or *root-mean-square* (rms) amplitude (see Section 3.3) rather than the peak amplitude for AC voltages and currents. In the case of a sinusoidal waveform, the effective voltage $\tilde{V} \equiv V_{\text{rms}}$ is related to the peak voltage V by:

$$\tilde{V} = V_{\text{rms}} = \frac{V}{\sqrt{2}} \tag{13.10}$$

Likewise, the effective current $\tilde{I} \equiv I_{\text{rms}}$ is related to the peak current I by:

$$\tilde{I} = I_{\text{rms}} = \frac{I}{\sqrt{2}} \tag{13.11}$$

The rms, or effective, value of an AC source is the DC value that produces the same average power to be dissipated by a common resistor.

The average power can be expressed in terms of effective voltage and current by plugging $V = \sqrt{2}\tilde{V}$ and $I = \sqrt{2}\tilde{I}$ into equation 13.7 to find:

$$P_{avg} = \tilde{V}\tilde{I}\cos(\theta_Z) = \frac{\tilde{V}^2}{|\mathbf{Z}|}\cos(\theta_Z) = \tilde{I}^2|\mathbf{Z}|\cos(\theta_Z) \tag{13.12}$$

Page 745

Voltage and current phasors are also represented with effective amplitudes by the notation:

$$\tilde{\mathbf{V}} = \tilde{V}e^{j\theta_V} = \tilde{V}\angle\theta_V \tag{13.13}$$

and

$$\tilde{\mathbf{I}} = \tilde{I}e^{j\theta_I} = \tilde{I}\angle\theta_I \tag{13.14}$$

It is critical to pay close attention to the *mathematical notation* that was first introduced in Chapter 3, namely that complex quantities, such as **V**, **I**, and **Z** are boldface. On the other hand, scalar quantities, such as $V, I, \tilde{V}$, and $\tilde{I}$ are italic. The relationship between these quantities is $V = |\mathbf{V}|$ and $\tilde{V} = |\tilde{\mathbf{V}}|$.

Impedance Triangle

Figure 13.4 illustrates the concept of the **impedance triangle**, which is an important graphical representation of impedance as a vector in the complex plane. Basic trigonometry yields:

$$R = |\mathbf{Z}|\cos\theta \tag{13.15}$$

$$X = |\mathbf{Z}|\sin\theta \tag{13.16}$$

where R is the *resistance* and X is the *reactance*. Notice that both R and P_{avg} are proportional to $\cos(\theta_Z)$, which suggests that a triangle similar to (i.e., the same shape as) the impedance triangle could be constructed with P_{avg} as one leg of a right triangle. In fact, such a triangle is known as a *power triangle*. The similarity of these two types of triangles is a powerful concept for problem solving, as is shown in Section 13.2.

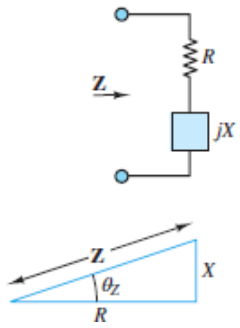


Figure 13.4 Impedance triangle

Power Factor

The phase angle θ_Z of the load impedance plays a very important role in AC power circuits. From equation 13.12, the average power dissipated by an AC load is proportional to $\cos(\theta_Z)$. For this reason, $\cos(\theta_Z)$ is known as the **power factor (pf)**. For purely resistive loads:

$$\theta_Z = 0 \quad \rightarrow \quad \text{pf} = 1 \qquad \text{Resistive load} \tag{13.17}$$

For purely inductive or capacitive loads:

$$\theta_Z = +\pi/2 \quad \rightarrow \quad \text{pf} = 0 \qquad \text{Inductive load} \tag{13.18}$$

$$\theta_Z = -\pi/2 \quad \rightarrow \quad \text{pf} = 0 \qquad \text{Capacitive load} \tag{13.19}$$

For loads with nonzero resistive (real) and reactive (imaginary) parts:

$$0 < |\theta_Z| < \pi/2 \quad \rightarrow \quad 0 < \text{pf} < 1 \qquad \text{Complex load} \tag{13.20}$$

Using the definition $\text{pf} = \cos(\theta_Z)$ the average power can be expressed as:

$$P_{\text{avg}} = \tilde{V}\tilde{I}\text{pf} \tag{13.21}$$

Thus, average power dissipated by a resistor is:

$$(P_{\text{avg}})_R = \tilde{V}_R \tilde{I}_R \text{pf}_R = \tilde{V}_R \tilde{I}_R \tag{13.22}$$

because $\text{pf}_R = 1$. By contrast, the average power dissipated by a capacitor or inductor is:

$$(P_{\text{avg}})_X = \tilde{V}_X \tilde{I}_X \text{pf}_X = 0 \tag{13.23}$$

because $pf_X = 0$, where the subscript X indicates a reactive element (i.e., either a capacitor or inductor). It is important to note that although capacitors and inductors are *lossless* (i.e., they store and release energy but do not dissipate energy), they do influence power dissipation in a circuit by affecting the voltage across and the current through resistors in the circuit.

When θ_Z is positive, the load is *inductive* and the power factor is said to be *lagging*; when θ_Z is negative, the load is *capacitive* and the power factor is said to be *leading*. It is important to keep in mind that $pf = \cos(\theta_Z) = \cos(-\theta_Z)$ because the cosine is an even function. Thus, while it may be important to know whether a load is inductive or capacitive, the value of the power factor only indicates the extent to which a load is inductive or capacitive. To know whether a load is inductive or capacitive, one must know whether the power factor is leading or lagging.

EXAMPLE 13.1 Computing Average and Instantaneous AC Power

Problem

Compute the average and instantaneous power dissipated by the load of Figure 13.5.

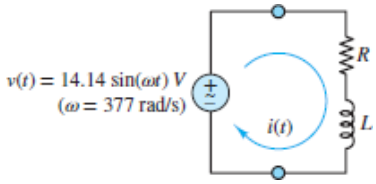


Figure 13.5

Solution

Known Quantities: Source voltage and frequency, load resistance and inductance values.

Find: P_{avg} and $p(t)$ for the RL load.

Schematics, Diagrams, Circuits, and Given Data: $v(t) = 14.14 \sin(377t)$V; $R = 4\Omega$; $L = 8$ mH.

Assumptions: None.

Analysis: The source voltage is expressed in terms of sin(377t). By convention, all time- domain sinusoids should be expressed as cosines. To convert sin(377t) to cos(377t + θ_V) recall that a sine equals a cosine shifted forward in time (to the right) by $\pi/2$ rad; that is, sin(377t) = cos(377t − $\pi/2$). Thus, at the angular frequency ω = 377 rad/s the source voltage is:

$$\tilde{\mathbf{V}} = 10\angle\left(-\frac{\pi}{2}\right) \text{ V rms}$$

where 14.14V = 10Vrms.

The equivalent impedance of the load is:

$$\mathbf{Z} = R + j\omega L \approx 4 + j3 = 5\angle(36.9°) = 5\angle(0.644 \text{ rad})\,\Omega$$

Page 747

The current in the loop is:

$$\tilde{\mathbf{I}} = \frac{\tilde{\mathbf{V}}}{\mathbf{Z}} \approx \frac{10\angle(-\pi/2)}{5\angle(0.644)} \approx 2\angle(-2.215) \text{ A rms}$$

It is instructive to compute the average power dissipated in the circuit in two ways:

1. The most straightforward and brute force approach is to compute:

$$P_{avg} = \tilde{V}\tilde{I}\cos(\theta_Z) = 10 \times 2 \times \cos(0.644) \approx 16 \text{ W}$$

2. Another approach is to realize that the average power dissipated by the inductor is zero. Thus, the total average power dissipated equals the average power dissipated by the resistor. Thus:

$$(P_{avg})_R = \tilde{I}^2 R\,\text{pf}_R = \tilde{I}^2 R \approx (2)^2 \times 4 = 16 \text{ W}$$

The instantaneous power is given by:

$$p(t) = v(t) \times i(t) = \sqrt{2} \times 10 \sin(377t) \times \sqrt{2} \times 2 \cos(377t - 2.215) \text{ W}$$

The instantaneous voltage and current waveforms and the instantaneous and average power are plotted in Figure 13.6.

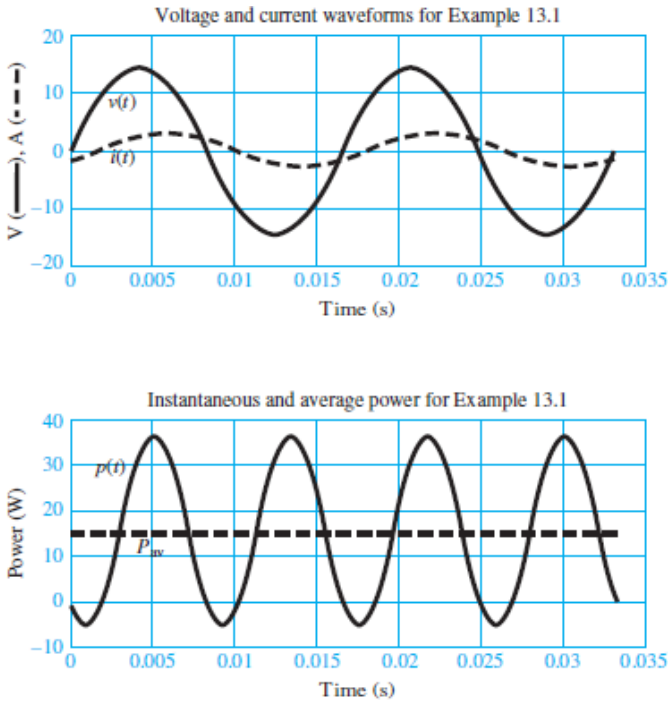


Figure 13.6 Voltage, current and power waveforms for Example 13.1.

Comment: It is standard procedure in electrical engineering practice to use rms values in power calculations. Also, note that the instantaneous power can be negative at times even though the average power is positive. This result reflects the fact that although the average power of an inductor is identically zero, the instantaneous power of an inductor can be positive or negative as the inductor charges or discharges with the sinusoidal source.

Page 748

EXAMPLE 13.2 Computing Average AC Power

Problem

Compute the average power dissipated by the load of Figure 13.7.

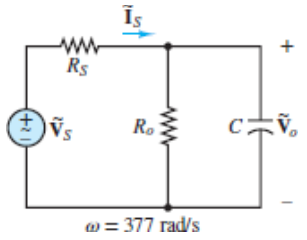


Figure 13.7 Circuit for Example 13.2.

Solution

Known Quantities: Source voltage, internal resistance, load resistance, capacitance, and frequency.

Find: *P*avg for the R_o | | C_o load.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 110\angle 0°$ V rms; $R_S = 2\ \Omega$; $R_o = 16\ \Omega$; $C = 100\ \mu$F; $\omega = 377$ rad/s.

Assumptions: None.

Analysis: First, compute the impedance of the load at the angular frequency $\omega = 377$ rad/s:

$$\mathbf{Z}_o = R_o \| \frac{1}{j\omega C} = \frac{R_o}{1 + j\omega C R_o} = \frac{16}{1 + j0.6032} = 13.7\angle(-0.543)\,\Omega$$

where the angle is given in radians. Next, apply voltage division to compute the load voltage:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{13.7\angle(-0.543)}{2 + 13.7\angle(-0.543)}110\angle 0 = 97.6\angle(-0.067)\text{ V rms}$$

Finally, compute the average power using equation 13.12:

$$P_{avg} = \frac{|\tilde{\mathbf{V}}_o|^2}{|\mathbf{Z}_o|}\cos(\theta_Z) = \frac{97.6^2}{13.7}\cos(-0.543) = 595\text{ W}$$

Alternatively, compute the source current $\tilde{\mathbf{I}}_S$ and then use equation 13.12 to compute the average power:

$$\tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = 7.12\angle 0.476\text{ A rms}$$

$$P_{avg} = |\tilde{\mathbf{I}}_S|^2|\mathbf{Z}_o|\cos(\theta) = 7.12^2 \times 13.7 \times \cos(-0.543) = 595\text{ W}$$

EXAMPLE 13.3 Computing Average AC Power

Problem

Compute the average power dissipated by the load of Figure 13.8.

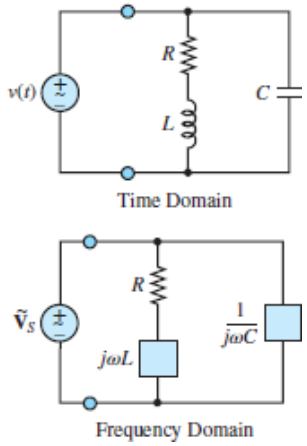


Figure 13.8

Solution

Known Quantities: Source voltage, internal resistance, load resistance, capacitance and inductance values, and frequency.

Find: P_{avg} for the complex load.

Page 749

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_s = 110\angle 0$ V rms; $R = 10\ \Omega$; $L = 0.05$ H; $C = 470\ \mu$F; $\omega = 377$ rad/s. Figure 13.8. Figure 13.8.

Assumptions: None.

Analysis: First, compute the impedance of the load $\mathbf{Z}_o$ at the angular frequency ω = 377 rad/s:

$$\mathbf{Z}_o = (R + j\omega L) \| \frac{1}{j\omega C} = \frac{(R + j\omega L)/j\omega C}{R + j\omega L + 1/j\omega C}$$

$$= \frac{R + j\omega L}{1 - \omega^2 LC + j\omega CR} = 1.16 - j7.18$$

$$= 7.27\angle(-1.41)\ \Omega$$

Note that the equivalent load impedance at ω = 377 rad/s has a negative imaginary part, which is a feature of a *capacitive load*, as shown in Figure 13.9. The average power is:

$$P_{avg} = \frac{|\tilde{\mathbf{V}}_s|^2}{|\mathbf{Z}_o|}\cos(\theta) = \frac{110^2}{7.27}\cos(-1.41) = 266\ \text{W}$$

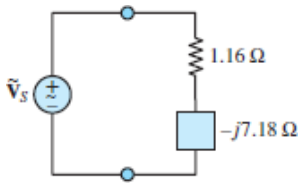


Figure 13.9

Comment: At ω = 377 rad/s, the capacitance has a larger impact on the total equivalent impedance than the inductance. At lower frequencies, where the impedance of the capacitor is large compared to $R + j\omega L$, the parallel equivalent impedance will be inductive. It is instructive to determine the frequencies when the parallel equivalent impedance has a zero imaginary part.

CHECK YOUR UNDERSTANDING

Consider the circuit shown in Figure 13.10. Find the impedance of the load "seen" by the voltage source, and compute the average power dissipated by the load. The constant 155.6 multiplying the cosine function is always the peak amplitude, not the rms amplitude.

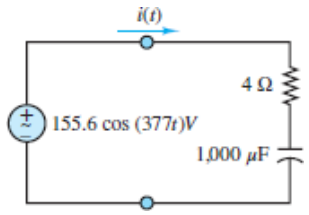


Figure 13.10

Answer: $\mathbf{Z} = 4.8\angle -33.5^{\circ}\ \Omega$; $P_{\text{avg}} = 2{,}103.4\ \text{W}$

CHECK YOUR UNDERSTANDING

For Example 13.2, compute the average power dissipated by the internal source resistance R_S.

Answer: 101.46 W; 595 W

13.2 COMPLEX POWER

The computation of AC power is simplified by defining a **complex power S**, where:

$$\mathbf{S} = \tilde{\mathbf{V}}\tilde{\mathbf{I}}^{*} \quad \text{Complex power} \tag{13.24}$$

Page 750

where the asterisk denotes the complex conjugate (see Appendix A). Note that the effect of taking the complex conjugate of a phasor is to multiply its phase angle by −1. In other words, $\angle\mathbf{S} = \angle\tilde{\mathbf{V}} + \angle\tilde{\mathbf{I}}^{*} = \angle\tilde{\mathbf{V}} - \angle\tilde{\mathbf{I}} = \theta_Z$. The definition of complex power leads to:

$$\begin{aligned}\mathbf{S} &= \tilde{V}\tilde{I}\cos(\theta_Z) + j\tilde{V}\tilde{I}\sin(\theta_Z)\\ &= \tilde{I}^2|\mathbf{Z}|\cos(\theta_Z) + j\tilde{I}^2|\mathbf{Z}|\sin(\theta_Z)\\ &= \tilde{I}^2 R + j\tilde{I}^2 X = \tilde{I}^2\mathbf{Z}\end{aligned} \tag{13.25}$$

where $R = |\mathbf{Z}|\cos(\theta_Z)$ and $X = |\mathbf{Z}|\sin(\theta_Z)$ are the resistance and reactance of the impedance triangle shown in Figure 13.11. The real and imaginary parts of **S** are the **real power** $P_{\text{avg}} = \tilde{V}\tilde{I}\cos(\theta_Z)$ and the **reactive power** $Q = \tilde{V}\tilde{I}\sin(\theta_Z)$, respectively, such that:

$$\mathbf{S} = P_{avg} + jQ \tag{13.26}$$

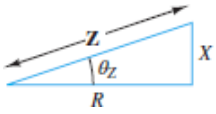
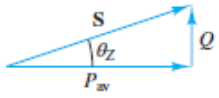


Figure 13.11 The impedance triangle

The magnitude $|\mathbf{S}|$ of the complex power is called the **apparent power** S and is measured in units of **volt-amperes (VA)**. The units of Q are **volt-amperes reactive**, or **VAR**.

The relationship between S, P, and Q is summarized by a *power triangle* as shown in Figure 13.12. It is important to note that the impedance and power triangles are similar; that is, they have the same shape. This result is helpful in problem solving. Table 13.1 shows the general expressions for calculating P and Q.



$|\mathbf{S}| = \sqrt{P_{av}^2 + Q^2} = \tilde{V} \cdot \tilde{I}$

$P_{av} = \tilde{V}\tilde{I} \cos\theta$

$Q = \tilde{V}\tilde{I} \sin\theta$

Figure 13.12 The complex power triangle

Table 13.1 Real and reactive power

Real power P_{avg}	Reactive power Q
$\tilde{V}\tilde{I}\cos(\theta)$	$\tilde{V}\tilde{I}\sin(\theta)$
$\tilde{I}^2 R$	$\tilde{I}^2 X$

The complex power can also be expressed as:

$$\mathbf{S} = \tilde{I}^2\mathbf{Z} = \tilde{I}^2 R + j\tilde{I}^2 X \tag{13.27}$$

Furthermore, since $\tilde{V} = \tilde{I}|\mathbf{Z}|$ and $|\mathbf{Z}|^2 = \mathbf{Z}\mathbf{Z}^*$, the complex power can be re- expressed as:

$$\mathbf{S} = \tilde{I}^2\mathbf{Z} = \frac{\tilde{I}^2\mathbf{Z}\mathbf{Z}^*}{\mathbf{Z}^*} = \frac{\tilde{V}^2}{\mathbf{Z}^*} \tag{13.28}$$

As previously stated, capacitors and inductors (reactive loads) do not dissipate energy themselves; they are lossless elements. However, they do influence power dissipation in a circuit by affecting the voltage across and current through resistors, which do dissipate energy. This influence is now quantified by the reactive power, Q, which is due entirely to capacitance and inductance in a circuit. It is worth noting that $Q = 0$, pf = 1, and therefore $P = S$ in purely resistive networks. It is also important to realize that P represents the real work done (per unit time) by a circuit. For example, the real power P of an electric motor represents the work done (per unit time) by the motor to perform some useful task. From the perspective of the utility company that provides the electric power for the motor and of the owner of the motor who has to pay the utility bill, it would be best if all the apparent power S provided by the utility company was converted to useful power P. (Why?) However, all electric motors have some inductance (e.g., coils of wire) such that $Q \neq 0$, pf < 1, and $P < S$. It is possible to *correct* the effect of a motor's inductance by adding capacitance in parallel with the motor so as to decrease Q and thereby decrease the apparent power S that must be provided for a given P required by the task.

Page 751

FOCUS ON PROBLEM SOLVING

COMPLEX POWER COMPUTATION

1. Use AC circuit analysis methods to compute (as phasors) the voltage across current through the load. Convert peak amplitudes to effective (rms) values.

$$\tilde{\mathbf{V}} = \tilde{V}\angle\theta_V \quad \text{and} \quad \tilde{\mathbf{I}} = \tilde{I}\angle\theta_I$$

2. Compute $\theta_Z = \theta_V - \theta_I$ and the power factor pf = $\cos(\theta_Z)$. Draw the imped triangle, as shown in Figure 13.11.
3. Use one of the two following methods to compute P_{avg} and Q.

- Compute the complex power $\mathbf{S} = \tilde{\mathbf{V}}\tilde{\mathbf{I}}^*$ such that $P = P_{avg} = \text{Re}(\mathbf{S})$, $Q = \text{Im}(\mathbf{S})$, $S = |\mathbf{S}|$. The effect of taking the complex conjugate of a phasor is to mult its phase angle by −1, such that $\angle\mathbf{S} = \angle\tilde{\mathbf{V}} - \angle\tilde{\mathbf{I}} = \theta_Z$.

- Compute the apparent power $s = |\mathbf{S}| = \tilde{V}\tilde{I}$ such that $P = P_{avg} = S_{pf}$ and Q $\sin(\theta_Z)$.

4. Draw the power triangle, as shown in Figure 13.12, and confirm that $S^2 =$ Q^2 and that $\tan(\theta_Z) = Q/P$.
5. If Q is negative, the load is capacitive and the power factor is *leading*; if positive, the load is inductive and the power factor is *lagging*.

EXAMPLE 13.4 Complex Power Calculations

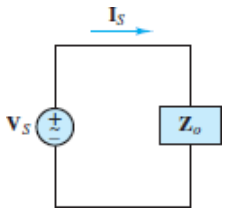


Figure 13.13 Circuit for Example 13.4.

Problem

Compute the complex power for the load $\mathbf{Z}o$ of Figure 13.13.

Solution

Known Quantities: Source, load voltage, and current.

Find: $\mathbf{S} = P_{avg} + jQ$ for the complex load.

Schematics, Diagrams, Circuits, and Given Data: $v(t) = 100\cos(\omega t + 0.262)$ V; $i(t) = 2\cos(\omega t - 0.262)$ A; $\omega = 377$ rad/s.

Assumptions: All angles are given in units of radians unless indicated otherwise.

Analysis: First, realize that the constants multiplying the cosine functions are always peak, not rms, values. These functions can be converted to phasor quantities with rms amplitudes as follows:

$$\bar{\mathbf{V}} = \frac{100}{\sqrt{2}}\angle 0.262 \text{ V} \qquad \bar{\mathbf{I}} = \frac{2}{\sqrt{2}}\angle(-0.262) \text{ A}$$

Page 752

Compute the phase angle of the load, and the real and reactive power, using the definitions of equation 13.12:

$$\theta_Z = \angle(\bar{\mathbf{V}}) - \angle(\bar{\mathbf{I}}) = 0.524 \text{ rad}$$
$$P_{avg} = |\bar{\mathbf{V}}||\bar{\mathbf{I}}| \cos(\theta_Z) = \frac{200}{2} \cos(0.524) = 86.6 \text{ W}$$
$$Q = |\bar{\mathbf{V}}||\bar{\mathbf{I}}| \sin(\theta_Z) = \frac{200}{2} \sin(0.524) = 50 \text{ VAR}$$

Apply the definition of complex power (equation 13.24) to repeat the same calculation:

$$\mathbf{S} = \bar{\mathbf{V}}\bar{\mathbf{I}}^* = \frac{100}{\sqrt{2}}\angle 0.262 \times \frac{2}{\sqrt{2}}\angle -(-0.262) = 100\angle 0.524$$
$$= (86.6 + j50) \text{ VA}$$

Therefore

$$P_{avg} = 86.6 \text{ W} \qquad Q = 50 \text{ VAR}$$

Comments: Note how the definition of complex power yields both quantities at one time.

EXAMPLE 13.5 Real and Reactive Power Calculations

Problem

Compute the complex power for the load of Figure 13.14.

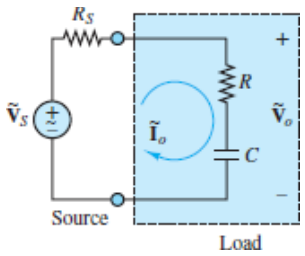


Figure 13.14 Circuit for Example 13.5.

Solution

Known Quantities: Source voltage and resistance; load impedance.

Find: $\mathbf{S} = P + jQ$ for the complex load.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 110\angle 0°$ V; $R_S = 2\ \Omega$; $R = 5\ \Omega$; $C = 2{,}000\ \mu$F; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis: The load impedance is:

$$\mathbf{Z}_o = R + \frac{1}{j\omega C} = (5 - j1.326)\Omega = 5.173\angle(-0.259)\Omega$$

Next, apply voltage division and the generalized Ohm's law to compute the load voltage and current:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{5 - j1.326}{7 - j1.326} \times 110 = 79.86\angle(-0.072)\ \text{V}$$

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{79.86\angle(-0.072)}{5.173\angle(-0.259)} = 15.44\angle 0.187\ \text{A}$$

Finally, compute the complex power, as defined in equation 13.24:

$$\mathbf{S} = \tilde{\mathbf{V}}_o\tilde{\mathbf{I}}_o^* = 79.9\angle(-0.072) \times 15.44\angle(-0.187) = 1{,}233\angle(-0.259)$$
$$= (1{,}192 - j316)\text{VA}$$

Page 753

Therefore:

$$P = 1{,}192\ \text{W} \qquad Q = -316\ \text{VAR}$$

Comment: Is the reactive power capacitive or inductive? Since $Q < 0$, the reactive power is capacitive!

EXAMPLE 13.6 Real Power Transfer for Complex Loads

Problem

Compute the complex power for the load between terminals *a* and *b* of Figure 13.15. Repeat the computation with the inductor removed from the load, and compare the real power for the two cases.

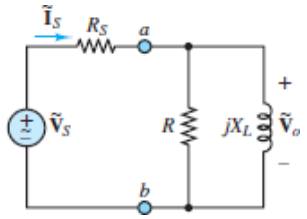


Figure 13.15 Circuit for Example 13.6.

Solution

Known Quantities: Source voltage and resistance; load impedance.

Find:

1. $\mathbf{S}_1 = P_1 + jQ_1$ for the complex load.
2. $\mathbf{S}_2 = P_2 + jQ_2$ for the real load.
3. For each case, compute the ratio of the real power dissipated by the load to the overall real power dissipated by the circuit.

Schematics, Diagrams, Circuits, and Given Data:

$\tilde{\mathbf{V}}_S = 110\angle 0°$ V; $R_S = 4\ \Omega$; $R = 10\ \Omega$; $jX_L = j6\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. With the inductor included in the load, its impedance $\mathbf{Z}_o$ is:

$$\mathbf{Z}_o = R \| j\omega L = \frac{10 \times j6}{10 + j6} = 5.145\angle 1.03\ \Omega$$

Apply voltage division to compute the load voltage $\tilde{\mathbf{V}}_o$ and the generalized Ohm's law to compute the current $\tilde{\mathbf{I}}_o = \tilde{\mathbf{I}}_S$.

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{5.145\angle 1.03}{4 + 5.145\angle 1.03} \times 110 = 70.9\angle 0.444 \text{ V}$$

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{70.9\angle 0.444}{5.145\angle 1.03} = 13.8\angle(-0.586)\text{A}$$

Finally, compute the complex power, as defined in equation 13.24:

$$\mathbf{S}_1 = \tilde{\mathbf{V}}_o\tilde{\mathbf{I}}_o^* = 70.9\angle 0.444 \times 13.8\angle 0.586 = 978\angle 1.03 \text{ VA}$$
$$= (503 + j839)\text{VA}$$

Therefore:

$$P_1 = 503 \text{ W} \qquad Q_1 = 839 \text{ VAR}$$

Page 754

2. With the inductor excluded from the load (Figure 13.16), its impedance is:

$$\mathbf{Z}_o = R = 10\ \Omega$$

Compute the load voltage and current:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{R_S + \mathbf{Z}_o}\tilde{\mathbf{V}}_S = \frac{10}{4 + 10} \times 110 = 78.57\angle 0 \text{ V}$$

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{78.57\angle 0}{10} = 7.857\angle 0 \text{ A}$$

Finally, compute the complex power, as defined in equation 13.24:

$$\mathbf{S}_2 = \tilde{\mathbf{V}}_o\tilde{\mathbf{I}}_o^* = 78.57\angle 0 \times 7.857\angle 0 = 617\angle 0 = (617 + j0)\text{VA}$$

Therefore:

$$P_2 = 617 \text{ W} \qquad Q_2 = 0 \text{ VAR}$$

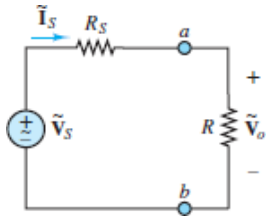


Figure 13.16 Circuit for Example 13.6 with inductor removed.

3. To compute the overall real power P_{total} dissipated by the circuit, it is necessary to include the impact of the line resistance R_S and compute for each case:

$$\mathbf{S}_{\text{total}} = \tilde{\mathbf{V}}_S \tilde{\mathbf{I}}_S^* = P_{\text{total}} + jQ_{\text{total}}$$

For case 1:

$$\tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_{\text{total}}} = \frac{\tilde{\mathbf{V}}_S}{R_S + \mathbf{Z}_o} = \frac{110}{4 + 5.145\angle 1.03} = 13.8\angle(-0.586)\,\text{A}$$

$$\mathbf{S}_{1_{\text{total}}} = \tilde{\mathbf{V}}_S \tilde{\mathbf{I}}_S^* = 110 \times 13.8\angle(+0.586) = (1{,}264 + j838)\text{VA} = P_{1_{\text{total}}} + jQ_{1_{\text{total}}}$$

The percent real power transfer is:

$$100 \times \frac{P_1}{P_{1_{\text{total}}}} = \frac{503}{1{,}264} = 39.8\%$$

For case 2:

$$\tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_{\text{total}}} = \frac{\tilde{\mathbf{V}}_S}{R_S + R} = \frac{110}{4 + 10} = 7.857\angle 0\ \text{A}$$

$$\mathbf{S}_{2_{\text{total}}} = \tilde{\mathbf{V}}_S \tilde{\mathbf{I}}_S^* = 110 \times 7.857 = (864 + j0)\text{VA} = P_{2_{\text{total}}} + jQ_{2_{\text{total}}}$$

The percent real power transfer is:

$$100 \times \frac{P_2}{P_{2_{\text{total}}}} = \frac{617}{864} = 71.4\%$$

Comments: If it were possible to eliminate the reactive part of the impedance, the percentage of real power transferred from the source to the load would be increased significantly. The procedure to accomplish this goal is called *power factor correction.*

EXAMPLE 13.7 Complex Power and Power Triangle

Problem

Find the reactive and real power for the load of Figure 13.17. Draw the associated power triangle.

Page 755

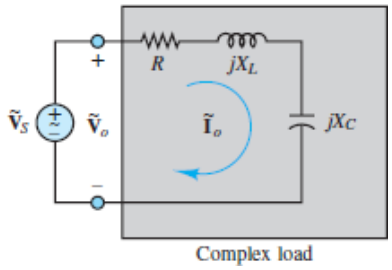


Figure 13.17 Circuit for Example 13.7.

Solution

Known Quantities: Source voltage; load impedance.

Find: $\mathbf{S} = P_{\text{avg}} + jQ$ for the complex load.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 60\angle 0$ V; $R = 3\ \Omega$; $jX_L = j9\ \Omega$; $jX_C = -j5\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians.

Analysis: First, compute the load current:

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{60\angle 0}{3 + j9 - j5} = \frac{60\angle 0}{5\angle 0.9273} = 12\angle(-0.9273)\,\text{A}$$

Next, compute the complex power, as defined in equation 13.24:

$$\mathbf{S} = \tilde{\mathbf{V}}_o \tilde{\mathbf{I}}_o^* = 60\angle 0 \times 12\angle 0.9273 = 720\angle 0.9273 = (432 + j576)\text{VA}$$

Therefore:

$$P = 432\ \text{W} \qquad Q = 576\ \text{VAR}$$

The total reactive power must be the sum of the reactive powers in each of the elements, such that $Q = Q_C + Q_L$. Compute these two quantities as follows:

$$Q_C = |\tilde{\mathbf{I}}_o|^2 \times X_C = (144)(-5) = -720\ \text{VAR}$$
$$Q_L = |\tilde{\mathbf{I}}_o|^2 \times X_L = (144)(9) = 1{,}296\ \text{VAR}$$

and

$$Q = Q_L + Q_C = 576 \text{ VAR}$$

Comments: The power triangle corresponding to this circuit is drawn in Figure 13.18. The vector diagram shows how the complex power **S** results from the vector addition of the three components P, Q_C, and Q_L.

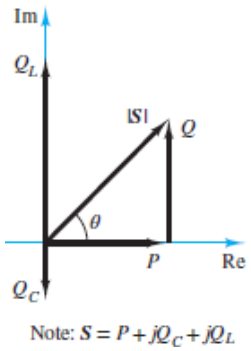


Figure 13.18 Power triangle for Example 13.7.

CHECK YOUR UNDERSTANDING

Compute the real and reactive power for the load of Example 13.2.

Answer: $P_{avg} = 595$ W; $Q = -359$ VAR

CHECK YOUR UNDERSTANDING

Compute the real and reactive power for the load of Figure 13.10.

Answer: $P_{avg} = 2.1$ kW; $Q = 1.39$ kVAR

CHECK YOUR UNDERSTANDING

Refer to Example 13.6, and compute the percent of real power transfer for the case where the inductance of the load is one-half of the original value.

Answer: 29.3%

CHECK YOUR UNDERSTANDING

Compute the power factor for the load of Example 13.7 with and without the inductor in the circuit.

Answer: pf = 0.6, lagging (with L in circuit); pf = 0.5145, leading (without L)

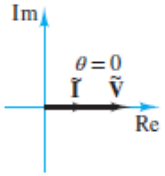
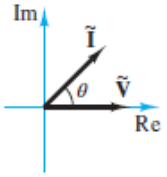
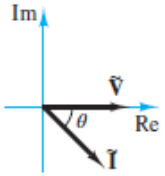
13.3 POWER FACTOR CORRECTION

A power factor close to unity signifies an efficient transfer of energy from the AC source to the load while a small power factor corresponds to inefficient use of energy, as illustrated in Example 13.6. If a load requires a given real power P, the current required by the load will be minimized when the power factor is maximized, that is, when pf = $\cos(\theta_Z) \rightarrow 1$. When pf < 1, it is possible to increase it (i.e., *correct* it) by adding, as appropriate, reactance (e.g., capacitance) to the load. When pf is leading, inductance must be added; when pf is lagging, capacitance must be added.

If $\theta_Z > 0$, then $Q > 0$, the load is inductive, the load current *lags* the load voltage, and the power factor pf is lagging. Alternatively, if $\theta_Z < 0$, then $Q < 0$, the load is capacitive, the load current *leads* the load voltage, and the power factor pf is leading.

Table 13.2 illustrates and summarizes these concepts. For simplicity, the phase angle of the voltage phasor $\tilde{\mathbf{V}}$ shown in the table is zero and acts as a reference angle for the current phasor.

Table 13.2 Important facts related to complex power

	Resistive load	**Capacitive load**	**Inductive load**
Ohm's law	$\tilde{\mathbf{V}} = \tilde{\mathbf{I}}\mathbf{Z}$	$\tilde{\mathbf{V}} = \tilde{\mathbf{I}}\mathbf{Z}$	$\tilde{\mathbf{V}} = \tilde{\mathbf{I}}\mathbf{Z}$
Complex impedance	$\mathbf{Z} = R$	$\mathbf{Z} = R + jX$ $X < 0$	$\mathbf{Z} = R + jX$ $X > 0$
Phase angle	$\theta = 0$	$\theta < 0$	$\theta > 0$
Complex plane sketch			
Explanation	The current is in phase with the voltage.	The current "leads" the voltage.	The current "lags" the voltage.
Power factor	Unity	Leading, < 1	Lagging, < 1
Reactive power	0	Negative	Positive

In practice, the load designed for a useful industrial task is often inductive because of the presence of electric motors. The power factor of an inductive load can be *corrected* by adding capacitance in parallel with the load. This procedure is called *power factor correction*.

The measurement and correction of the power factor for the load are an extremely important aspect of any industrial engineering application that requires the use of substantial quantities of electric power. In particular, industrial plants, construction sites, heavy machinery, and other heavy users of electric power must be aware of the power factor that their loads present to the electric utility company. As was already observed, a low power factor results in greater current draw from the electric utility and greater line losses. Thus, computations related to the power factor of complex loads are of great utility to any practicing engineer.

FOCUS ON PROBLEM SOLVING

POWER FACTOR CORRECTION

1. Follow the steps outlined in the Focus on Problem Solving box, "Com Power Computation," to find the initial phase angle of the load θ_{Z_i}, power fa pf_i, real power P_i, and reactive power Q_i. If both P_i and either pf or θ_Z are gi compute Q directly using $Q = P\tan(\theta_Z)$. An initial power triangle is helpfu visualizing this information.
2. For a lagging power factor, augment the load with a parallel capacitor such th

$$\Delta Q = Q_C = \frac{\tilde{\mathbf{V}}^2}{|\mathbf{Z}_C|}\sin(\theta_Z) = -\omega C\,\tilde{\mathbf{V}}^2$$

3. Express the final reactive power Q_f as:

$$Q_f = Q_i + \Delta Q$$

4. The real power is unchanged by the addition of the capacitor in parallel. Thu $= P_i$ and the final (corrected) phase angle of the augmented load is:

$$\theta_{Z_f} = \tan^{-1}\left(\frac{Q_f}{P_f}\right)$$

   It is helpful to draw a final power triangle to visualize the effect of the par capacitor.
5. The final corrected power factor is:

$$\text{pf}_f = \cos(\theta_{Z_f})$$

EXAMPLE 13.8 Power Factor Correction

Problem

Calculate the power factor for the circuit of Figure 13.19. Correct it to unity by adding a capacitor in parallel with the load.

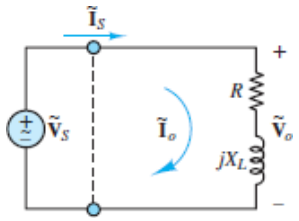


Figure 13.19 Circuit for Example 13.8.

Solution

Known Quantities: Source voltage; load impedance.

Find:

1. $\mathbf{S} = P + jQ$ for the complex load.
2. Value of parallel capacitance that results in pf = 1.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 117\angle 0$ V rms; $R = 50\ \Omega$; $jX_L = j86.7\ \Omega$; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. First, compute the load impedance:

$$\mathbf{Z}_o = R + jX_L = 50 + j86.7 = 100\angle 1.047\ \Omega$$

Next, compute the load current $\tilde{\mathbf{I}}_o = \tilde{\mathbf{I}}_S$:

$$\tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_o}{\mathbf{Z}_o} = \frac{117\angle 0}{50 + j86.7} = \frac{117\angle 0}{100\angle 1.047} = 1.17\angle(-1.047)\,\text{A}$$

The complex power, as defined in equation 13.24, is:

$$\mathbf{S} = \tilde{\mathbf{V}}_o \tilde{\mathbf{I}}_o^* = 117\angle 0 \times 1.17\angle 1.047 = 137\angle 1.047 = (68.4 + j118.5)\,\text{VA}$$

Therefore:

$$P = 68.4\ \text{W} \qquad Q = 118.5\ \text{VAR}$$

The power triangle corresponding to this circuit is drawn in Figure 13.20. The vector diagram shows how the complex power **S** results from the vector addition of the two components P and Q.

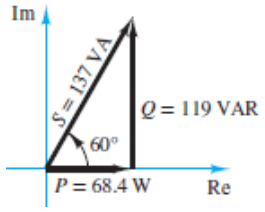


Figure 13.20 Power triangle for Example 13.8.

2. To correct the power factor to unity it is necessary to subtract 118.5 VAR. This goal can be accomplished by adding in parallel a capacitor with $Q_C = -118.5$ VAR. The required capacitance is found by:

$$X_C = \frac{|\tilde{\mathbf{V}}_o|^2}{Q_C} = -\frac{(117)^2}{118.5} = -115\ \Omega$$

The reactance X_C is related to the capacitance by:

$$jX_C = \frac{1}{j\omega C} = -\frac{j}{\omega C}$$

Thus, the result is:

$$C = -\frac{1}{\omega X_C} = -\frac{1}{377(-115)} = 23.1\,\mu\text{F}$$

Page 759

3. The total current required of the source is $\tilde{\mathbf{I}}_S = \tilde{\mathbf{I}}_o + \tilde{\mathbf{I}}_c$, where:

$$\tilde{\mathbf{I}}_c = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_c} = (j\omega C)(117\angle 0) = (377)(23.1\,\mu\text{F})(117)\angle(\pi/2) \approx 1.02\angle 90^\circ\ \text{A}$$

Notice that $|\tilde{\mathbf{I}}_c| = |\tilde{\mathbf{V}}_S|/|X_c| \approx 117/115 \approx 1.02$ A. The total current is computed by phasor addition to be:

$$\tilde{\mathbf{I}}_S \approx 1.17\angle(-1.047) + 1.02\angle(\pi/2) \approx 0.585\angle 0\ \text{A}$$

The corrected power factor pf = 1 implies that the impedance of the load is now purely real; that is, $\theta_Z = 0$. Thus, the source current must now be *in phase* with the source voltage; and it is.

Comments: Notice that the magnitude of the source current is reduced by increasing the power factor. The power factor correction, which is a very common procedure in electric power systems, is illustrated in Figure 13.21.

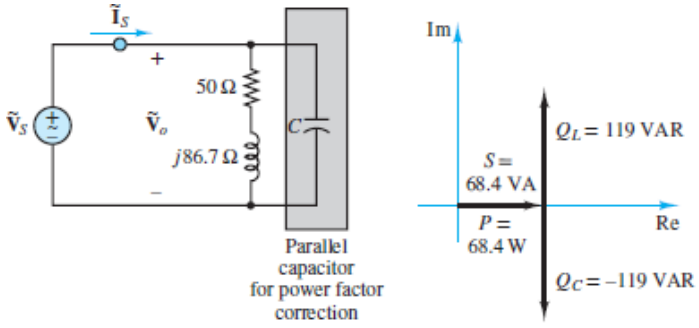


Figure 13.21 Power factor correction

EXAMPLE 13.9 Can a Series Capacitor Be Used for Power Factor Correction?

Problem

The circuit of Figure 13.22 suggests the use of a series capacitor for power factor correction. Why is this approach *not* a feasible alternative to the parallel capacitor approach demonstrated in Example 13.8?

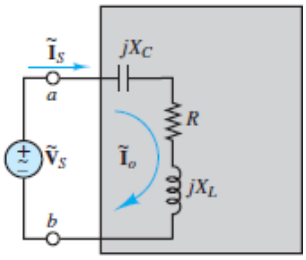


Figure 13.22 Circuit for Example 13.9.

Solution

Known Quantities: Source voltage; load impedance.

Find: Load (source) current.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 117\angle 0$ V; $R = 50\ \Omega$; $jX_L = j86.7\ \Omega$; $jX_C = -j86.7\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis: First, compute the impedance of the load between terminals *a* and *b*:

$$\mathbf{Z}_o = R + jX_L + jX_C = 50 + j86.7 - j86.7 = 50\ \Omega$$

Notice that the reactance of the capacitor was chosen so as to make the total load purely resistive. Thus, $\theta_Z = 0$ and the corrected power factor is pf = 1. So far, so good.

Next, compute the current through the series load:

$$\tilde{\mathbf{I}}_o = \tilde{\mathbf{I}}_S = \frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_o} = \frac{117\angle 0}{50} = 2.34\ \text{A}$$

The corrected power factor pf = 1 implies that the impedance of the load is now purely real; that is, $\theta_Z = 0$. Thus, the source current must now be *in phase* with the source voltage; and it is.

The problem with this approach to power factor correction is revealed by computing the initial current through the load, prior to the addition of the capacitor.

$$(\tilde{\mathbf{I}}_o)_{\text{initial}} = \frac{\tilde{\mathbf{V}}_S}{R + jX_L} = \frac{117\angle 0}{50 + j86.7} \approx 1.17\angle(-\pi/3)\,\text{A}$$

Comments: Notice the twofold increase in the source current as a result of the additional capacitor in series. Consequently, the power required by the source doubled as well. In practice, adding capacitance in parallel can be accomplished relatively easily with one large bank located somewhere on an industrial site and away from the production motors themselves. Electric utilities motivate industries to raise power factors by offering discounted rates ($/kWh).

EXAMPLE 13.10 Power Factor Correction

Problem

A capacitor is used to correct the power factor of the 100 kW and lagging pf = 0.7 load of Figure 13.23. Determine the reactive power of the load alone, and compute the capacitance required for a corrected power factor pf = 1.

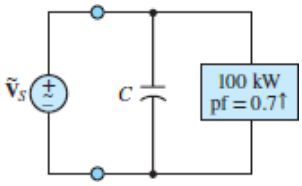


Figure 13.23 Circuit for Example 6.10.

Solution

Known Quantities: Source voltage; load power and power factor.

Find:

1. The reactive power Q of the load alone.
2. The capacitance C required for a corrected power factor pf = 1.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 480\angle 0$ V rms; $P = 10^5$ W; pf = 0.7 lagging for the load; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. For the load alone, pf = 0.7 lagging or $\cos(\theta_Z) = 7/10$, and the power triangle has the *shape* shown in Figure 13.24. The real power is given as $P = 100$ kW, so the reactive power of the load can be computed using the relative triangle dimensions to be:

$$Q = P\tan(\theta_Z) = (100\,\text{kW})(\sqrt{51}/7) = 102\ \text{kVAR}$$

Since the power factor is lagging, the reactive power is positive as indicated in Table 13.2 and shown in the power triangle of Figure 13.25.

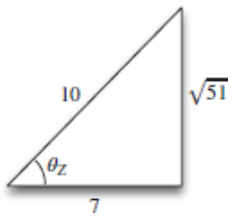


Figure 13.24 Relative dimensions of power triangle

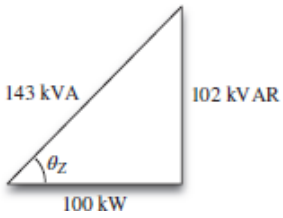


Figure 13.25 Power triangle

Page 761

2. To set the corrected power factor to pf = 1 the capacitance must contribute −102 kVAR of reactive power. That is:

$$Q_C = \Delta Q = Q_{\text{final}} - Q_{\text{initial}} = 0 - 102 \text{ kVAR} = -102 \text{ kVAR}$$

Since the voltage across capacitor $\tilde{\mathbf{V}}_C$ equals the source voltage $\tilde{\mathbf{V}}_S$, the reactive power of the capacitor is:

$$Q_C = \frac{|\tilde{\mathbf{V}}_C|^2}{|X_C|} \sin(-90°) = -(\omega C)|\tilde{\mathbf{V}}_S|^2 = -(377)(480^2)C$$

Thus, to correct the power factor to pf = 1 (zero total reactive power), the capacitor must satisfy:

$$Q_C = -(377)(480^2)C = -102 \text{ kVAR}$$

or

$$C = \frac{102 \text{ kVAR}}{(377)(480^2)} = 1{,}175 \ \mu\text{F}$$

Use trigonometry and/or the Pythagorean theorem to show that the apparent power $|S| = 143$ kVA, as indicated in Figure 13.25.

Comments: Note that it is not necessary to know the load impedance to perform power factor correction; however, it is a useful exercise to compute the equivalent

impedance seen by $\tilde{\mathbf{V}}_s$ and check that $\cos(\theta_Z) = 0.7$.

EXAMPLE 13.11 Power Factor Correction

Problem

Figure 13.26 shows a second load added to the circuit of Figure 13.23. Determine the capacitance required for an overall corrected power factor pf = 1. Draw the phasor diagram showing the relationship between $\tilde{\mathbf{I}}_C$, $\tilde{\mathbf{I}}_1$, and $\tilde{\mathbf{I}}_2$.

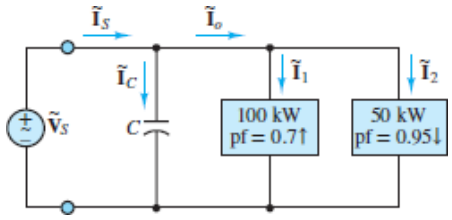


Figure 13.26 Circuit with two loads.

Solution

Known Quantities: Source voltage; load power and power factor.

Find:

1. The total reactive power of loads 1 and 2.
2. The capacitance C required for an overall power factor pf = 1.
3. $\tilde{\mathbf{I}}_C$, $\tilde{\mathbf{I}}_1$, and $\tilde{\mathbf{I}}_2$, and construct a phasor diagram of these currents.

Page 762

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 480\angle 0$ V rms; $P_1 = 100$ kW; $\text{pf}_1 = 0.7$ lagging; $P_2 = 50$ kW; $\text{pf}_2 = 0.95$ leading; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis:

1. Compute $\tilde{\mathbf{I}}_1$ and $\tilde{\mathbf{I}}_2$ using the relation $P = |\tilde{\mathbf{V}}||\tilde{\mathbf{I}}|\text{pf}$.

$$P_1 = |\tilde{\mathbf{V}}_S||\tilde{\mathbf{I}}_1|\cos(\theta_1) \quad \rightarrow \quad |\tilde{\mathbf{I}}_1| = \frac{P_1}{|\tilde{\mathbf{V}}_S|\cos(\theta_1)} \approx 298 \text{ A}$$

and

$$\angle\tilde{\mathbf{V}}_S = \angle\tilde{\mathbf{I}}_1 + \theta_{Z_1} \quad \rightarrow \quad \angle\tilde{\mathbf{I}}_1 = \angle\tilde{\mathbf{V}}_S - \theta_{Z_1} = 0 - \cos^{-1}(0.7) \approx -0.795 \text{ rad}$$

It is important to keep in mind that although inverse trigonometric functions are double-valued [e.g., $\cos^{-1}(0.7) \approx \pm 0.795$ rad], the power factor for load 1 is lagging such that $\theta_{Z_1} = +0.795$ rad is the correct choice.

Similarly, for load 2:

$$P_2 = |\tilde{\mathbf{V}}_S||\tilde{\mathbf{I}}_2|\cos(\theta_2) \quad \rightarrow \quad |\tilde{\mathbf{I}}_2| = \frac{P_2}{|\tilde{\mathbf{V}}_S|\cos(\theta_2)} \approx 110 \text{ A}$$

and

$$\angle\tilde{\mathbf{V}}_S = \angle\tilde{\mathbf{I}}_2 + \theta_{Z_2} \quad \rightarrow \quad \angle\tilde{\mathbf{I}}_2 = \angle\tilde{\mathbf{V}}_S - \theta_{Z_2} = 0 - \cos^{-1}(0.95) \approx +0.318 \text{ rad}$$

The power factor for load 2 is leading such that $\theta_{Z_2} = -0.318$ rad is the correct choice.

Now use the given data and the relation $Q = P\tan(\theta_Z)$ to compute the reactive power for each load.

$$Q_1 = P_1\tan(+0.795 \text{ rad}) \approx +102 \text{ kVAR}$$

and

$$Q_2 = P_2\tan(-0.318 \text{ rad}) \approx -16.4 \text{ kVAR}$$

The power triangles for the two loads are shown in Figures 13.27 and 13.28. The total reactive power is therefore $Q = Q_1 + Q_2 \approx 85.6$ kVAR.

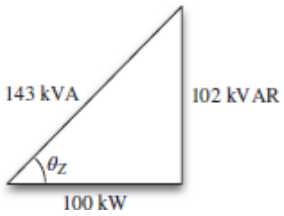


Figure 13.27 Power triangle for load 1

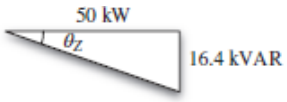


Figure 13.28 Power triangle for load 2

2. To set the corrected power factor to pf = 1 the capacitance must contribute −85.6 kVAR of reactive power. That is:

$$Q_C = \Delta Q = Q_{\text{final}} - Q_{\text{initial}} = 0 - 85.6 \text{ kVAR} = -85.6 \text{ kVAR}$$

For a capacitor alone its reactive power is:

$$Q_C = \frac{|\tilde{\mathbf{V}}_C|^2}{X_C} = -(\omega C)|\tilde{\mathbf{V}}_S|^2 = -(377)(480^2)C$$

Thus, to correct the power factor to pf = 1 (zero total reactive power), the capacitor must satisfy:

$$Q_C = -(377)(480^2)C = -85.6 \text{ kVAR}$$

or

$$C = \frac{85.6 \text{ kVAR}}{(377)(480^2)} \approx 985 \mu\text{F}$$

Page 763

3. To compute the capacitor current it is not possible to use $P = |\tilde{\mathbf{V}}||\tilde{\mathbf{I}}|\text{pf}$ because $P = 0$ and pf = 0 for a capacitor. Instead, the generalized Ohm's law provides an alternative approach.

$$\tilde{\mathbf{V}}_C = \tilde{\mathbf{I}}_C \mathbf{Z}_C \quad \rightarrow \quad |\tilde{\mathbf{I}}_C| = \frac{|\tilde{\mathbf{V}}_C|}{|\mathbf{Z}_C|} = \omega C|\tilde{\mathbf{V}}_C| \approx 178.2 \text{ A}$$

where $\tilde{\mathbf{V}}_C = \tilde{\mathbf{V}}_S$. The phase angle of $\tilde{\mathbf{I}}_C$ is:

$$\angle\tilde{\mathbf{I}}_C = \angle\tilde{\mathbf{V}}_C - \theta_{Z_C} = 0 - (-\pi/2) = +\pi/2 \text{ rad}$$

The current phasor diagram can now be drawn as shown in Figure 13.29.

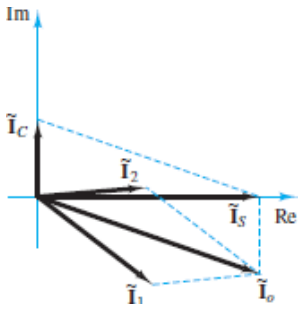


Figure 13.29

Comment: The power triangle suggests that the capacitor current can also be calculated using the relation $Q_C = |\tilde{\mathbf{V}}_C||\tilde{\mathbf{I}}_C| \sin(\theta_C)$, where $\theta_C = -\pi/2$ and $Q_C = |\tilde{\mathbf{V}}_C|^2/X_C = -(\omega C)|\tilde{\mathbf{V}}_C|^2$. Try it!

CHECK YOUR UNDERSTANDING

Two cases of the voltage across and the current through a load are given below. Determine the power factor of the load, and whether it is leading or lagging, for each case.

a. $v(t) = 540 \cos(\omega t + 15°)$ V, $i(t) = 2 \cos(\omega t + 47°)$ A
b. $v(t) = 155 \cos(\omega t - 15°)$ V, $i(t) = 2 \cos(\omega t - 22°)$ A

Answer: a. 0.848, leading; b. 0.9925, lagging

CHECK YOUR UNDERSTANDING

Determine if a load is capacitive or inductive, given the following facts:

a. pf = 0.87, leading
b. pf = 0.42, leading
c. $v(t) = 42 \cos(\omega t)$ V, $i(t) = 4.2 \sin(\omega t)$ A [*Hint:* $\sin(\omega t)$ lags $\cos(\omega t)$.]
d. $v(t) = 10.4 \cos(\omega t - 22°)$ V, $i(t) = 0.4 \cos(\omega t - 22°)$ A

Answer: a. Capacitive; b. capacitive; c. inductive; d. neither (resistive)

CHECK YOUR UNDERSTANDING

Compute the power factor for an inductive load with L = 100 mH in series with R = 0.4Ω. Assume ω = 377 rad/s.

Answer: pf = 0.0106, lagging

Page 764

FOCUS ON MEASUREMENTS

The Wattmeter

The instrument used to measure power is called a wattmeter. The external part of a wattmeter consists of four connections and a metering mechanism that displays the amount of real power dissipated by a circuit. The external and internal appearance of a wattmeter is depicted in Figure 13.30. Inside the wattmeter are two coils: a current-sensing coil and a voltage-sensing coil. In this example, we assume for simplicity that the impedance of the current-sensing coil $\mathbf{Z}_I$ is zero and that the impedance of

the voltage-sensing coil $\mathbf{Z}_V$ is infinite. In practice, this will not necessarily be true; some correction mechanism will be required to account for the impedance of the sensing coils.

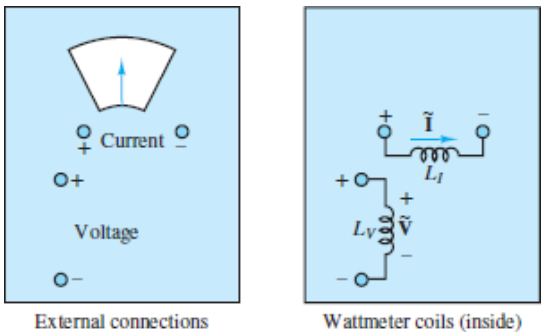


Figure 13.30 Wattmeter: external connections and internal layout.

A wattmeter should be connected as shown in Figure 13.31 to provide both current and voltage measurements. We see that the current-sensing coil is placed in series with the load and that the voltage-sensing coil is placed in parallel with the load. In this manner, the wattmeter is seeing the current through and the voltage across the load. Remember that the power dissipated by a circuit element is related to these two quantities. The wattmeter, then, is constructed to provide a readout of the real power absorbed by the load: $P = \text{Re}(S) = \text{Re}(\tilde{\mathbf{V}}\tilde{\mathbf{I}}^*)$.

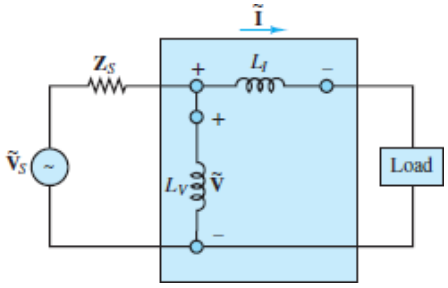


Figure 13.31 Wattmeter connection.

Problem:

1. For the circuit shown in Figure 13.32, show the connections of a wattmeter between the ideal voltage source and the load and find the power dissipated by the load.

2. Show the connections that will determine the power dissipated by R_2. What should the meter read?

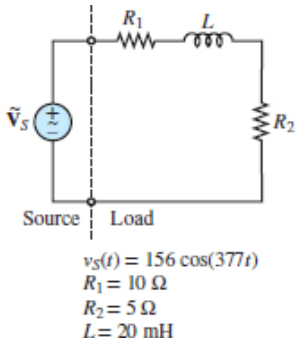


Figure 13.32 Wattmeter: example of power calculation.

Solution:

1. To measure the power dissipated by the load, we must know the current through and the voltage across the entire load circuit. This means that the wattmeter must be connected as shown in Figure 13.33. The wattmeter should read

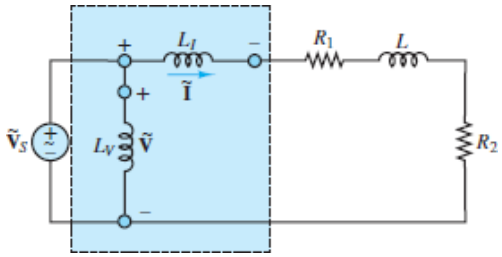


Figure 13.33 Wattmeter: example of power calculation.

$$P = \text{Re}\left[\tilde{\mathbf{V}}_S\tilde{\mathbf{I}}^*\right] = \text{Re}\left[\left(\frac{156}{\sqrt{2}}\angle 0\right)\left(\frac{(156/\sqrt{2})\angle 0}{R_1 + R_2 + j\omega L}\right)^*\right]$$

$$= \text{Re}\left[110\angle 0^\circ\left(\frac{110\angle 0}{15 + j7.54}\right)^*\right]$$

$$= \text{Re}\left[110\angle 0^\circ\left(\frac{110\angle 0}{16.79\angle 0.466}\right)^*\right] = \text{Re}\left[\frac{110^2}{16.79\angle(-0.466)}\right]$$

$$= \text{Re}\,[720.67\angle 0.466]$$

$$= 643.88\ \text{W}$$

2. To measure the power dissipated by R_2 alone, we must measure the current through R_2 and the voltage across R_2 *alone*. The connection is shown in Figure 13.34. The meter will read

$$P = |\tilde{\mathbf{I}}^2|R_2 = \left[\frac{110}{(15^2 + 7.54^2)^{1/2}}\right]^2 \times 5 = \frac{110^2}{15^2 + 7.54^2} \times 5$$
$$= 215 \text{ W}$$

Page 766

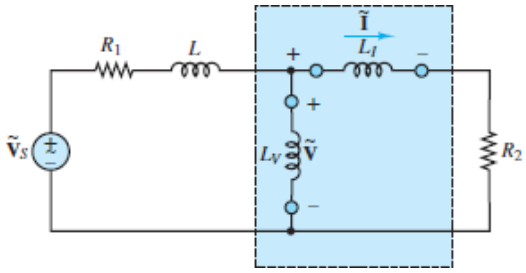


Figure 13.34 Circuit with wattmeter inserted to measure only the power dissipated by R_2

FOCUS ON MEASUREMENTS

Power Factor

Problem:

A capacitor is being used to correct the power factor of a load to unity, as shown in Figure 13.35. The capacitor value is varied, and measurements of the total current are

taken. Explain how it is possible to zero in on the capacitance value necessary to bring the power factor to unity just by monitoring the current $\tilde{\mathbf{I}}_S$.

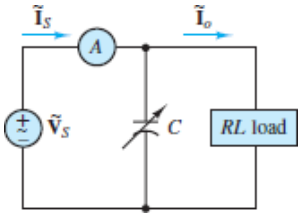


Figure 13.35 Circuit for illustration of power factor correction.

Solution:

The current through the load is

$$\tilde{\mathbf{I}}_o = \frac{\tilde{V}_S \angle 0^\circ}{R + j\omega L} = \frac{\tilde{V}_S}{R^2 + \omega^2 L^2}(R - j\omega L)$$

$$= \frac{\tilde{V}_S R}{R^2 + \omega^2 L^2} - j\frac{\tilde{V}_S\, \omega L}{R^2 + \omega^2 L^2}$$

The current through the capacitor is

$$\tilde{\mathbf{I}}_C = \frac{\tilde{V}_S \angle 0^\circ}{1/j\omega C} = j\tilde{V}_S \omega C$$

Page 767

The source current to be measured is

$$\tilde{\mathbf{I}}_S = \tilde{\mathbf{I}}_o + \tilde{\mathbf{I}}_C = \frac{\tilde{V}_S R}{R^2 + \omega^2 L^2} + j\left(\tilde{V}_S \omega C - \frac{\tilde{V}_S \omega L}{R^2 + \omega^2 L^2}\right)$$

The magnitude of the source current is

$$\tilde{I}_S = \sqrt{\left(\frac{\tilde{V}_S R}{R^2 + \omega^2 L^2}\right)^2 + \left(\tilde{V}_S \omega C - \frac{\tilde{V}_S \omega L}{R^2 + \omega^2 L^2}\right)^2}$$

We know that when the load is a pure resistance, then the current and voltage are in phase, the power factor is 1, and all the power delivered by the source is dissipated by the load as real power. This corresponds to equating the imaginary part of the expression for the source current to zero or, equivalently, to the following expression:

$$\frac{\tilde{V}_S \omega L}{R^2 + \omega^2 L^2} = \tilde{V}_S \omega C$$

in the expression for $|\tilde{I}_s|$. Thus, the magnitude of the source current is actually a minimum when the power factor is unity! It is therefore possible to "tune" a load to a unity pf by observing the readout of the ammeter while changing the value of the capacitor and selecting the capacitor value that corresponds to the lowest source current value.

13.4 TRANSFORMERS

Two separate AC circuits are often interfaced by a **transformer**, which acts as a magnetic coupling and *transforms* the voltage and current at the interface (e.g., by matching the high-voltage, low-current output of one circuit to the low-voltage, high-current input required by the other). Transformers play a major role in electric power engineering and are a necessary part of the electric power distribution network. The objective of this section is to introduce the ideal transformer and the concepts of impedance reflection and impedance matching.

The Ideal Transformer

The ideal transformer consists of two coils coupled to each other by a magnetic medium. There is no conducting electrical connection between the coils. The input side of a transformer is known as the **primary** while the output side is known as the **secondary**. The number of turns in the primary and secondary coils are designated n_1 and n_2, respectively. The **turns ratio** N is defined by:

$$N = \frac{n_2}{n_1} \tag{13.29}$$

Figure 13.36 illustrates the convention by which voltages and currents are usually assigned at a transformer. The solid black dots in Figure 13.36 are used to mark coil terminals that have the same polarity.

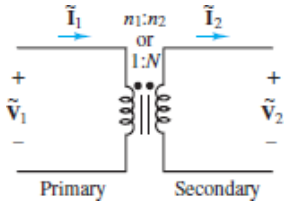


Figure 13.36 Ideal transformer

Page 768

Recall from Faraday's law that each coil experiences *self-induction* in that a time-varying current through a coil produces a time-varying magnetic flux through the coil

itself, which, in turn, induces a potential difference opposing the time-varying magnetic flux. The net effect of this self-induction is expressed by the inductance L of a coil. However, when two coils are present, as in a transformer, both coils also experience *mutual induction* in that some of the time-varying magnetic flux due to one coil passes through the other coil and induces another opposing potential difference. The net effect of the mutual induction is expressed by the mutual inductance M of the two coils. Both L and M contribute to the behavior of a transformer.

Notice the emphasis on *time variations* in the previous paragraph. One result of Faraday's law is that a leave in current through a coil, which generates a constant magnetic field, induces no opposing reaction within the coil itself (no self-induction) nor within any nearby coil (no mutual induction). Instead, a coil acts as a short-circuit in the presence of leave in current and transformers perform no useful function in DC circuits. See Chapter 14 for further discussion of Faraday's law as it relates to electromechanics.

As depicted in Figure 13.36, the relationships between primary and secondary currents and voltages in an ideal transformer are:

$$\begin{aligned} \tilde{\mathbf{V}}_2 &= N\tilde{\mathbf{V}}_1 \\ \tilde{\mathbf{I}}_2 &= \frac{\tilde{\mathbf{I}}_1}{N} \end{aligned} \qquad \text{Ideal transformer} \tag{13.30}$$

When $N > 1$, $|\tilde{\mathbf{V}}_2| > |\tilde{\mathbf{V}}_1|$ and a transformer is called a **step-up transformer**. When $N < 1$, $|\tilde{\mathbf{V}}_2| < |\tilde{\mathbf{V}}_1|$ and a transformer is called a **step-down transformer**. Either side of an ideal transformer can be used as the primary; thus, to produce a step-up transformer from a step-down transformer one only need exchange the primary and secondary connections. (Exchanging the primary and secondary by mistake can lead to significant dangers in a laboratory experiment!) Finally, when $N = 1$, a transformer is called an **isolation transformer**, which can be used to electrically couple or isolate two circuits and adjust the output and input impedances at the interface of two circuits.

A comparison of the complex power at the primary and secondary terminals of an ideal transformer reveals that they are the same:

$$\mathbf{S}_1 = \tilde{\mathbf{I}}_1^* \tilde{\mathbf{V}}_1 = N\tilde{\mathbf{I}}_2^* \frac{\tilde{\mathbf{V}}_2}{N} = \tilde{\mathbf{I}}_2^* \tilde{\mathbf{V}}_2 = \mathbf{S}_2 \tag{13.31}$$

That is, **ideal transformers conserve power**.

As shown in Figure 13.37, the secondary coil of many practical transformers is *center-tapped*, which splits the secondary voltage into two equal halves. This type of

transformer is found at the entry of a power line into a household, where a high-voltage primary is transformed to 240 V as well as split into two 120-V lines. Referring to Figure 13.37, $\tilde{\mathbf{V}}_2$ and $\tilde{\mathbf{V}}_3$ would both provide 120 V for common household appliances while $(\tilde{\mathbf{V}}_2 + \tilde{\mathbf{V}}_3)$ would provide 240 V for higher-powered devices, such as clothes dryers and electric ranges.

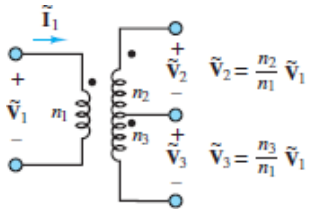


Figure 13.37 Center-tapped transformer

Page 769

Impedance Reflection

Transformers are commonly used to couple one AC circuit to another, as depicted in Figure 13.38, where an AC Thévenin source network is connected to a load $\mathbf{Z}_2$ by means of a transformer.

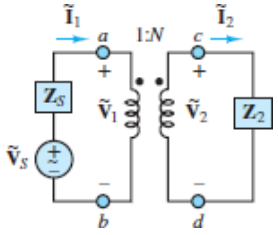


Figure 13.38 Operation of an ideal transformer

The equivalent impedance seen by the Thévenin source is that of the entire network to the right of terminals *a* and *b*. Applying the definition of equivalent impedance and using the ideal transformer relations from equation 13.30, the result is:

$$\mathbf{Z}_1 \equiv \frac{\tilde{\mathbf{V}}_1}{\tilde{\mathbf{I}}_1} = \frac{\tilde{\mathbf{V}}_2}{N}\frac{1}{N\tilde{\mathbf{I}}_2} = \frac{1}{N^2}\frac{\tilde{\mathbf{V}}_2}{\tilde{\mathbf{I}}_2} = \frac{1}{N^2}\mathbf{Z}_2 \tag{13.32}$$

Thus, the equivalent impedance seen by the AC Thévenin source is the load impedance $\mathbf{Z}_2$ reduced by the factor $1/N^2$.

Likewise, the equivalent network seen by $\mathbf{Z}_2$ is the Thévenin equivalent of the entire network to the left of terminals c and d. When $\mathbf{Z}_2$ is replaced by an open-circuit, $\tilde{\mathbf{I}}_2 = 0$ and the Thévenin (open-circuit) voltage is:

$$\tilde{\mathbf{V}}_T = (\tilde{\mathbf{V}}_2)_{OC} = N\tilde{\mathbf{V}}_1 \tag{13.33}$$

However, since $\tilde{\mathbf{I}}_1 = N\tilde{\mathbf{I}}_2 = 0$, the voltage drop across $\mathbf{Z}_S$ is zero such that $\tilde{\mathbf{V}}_1 = \tilde{\mathbf{V}}_S$ with the result:

$$\tilde{\mathbf{V}}_T = (\tilde{\mathbf{V}}_2)_{OC} = N\tilde{\mathbf{V}}_S \tag{13.34}$$

When $\mathbf{Z}_2$ is replaced by a short-circuit, $\tilde{\mathbf{V}}_2 = 0$ and the short-circuit current is:

$$(\mathbf{I}_2)_{SC} = \frac{\mathbf{I}_1}{N} \tag{13.35}$$

However, since $\tilde{\mathbf{V}}_1 = \tilde{\mathbf{V}}_2/N = 0$, the voltage drop across $\mathbf{Z}_S$ is $\tilde{\mathbf{V}}_S$ such that $\tilde{\mathbf{I}}_1 = \tilde{\mathbf{V}}_S/Z_S$ with the result:

$$(\tilde{\mathbf{I}}_2)_{SC} = \frac{1}{N}\frac{\tilde{\mathbf{V}}_S}{\mathbf{Z}_S} \tag{13.36}$$

Thus, the Thévenin equivalent impedance seen by $\mathbf{Z}_2$ is:

$$\mathbf{Z}_T = \frac{(\tilde{\mathbf{V}}_2)_{OC}}{(\tilde{\mathbf{I}}_2)_{SC}} = N\tilde{\mathbf{V}}_S\frac{N\mathbf{Z}_S}{\tilde{\mathbf{V}}_S} = N^2\mathbf{Z}_S \tag{13.37}$$

Thus, the equivalent impedance seen by $\mathbf{Z}_2$ is the source impedance $\mathbf{Z}_S$ multiplied by N^2.

Figure 13.39 summarizes and illustrates these effects, which are known as **impedance reflection** across a transformer and which play an important role in power transfer.

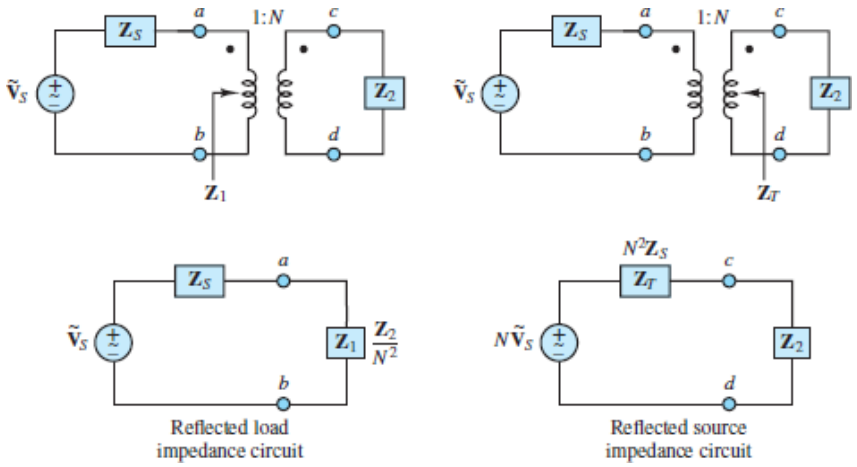


Figure 13.39 Impedance reflection across a transformer

Maximum Power Transfer

Recall that in resistive DC circuits, maximum power is transferred to a load when the load equals the Thévenin equivalent resistance of the source network. For AC circuits, the analogous maximum power transfer condition is known as **impedance matching**.

Consider the general form of an AC circuit, shown in Figure 13.40, and assume that the source impedance $\mathbf{Z}_T$ is:

$$\mathbf{Z}_T = R_T + jX_T \tag{13.38}$$

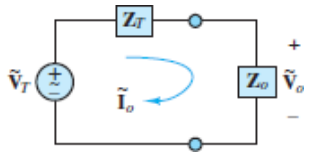


Figure 13.40 The maximum power transfer problem in AC circuits

What value of the load $\mathbf{Z}_o$ results in the maximum *real power* transfer to the load itself? The real power absorbed by the load is:

$$P_o = \tilde{V}_o \tilde{I}_o \cos\theta_{Z_o} = \mathrm{Re}\left(\tilde{\mathbf{V}}_o \tilde{\mathbf{I}}_o^*\right) \tag{13.39}$$

Apply voltage division and the generalized Ohm's law to find:

$$\tilde{\mathbf{V}}_o = \frac{\mathbf{Z}_o}{\mathbf{Z}_T + \mathbf{Z}_o}\tilde{\mathbf{V}}_T \qquad \tilde{\mathbf{I}}_o = \frac{\tilde{\mathbf{V}}_T}{\mathbf{Z}_T + \mathbf{Z}_o} \tag{13.40}$$

Let $\mathbf{Z}_o = R_o + jX_o = |\mathbf{Z}_o|\cos\theta_{Z_o} + j|\mathbf{Z}_o|\sin\theta_{Z_o}$, and since $\tilde{V}_o = |\tilde{\mathbf{V}}_o|$ and $\tilde{I}_o = |\tilde{\mathbf{I}}_o|$, the real power absorbed by the load can be expressed as:

$$P_o = \frac{|\mathbf{Z}_o|}{|\mathbf{Z}_T + \mathbf{Z}_o|}\tilde{V}_T \times \frac{1}{|\mathbf{Z}_T + \mathbf{Z}_o|}\tilde{V}_T \times \frac{R_o}{|\mathbf{Z}_o|} \tag{13.41}$$

Or, after simplification:

$$P_o = \frac{R_o}{|\mathbf{Z}_T + \mathbf{Z}_o|^2}\tilde{V}_T^2 = \frac{R_o}{(R_T + R_o)^2 + (X_T + X_o)^2}\tilde{V}_T^2 \tag{13.42}$$

The condition for the maximum value of P_o can be found by solving:

$$dP_o = \frac{\partial P_o}{\partial R_o}dR_o + \frac{\partial P_o}{\partial X_o}dX_o = 0 \tag{13.43}$$

or

$$\frac{\partial P_o}{\partial R_o} = 0 \quad \text{and} \quad \frac{\partial P_o}{\partial X_o} = 0 \tag{13.44}$$

Both of these conditions are satisfied when $R_o = R_T$ and $X_o = -X_T$. That is, the condition for maximum real power transfer to a load is $\mathbf{Z}_o = \mathbf{Z}_T^*$.

$$\boxed{\mathbf{Z}_o = \mathbf{Z}_T^* \quad \text{Maximum power transfer}} \tag{13.45}$$

Maximum power is transferred to the load when its impedance equals the complex conjugate of the Thévenin equivalent impedance of the source. When this condition is satisfied, the load and source impedances are *matched*.

In some cases, it may not be possible to match the load to the source because of practical limitations. In these situations, it may be possible to use a transformer as the interface between the source and the load to achieve maximum power transfer. Figure 13.41 illustrates how the reflected load impedance, as seen by the source, is equal to $\mathbf{Z}_o/N^2$, such that the condition for maximum power transfer is:

$$\frac{\mathbf{Z}_o}{N^2} = \mathbf{Z}_S^* \tag{13.46}$$

$$R_o = N^2 R_S$$

$$X_o = -N^2 X_S$$

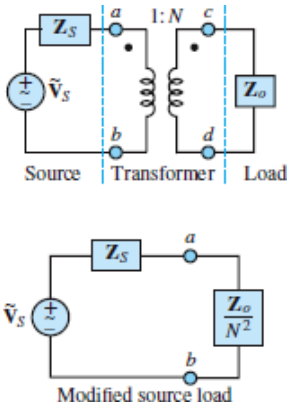


Figure 13.41 Maximum power transfer in an AC circuit with a transformer

EXAMPLE 13.12 Ideal Transformer Turns Ratio

Problem

We require a transformer (see Figure 13.42) to output 500 mA at 24 V from a 120 V rms input line source. The primary has n_1 = 3,000 turns. How many turns are required in the secondary? What is the primary current?

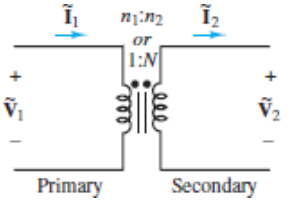


Figure 13.42 Example 13.12

Solution

Known Quantities: Primary and secondary voltages; secondary current; number of turns in the primary coil.

Find: n_2 and $\tilde{\mathbf{I}}_1$.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_1 = 120$ V; $\tilde{\mathbf{V}}_2 = 24$ V; $\tilde{\mathbf{I}}_2 = 500$ mA; $n_1 = 3{,}000$ turns.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians.

Analysis: Use equation 13.30 to compute the number of turns in the secondary coil:

$$\frac{\tilde{\mathbf{V}}_1}{n_1} = \frac{\tilde{\mathbf{V}}_2}{n_2} \qquad n_2 = n_1 \frac{\tilde{\mathbf{V}}_2}{\tilde{\mathbf{V}}_1} = 3{,}000 \times \frac{24}{120} = 600 \text{ turns}$$

Page 772

Again, use equations 13.29 and 13.30 to compute the primary current:

$$n_1 \tilde{\mathbf{I}}_1 = n_2 \tilde{\mathbf{I}}_2 \qquad \tilde{\mathbf{I}}_1 = \frac{n_2}{n_1} \tilde{\mathbf{I}}_2 = \frac{600}{3{,}000} \times 500 = 100 \text{ mA}$$

Comment: Note that since the transformer does not affect the phase of the voltages and currents, it was possible to solve the problem using only the rms amplitudes.

EXAMPLE 13.13 Center-Tapped Transformer

Problem

An ideal center-tapped power transformer (Figure 13.43) has a 4,800-V primary and a 240-V secondary. The center-tap is located such that $\tilde{v}_2 = \tilde{v}_3 = 120$ V.. Three resistive loads are attached to the secondary terminals. Compute the current in the primary assuming that R_2, R_3, and R_4 each absorb P_2, P_3, and P_4, respectively. Also compute the current through each load and the resistance of each load.

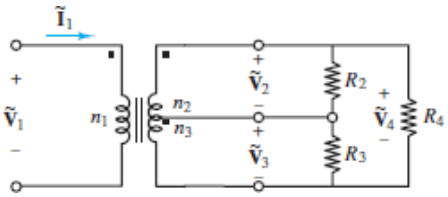


Figure 13.43 Example 13.13

Solution

Known Quantities: Primary and secondary voltages; load power ratings.

Find: $I_{\text{primary}} = |\tilde{\mathbf{I}}|$

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_1 = 4{,}800$ V rms; $\tilde{\mathbf{V}}_2 = 120$ V rms; $\tilde{\mathbf{V}}_3 = 120$ V rms; $P_2 = 5{,}000$ W; $P_3 = 1{,}000$ W; $P_4 = 1{,}500$ W.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise. The transformer is ideal.

Analysis: Power is conserved for an ideal transformer; thus:

$$\mathbf{S}_{\text{primary}} = \mathbf{S}_{\text{secondary}}$$

Since each load is purely resistive, $\theta_Z = 0$ and $\text{pf} = \cos\theta_Z = 1$ such that:

$$|\mathbf{S}|_{\text{secondary}} = P_{\text{secondary}} = P_2 + P_3 + P_4 = 7{,}500 \text{ W}$$

Since $|\mathbf{S}|_{\text{primary}} = |\mathbf{S}|_{\text{secondary}}$:

$$\tilde{V}_{\text{primary}} \times \tilde{I}_{\text{primary}} = 7{,}500 \text{ W}$$

Thus:

$$\tilde{I}_{\text{primary}} = \frac{7{,}500 \text{ W}}{4{,}800 \text{ V rms}} = 1.5625 \text{ A rms}$$

Page 773

The current through each resistor is simply:

$$\tilde{I}_2 = \frac{P_2}{\tilde{V}_2} = \frac{5{,}000 \text{ W}}{120 \text{ V rms}} \approx 41.7 \text{ A rms}$$

$$\tilde{I}_3 = \frac{P_3}{\tilde{V}_3} = \frac{1{,}000 \text{ W}}{120 \text{ V rms}} \approx 8.3 \text{ A rms}$$

$$\tilde{I}_4 = \frac{P_4}{\tilde{V}_2 + \tilde{V}_3} = \frac{1{,}500 \text{ W}}{240 \text{ V rms}} = 6.25 \text{ A rms}$$

The resistor values are:

$$\tilde{R}_2 = \frac{P_2}{\tilde{I}_2^2} = 2.88\ \Omega$$

$$\tilde{R}_3 = \frac{P_3}{\tilde{I}_3^2} = 14.4\ \Omega$$

$$\tilde{R}_4 = \frac{P_4}{\tilde{I}_4^2} = 38.4\ \Omega$$

Comments: The calculations in this example were particularly straightforward because the load was purely resistive, such that $\theta_Z = 0$, the power triangle is flat, and the apparent power S equals the real power P. When the load is complex, $\theta_Z > 0$, the power triangle is not flat, and the apparent power S equals $P\cos\theta_Z$. Then, the calculations are more complicated.

Also, KCL can be used to determine the current drawn from/to the outside and center taps. Try it!

EXAMPLE 13.14 Use of Transformers to Improve Power Line Efficiency

Problem

Figure 13.44 illustrates the use of transformers in electric power transmission lines. The line voltage is transformed before and after being transmitted over long distances. This example illustrates the efficiency gained through the use of transformers. For the sake of simplicity, ideal transformers and simple resistive models for the generator, transmission line, and load have been assumed.

Solution

Known Quantities: Values of circuit elements.

Find: Calculate the power transfer efficiency for the two circuits of Figure 13.44.

Schematics, Diagrams, Circuits, and Given Data: Step-up transformer turns ratio is N, step-down transformer turns ratio is $M = 1/N$. All transformers are ideal.

Assumptions: None.

Analysis: Since the load and source currents are equal in Figure 13.44(a), the power transmission efficiency is:

$$\eta = \frac{P_{load}}{P_{source}} = \frac{\tilde{V}_{load}\tilde{I}_{load}}{\tilde{V}_{source}\tilde{I}_{load}} = \frac{\tilde{V}_{load}}{\tilde{V}_{source}} = \frac{R_{load}}{R_{source} + R_{line} + R_{load}}$$

Page 774

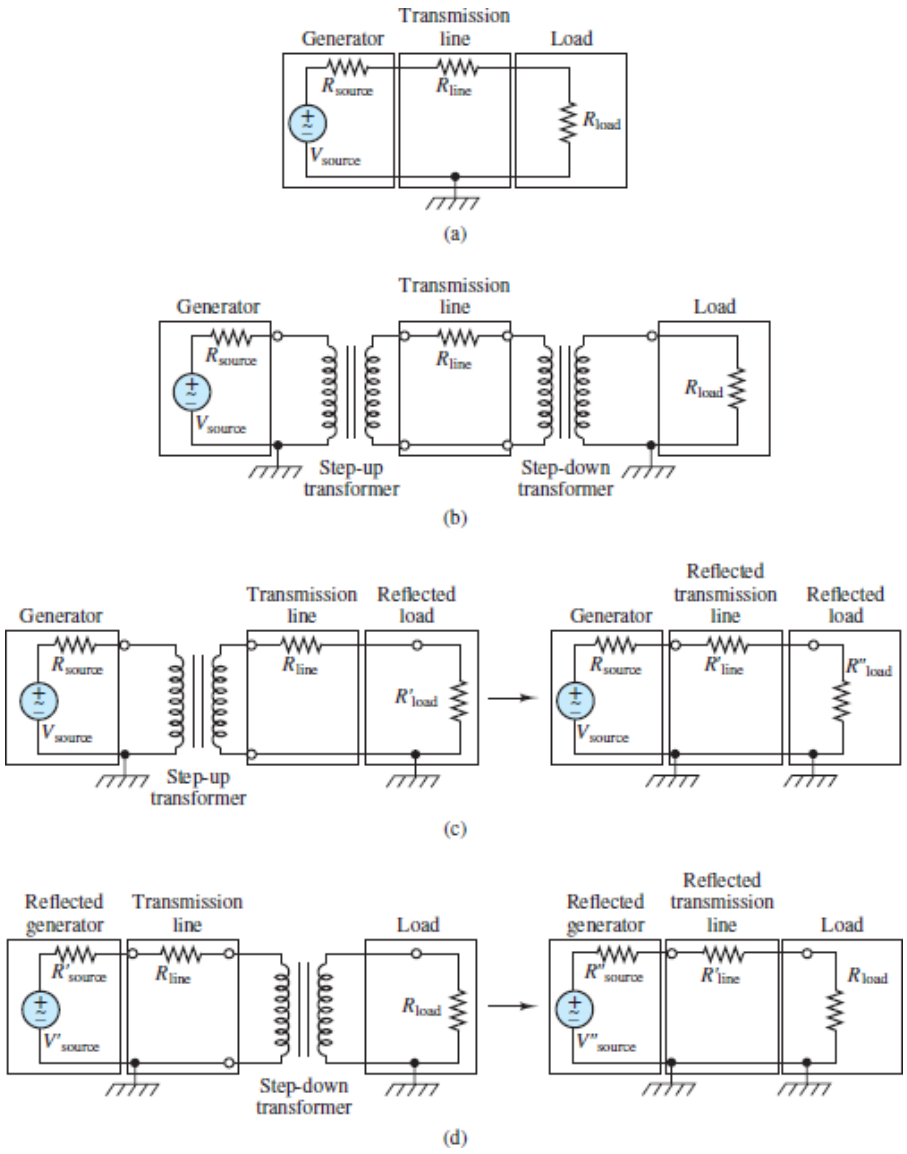


Figure 13.44 Electric power transmission: (a) direct power transmission; (b) power transmission with transformers; (c) equivalent circuit seen by generator; (d) equivalent circuit seen by load

In Figure 13.44(b), transformers are introduced between each of the three portions of the overall circuit. The equivalent load resistance seen by the transmission line (or "reflected" by the step-down transformer) is found from equation 13.32 to be:

$$R'_{\text{load}} = \frac{1}{M^2} R_{\text{load}} = N^2 R_{\text{load}}$$

Now, the step-up transformer sees the equivalent impedance $R'_{\text{load}} + R_{\text{line}}$. The resistance seen by the generator (or "reflected" by the step-up transformer) is:

$$R''_{\text{load}} = \frac{1}{N^2}(R'_{\text{load}} + R_{\text{line}}) = R_{\text{load}} + \frac{1}{N^2} R_{\text{line}}$$

Page 775

These transformations are depicted in Figure 13.44(c). The effect of the two transformers is to reduce the line resistance seen by the source by N_2. The source current is:

$$\tilde{I}_{\text{source}} = \frac{\tilde{V}_{\text{source}}}{R_{\text{source}} + R''_{\text{load}}} = \frac{\tilde{V}_{\text{source}}}{R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}}$$

The source power is:

$$P_{\text{source}} = \frac{\tilde{V}^2_{\text{source}}}{R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}}$$

The same process can be repeated starting from the left and reflecting the source circuit to the right of the step-up transformer:

$$\tilde{V}'_{\text{source}} = N\tilde{V}_{\text{source}} \quad \text{and} \quad R'_{\text{source}} = N^2 R_{\text{source}}$$

Now the circuit to the left of the step-down transformer comprises the series combination of $\tilde{V}'_{\text{source}}$, R'_{source}, and R_{line}, which can be reflected to the right of the step-down transformer to obtain $\tilde{V}''_{\text{source}} = M\tilde{V}'_{\text{source}} = \tilde{V}_{\text{source}}$, $R'_{\text{source}} = M^2 R'_{\text{source}} = R_{\text{source}}$, $R'_{\text{line}} = M^2 R_{\text{line}}$, , and R_{load} in series. These transformations are depicted in Figure 13.44(d). Thus, the load voltage, current, and power are:

$$\tilde{I}_{\text{load}} = \frac{\tilde{V}_{\text{source}}}{R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}}$$

$$\tilde{V}_{\text{load}} = \tilde{V}_{\text{source}} \frac{R_{\text{load}}}{R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}}$$

$$P_{\text{load}} = \tilde{I}_{\text{load}}\tilde{V}_{\text{load}} = \frac{\tilde{V}^2_{\text{source}} R_{\text{load}}}{\left[R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}\right]^2}$$

Finally, the power efficiency can be computed as the ratio of the load to source power:

$$\eta = \frac{P_{\text{load}}}{P_{\text{source}}} = \frac{\tilde{V}^2_{\text{source}} R_{\text{load}}}{\left[R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}\right]^2} \frac{R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}}{\tilde{V}^2_{\text{source}}}$$

$$= \frac{R_{\text{load}}}{R_{\text{source}} + (1/N^2)R_{\text{line}} + R_{\text{load}}}$$

Notice that the power transmission efficiency calculated for Figure 13.44(a) was improved by reducing the effect of the line resistance by a factor of $1/N^2$.

EXAMPLE 13.15 Maximum Power Transfer Through a Transformer

Problem

Find the transformer turns ratio N and the load reactance X_o that results in maximum power transfer in the transformer shown in Figure 13.45.

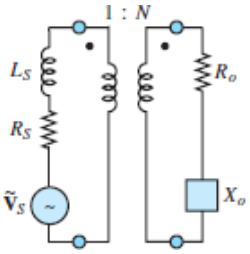


Figure 13.45 Circuit for Example 13.15.

Solution

Known Quantities: Source voltage, frequency, and impedance; load resistance.

Find: Transformer turns ratio and load reactance.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_S = 240\angle 0$ V rms; $R_S = 10\ \Omega$; $L_S = 0.1$ H; $R_o = 400\ \Omega$; $\omega = 377$ rad/s.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians. The transformer is ideal.

Analysis: The requirements for maximum power transfer, as given by equation 13.46, are $R_o = N^2 R_S$ and $X_o = -N^2 X_S = -N^2(\omega \times 0.1)$. Thus:

$$N^2 = \frac{R_o}{R_S} = \frac{400}{10} = 40 \qquad N = \sqrt{40} = 6.325$$

$$X_o = -40 \times 37.7 = -1{,}508\ \Omega$$

Thus, the load reactance should be a capacitor with value:

$$C = -\frac{1}{X_o \omega} = -\frac{1}{(-1{,}508)(377)} = 1.76\ \mu\text{F}$$

CHECK YOUR UNDERSTANDING

With reference to Example 13.12, compute the number of primary turns required if n_2 = 600 but the transformer is required to deliver 1 A. What is the primary current now?

Answer: $n_1 = 3{,}000$; $I_1 = 200$ mA

CHECK YOUR UNDERSTANDING

If the transformer of Example 13.12 has 300 turns in the secondary coil, how many turns will the primary require?

Answer: $n_2 = 6{,}000$

CHECK YOUR UNDERSTANDING

Assume that the generator produces a source voltage of 480 V rms, and that $N = 300$. Further assume that the source impedance is 2 Ω, the line impedance is also 2 Ω, and that the load impedance is 8 Ω. Calculate the efficiency improvement for the circuit of Figure 13.37(b) over the circuit of Figure 13.37(a).

Answer: 80% vs. 67%

CHECK YOUR UNDERSTANDING

The transformer shown in Figure 13.46 is ideal. Assume that $\mathbf{Z}_S = 1{,}800\Omega$ and $\mathbf{Z}_o = 8\Omega$ to find the turns ratio N that will ensure maximum power transfer to the load.

Page 777

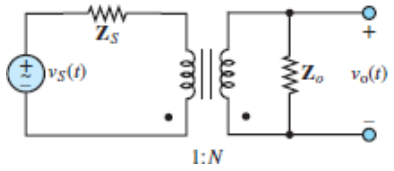


Figure 13.46 Ideal transformer—calculation of maximum power transfer.

Now assume that $N = 5.4$ and $\mathbf{Z}_o = 2 + j10\Omega$. Find the source impedance $\mathbf{Z}_S$ that will ensure maximum power transfer to the load.

Answer: $N = 0.0667$; $\mathbf{Z}_S = 0.0686 - j0.3429\ \Omega$

13.5 THREE-PHASE POWER

The material presented so far in this chapter has dealt exclusively with **single-phase AC power**, which implies a single sinusoidal source. However, most of the AC power used today is generated and distributed as **three-phase power**, which implies three sinusoidal sources, each out of phase with the other. The primary benefit is efficiency: The weight of the conductors and other components in a three-phase system is much lower than that in a single-phase system delivering the same amount

of power. Further, while the power produced by a single-phase system has a pulsating nature (recall the results of Section 13.1), a three-phase system can deliver a steady, constant supply of power. For example, later in this section it will be shown that a three-phase generator producing three **balanced voltages**—that is, voltages of equal amplitude and frequency displaced in phase by 120°—has the property of delivering constant instantaneous power.

The change to three-phase AC power systems from the early DC system proposed by Edison was due to a number of reasons: the efficiency resulting from transforming voltages up and down to minimize transmission losses over long distances, the ability to deliver constant power, a more efficient use of conductors, and the ability to provide starting torque for industrial motors.

Consider a three-phase source connected in a **wye (Y) configuration**, as shown in Figure 13.47. Each of the three voltages is 120° out of phase with the Page 778others, such that:

$$\begin{aligned} \tilde{\mathbf{V}}_{an} &= \tilde{V}_{an}\angle 0^\circ \\ \tilde{\mathbf{V}}_{bn} &= \tilde{V}_{bn}\angle -(120^\circ) \\ \tilde{\mathbf{V}}_{cn} &= \tilde{V}_{cn}\angle (-240^\circ) = \tilde{V}_{cn}\angle 120^\circ \end{aligned} \qquad \text{Phase voltages} \tag{13.47}$$

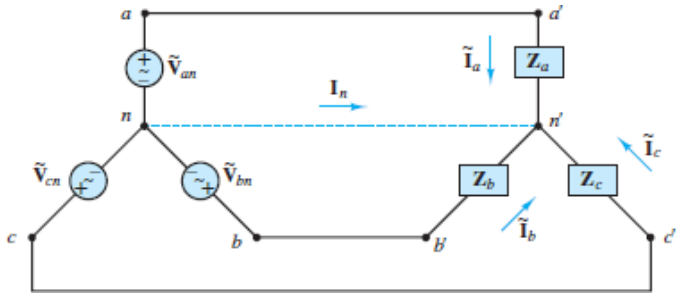


Figure 13.47 Balanced three-phase AC circuit

If the three-phase source is *balanced*, then:

$$\tilde{\mathbf{V}}_{an} + \tilde{\mathbf{V}}_{bn} + \tilde{\mathbf{V}}_{cn} = 0 \qquad \text{Balanced phase voltages} \tag{13.48}$$

For three balanced **phase voltages** each separated by 120°, the phase amplitudes are also equal:

$$\tilde{V}_{an} = \tilde{V}_{bn} = \tilde{V}_{cn} = \tilde{V} \tag{13.49}$$

The result is the so-called **positive abc sequence**, as shown in Figure 13.48. In the wye configuration, the three phase voltages share a common *neutral node*, denoted by n.

It is also possible to define **line voltages** as the potential differences between lines aa' and bb', lines aa' and cc', and lines bb' and cc'. Each line voltage is related to the phase voltages by:

$$\begin{aligned}\tilde{\mathbf{V}}_{ab} &= \tilde{\mathbf{V}}_{an} - \tilde{\mathbf{V}}_{bn} = \sqrt{3}\tilde{V}\angle 30^\circ \\ \tilde{\mathbf{V}}_{bc} &= \tilde{\mathbf{V}}_{bn} - \tilde{\mathbf{V}}_{cn} = \sqrt{3}\tilde{V}\angle(-90^\circ) \\ \tilde{\mathbf{V}}_{ca} &= \tilde{\mathbf{V}}_{cn} - \tilde{\mathbf{V}}_{an} = \sqrt{3}\tilde{V}\angle 150^\circ\end{aligned} \quad \text{Line voltages} \tag{13.50}$$

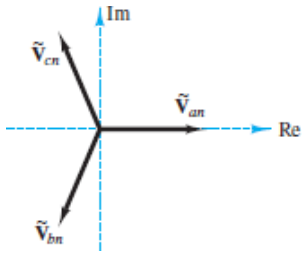


Figure 13.48 Positive, or *abc*, sequence for balanced three-phase voltages

It is instructive to note that the circuit of Figure 13.47 can be redrawn as shown in Figure 13.49, where it is clear that the three branches are in parallel.

When $\mathbf{Z}_a = \mathbf{Z}_b = \mathbf{Z}_c = \mathbf{Z}$, the wye load configuration is also balanced. When both the source and load networks are balanced, KCL requires that the current $\tilde{\mathbf{I}}_n$ in the neutral line $n - n'$ be identically zero.

$$\tilde{\mathbf{I}}_n = \tilde{\mathbf{I}}_a + \tilde{\mathbf{I}}_b + \tilde{\mathbf{I}}_c = \frac{\tilde{\mathbf{V}}_{an} + \tilde{\mathbf{V}}_{bn} + \tilde{\mathbf{V}}_{cn}}{\mathbf{Z}} = 0 \tag{13.51}$$

Another important characteristic of a balanced three-phase power system is illustrated by a simplified version of Figure 13.49, where the balanced load impedances are replaced by three equal resistors R. Since $\theta_R = 0$, the instantaneous power $p(t)$ delivered to each resistor is given by equation 13.4 [with $\theta_V = \theta_I$ and with the freely chosen reference $(\theta_V)_a = 0$] to be:

$$p_a(t) = \frac{\tilde{V}^2}{R}(1 + \cos 2\omega t)$$
$$p_b(t) = \frac{\tilde{V}^2}{R}[1 + \cos(2\omega t - 120°)] \tag{13.52}$$
$$p_c(t) = \frac{\tilde{V}^2}{R}[1 + \cos(2\omega t + 120°)]$$

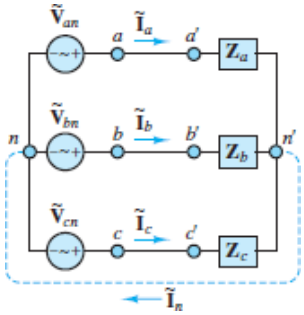


Figure 13.49 Balanced three-phase AC circuit (redrawn)

Page 779

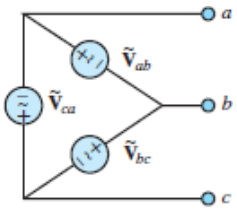
The total instantaneous power $p(t)$ delivered to the total load is the sum:

$$\begin{aligned} p(t) &= p_a(t) + p_b(t) + p_c(t) \\ &= \frac{\tilde{V}^2}{R}[3 + \cos 2\omega t + \cos(2\omega t - 120°) + \cos(2\omega t + 120°)] \\ &= \frac{3\tilde{V}^2}{R} = \text{constant!} \end{aligned} \tag{13.53}$$

It is worthwhile to verify that the sum of the three cosine terms is identically zero. (*Hint:* Consider the phasor sum of $e^{j(2\omega t)}$, $e^{j(2\omega t - \pi/3)}$, and $e^{j(2\omega t + \pi/3)}$.)

Thus, with the simplified balanced resistive load, the total power delivered to the load by the balanced three-phase source is constant. This is an extremely important result, for a very practical reason: Delivering power in a steady fashion (as opposed to the pulsating nature of single-phase power) reduces "wear and tear" on the source and load.

It is also possible to connect three AC sources in a **delta (Δ) configuration**, as shown in Figure 13.50 although it is rarely used in practice.



A delta-connected three-phase generator with line voltages V_{ab}, V_{bc}, V_{ca}

Figure 13.50 Delta configuration

Balanced Wye Loads

These results for purely resistive loads can be generalized for any arbitrary balanced complex load. Consider again in Figure 13.47, where now the balanced load consists of three complex impedances:

$$\mathbf{Z}_a = \mathbf{Z}_b = \mathbf{Z}_c = \mathbf{Z}_y = |\mathbf{Z}_y|\angle\theta \tag{13.54}$$

Because of the common neutral line $n - n'$, each impedance sees the corresponding phase voltage across itself. Therefore, since $\tilde{V}_{an} = \tilde{V}_{bn} = \tilde{V}_{cn}$, it is also true that $\tilde{I}_a = \tilde{I}_b = \tilde{I}_c$ and the phase angles of the currents will differ by ±120°. Consequently, it is possible to compute the power for each phase from the phase voltage and the associated line current. Denote the complex power for each phase by **S**, where:

$$\begin{aligned}\mathbf{S} &= P + jQ \\ &= \tilde{V}\tilde{I}\cos\theta + j\tilde{V}\tilde{I}\sin\theta\end{aligned} \tag{13.55}$$

The total real power delivered to the balanced wye load is $3P$, and the total reactive power is $3Q$. The total complex power $\mathbf{S}_T$ is:

$$\begin{aligned}\mathbf{S}_T &= P_T + jQ_T = 3P + j3Q \\ &= \sqrt{(3P)^2 + (3Q)^2}\angle\theta\end{aligned} \tag{13.56}$$

The apparent power $|\mathbf{S}_T|$ is:

$$\begin{aligned}|\mathbf{S}_T| &= 3\sqrt{(\tilde{V}\tilde{I})^2\cos^2\theta + (\tilde{V}\tilde{I})^2\sin^2\theta} \\ &= 3\tilde{V}\tilde{I}\end{aligned} \tag{13.57}$$

such that:

$$\begin{aligned}P_T &= |\mathbf{S}_T|\cos\theta \\ Q_T &= |\mathbf{S}_T|\sin\theta\end{aligned} \tag{13.58}$$

Balanced Delta Loads

It is also possible to assemble a balanced load in a delta configuration. A wye generator and a delta load are shown in Figure 13.51.

Page 780

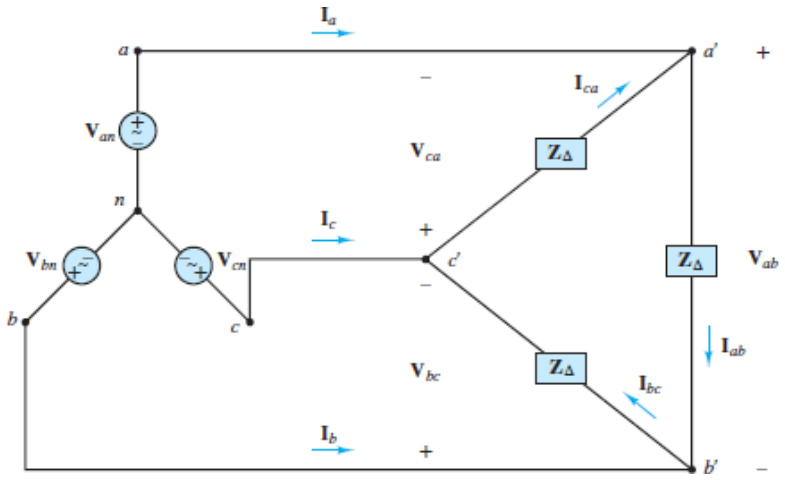


Figure 13.51 Balanced wye generators with balanced delta load

Note immediately that each impedance $\mathbf{Z}_\Delta$ sees a corresponding line voltage, rather than a phase voltage. For example, the voltage across $\mathbf{Z}_{c'a'}$ is $\tilde{\mathbf{V}}_{ca}$. Thus, the three load currents are:

$$\tilde{\mathbf{I}}_{ab} = \frac{\tilde{\mathbf{V}}_{ab}}{\mathbf{Z}_\Delta} = \frac{\sqrt{3}\tilde{V}\angle(\pi/6)}{|\mathbf{Z}_\Delta|\angle\theta}$$
$$\tilde{\mathbf{I}}_{bc} = \frac{\tilde{\mathbf{V}}_{bc}}{\mathbf{Z}_\Delta} = \frac{\sqrt{3}\tilde{V}\angle(-\pi/2)}{|\mathbf{Z}_\Delta|\angle\theta} \qquad (13.59)$$
$$\tilde{\mathbf{I}}_{ca} = \frac{\tilde{\mathbf{V}}_{ca}}{\mathbf{Z}_\Delta} = \frac{\sqrt{3}\tilde{V}\angle(5\pi/6)}{|\mathbf{Z}_\Delta|\angle\theta}$$

The relationship between a delta load and a wye load can be illustrated by determining the delta load $\mathbf{Z}_\Delta$ that would draw the same amount of current as a wye load $\mathbf{Z}_y$, assuming a given source voltage. Consider the circuits shown in Figures 13.47 and 13.51. For instance, the line current drawn in phase *a* by *a* wye load is:

$$(\tilde{\mathbf{I}}_a)_y = \frac{\tilde{\mathbf{V}}_{an}}{\mathbf{Z}} = \frac{\tilde{V}}{|\mathbf{Z}_y|}\angle(-\theta) \qquad (13.60)$$

The current drawn by a delta load is:

$$
\begin{aligned}
(\tilde{\mathbf{I}}_a)_\Delta &= \tilde{\mathbf{I}}_{ab} - \tilde{\mathbf{I}}_{ca} \\
&= \frac{\tilde{\mathbf{V}}_{ab}}{\mathbf{Z}_\Delta} - \frac{\tilde{\mathbf{V}}_{ca}}{\mathbf{Z}_\Delta} \\
&= \frac{1}{\mathbf{Z}_\Delta}(\tilde{\mathbf{V}}_{an} - \tilde{\mathbf{V}}_{bn} - \tilde{\mathbf{V}}_{cn} + \tilde{\mathbf{V}}_{an}) \\
&= \frac{1}{\mathbf{Z}_\Delta}(2\tilde{\mathbf{V}}_{an} - \tilde{\mathbf{V}}_{bn} - \tilde{\mathbf{V}}_{cn}) \\
&= \frac{3\tilde{\mathbf{V}}_{an}}{\mathbf{Z}_\Delta} = \frac{3\tilde{\mathbf{V}}}{|\mathbf{Z}_\Delta|}\angle(-\theta)
\end{aligned}
\tag{13.61}
$$

Page 781

The two currents $(\tilde{\mathbf{I}}_a)_\Delta$ and $(\tilde{\mathbf{I}}_a)_y$ are equal if:

$$
\mathbf{Z}_\Delta = 3\mathbf{Z}_y \tag{13.62}
$$

This result also implies that a delta load will draw three times as much current and absorb three times as much power as a wye load with the same branch impedance.

EXAMPLE 13.16 Per-Phase Solution of Balanced Wye-Wye Circuit

Problem

Compute the power delivered to the load by the three-phase generator in the circuit shown in Figure 13.52.

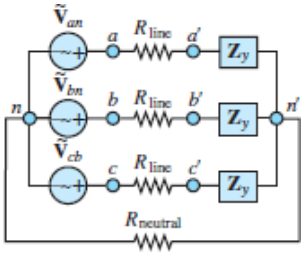


Figure 13.52 Circuit for Example 13.16.

Solution

Known Quantities: Source voltage, line resistance, load impedance.

Find: Power delivered to the load P_{load}.

Schematics, Diagrams, Circuits, and Given Data: $\tilde{\mathbf{V}}_{an} = 480\angle 0$ V rms; $\tilde{\mathbf{V}}_{bn} = 480\angle(-2\pi/3)$ V rms; $\tilde{\mathbf{V}}_{cn} = 480\angle(2\pi/3)$ V rms; $R_{\text{line}} = 2\ \Omega$; $R_{\text{neutral}} = 10\ \Omega$; $\mathbf{Z}_y = R_o + jX_o = 2 + j4 = 4.47\angle 1.107\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians unless indicated otherwise.

Analysis: Since the circuit is balanced, $\tilde{\mathbf{V}}_{n-n'} = 0$ and the current through the neutral line is zero. As a result, each phase has the structure shown in Figure 13.53. For example, the real power absorbed by the load in phase *a* is:

$$P_a = |\tilde{\mathbf{I}}_a|^2 R_o$$

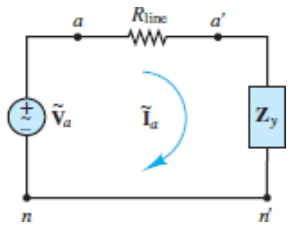


Figure 13.53 One phase of the three-phase circuit

where

$$|\tilde{\mathbf{I}}_a| = \left|\frac{\tilde{\mathbf{V}}_a}{\mathbf{Z}_y + R_{\text{line}}}\right| = \left|\frac{480\angle 0}{2 + j4 + 2}\right| = \left|\frac{480\angle 0}{5.66\angle(\pi/4)}\right| = 84.85\ \text{A rms}$$

and $Pa = (84.85\text{A})^2(2\Omega) = 14.4$ kW. Since the circuit is balanced, the results for phases *b* and *c* are identical, such that:

$$P_{\text{load}} = 3P_a = 43.2\ \text{kW}$$

EXAMPLE 13.17 Parallel Wye-Delta Load Circuit

Problem

Compute the power delivered to the wye-delta load by the three-phase generator in the circuit shown in Figure 13.54.

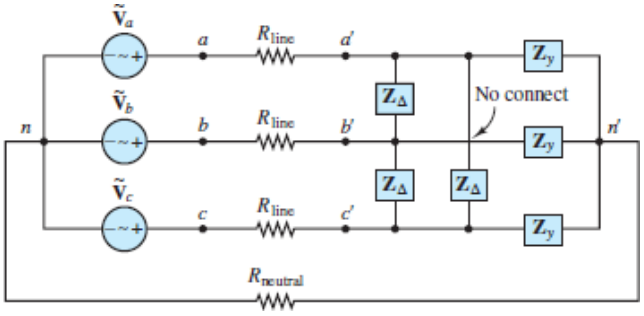


Figure 13.54 AC circuit with delta and wye loads

Solution

Known Quantities: Source voltage, line resistance, load impedance.

Page 782

Find: Power delivered to the load P_{load}.

Schematics, Diagrams, Circuits, and Given Data:

$\tilde{\mathbf{V}}_{an} = 480\angle 0$ V rms; $\tilde{\mathbf{V}}_{bn} = 480\angle(-2\pi/3)$ V rms;

$\tilde{\mathbf{V}}_{cn} = 480\angle(2\pi/3)$ V rms; $\mathbf{Z}_y = 2 + j4 = 4.47\angle 1.107\ \Omega$; $\mathbf{Z}_\Delta = 5 - j2 = 5.4\angle(-0.381)\Omega$; $R_{\text{line}} = 2\ \Omega$; $R_{\text{neutral}} = 10\ \Omega$.

Assumptions: All amplitudes are effective (rms) values. All angles are given in units of radians.

Analysis: First, convert the balanced delta load to an equivalent wye load, according to equation 13.62. Figure 13.55 illustrates the effect of this conversion.

$$\mathbf{Z}_{\Delta-y} = \frac{\mathbf{Z}_\Delta}{3} = 1.667 - j0.667 = 1.8\angle(-0.381)\Omega$$

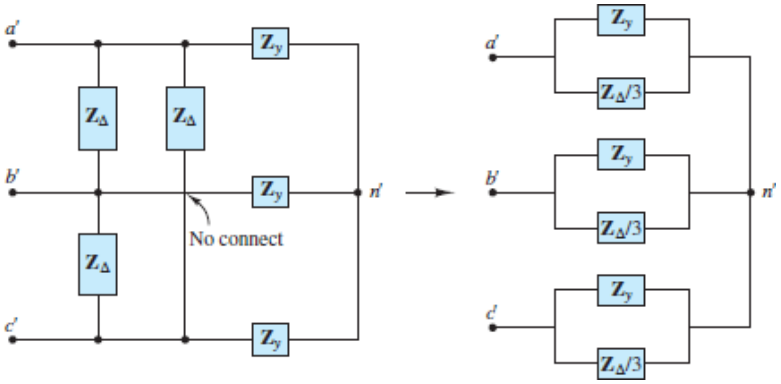


Figure 13.55 Conversion of delta load to equivalent wye load

Since the circuit is balanced, $\tilde{\mathbf{V}}_{n-n'} = 0$ and the current through the neutral line is zero. The resulting per-phase circuit is shown in Figure 13.56. For example, the real power absorbed by the load in phase *a* is:

$$P_a = |\tilde{\mathbf{I}}_a|^2 R_a = |\tilde{\mathbf{I}}_a|^2 \text{Re}(\mathbf{Z}_a)$$

where

$$\mathbf{Z}_a = \mathbf{Z}_y \| \mathbf{Z}_{\Delta-y} = \frac{\mathbf{Z}_y \times \mathbf{Z}_{\Delta-y}}{\mathbf{Z}_y + \mathbf{Z}_{\Delta-y}} = 1.62 - j0.018 = 1.62\angle(-0.011)\Omega$$

The load current $|\tilde{\mathbf{I}}_a|$ is:

$$|\tilde{\mathbf{I}}_a| = \left|\frac{\tilde{\mathbf{V}}_a}{\mathbf{Z}_o + R_{\text{line}}}\right| = \left|\frac{480\angle 0}{1.62 + j0.018 + 2}\right| = 132.6 \text{ A rms}$$

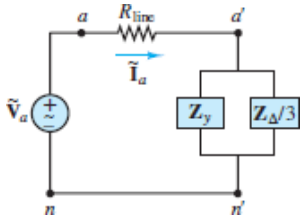


Figure 13.56 Per-phase circuit

Page 783

Thus, $P_a = (132.6)^2 \times \text{Re}(\mathbf{Z}_o) = 28.5$ kW. Since the circuit is balanced, the results for phases *b* and *c* are identical, such that:

$$P_{\text{load}} = 3P_a = 85.5 \text{ kW}$$

CHECK YOUR UNDERSTANDING

Find the power lost in the line resistance shown in Example 13.16.

Compute the complex power $\mathbf{S}_o$ delivered to the balanced load of Example 13.16 if the lines have zero resistance and $\mathbf{Z}_y = 1 + j3\Omega$.

Show that the voltage across each branch of the wye load is equal to the corresponding phase voltage (e.g., the voltage across $\mathbf{Z}_a$ is $\tilde{\mathbf{V}}_a$).

Prove that the sum of the instantaneous powers absorbed by the three branches in a balanced wye-connected load is constant and equal to $3\,\tilde{\mathbf{V}}\tilde{\mathbf{I}}\cos\theta$.

Answer: $P_{line} = 43.2\text{ kW}, S_o = 69.12\text{kW} + j207.4\text{kVA}$

13.6 RESIDENTIAL WIRING; GROUNDING AND SAFETY

Common residential electric power service consists of a three-wire AC system supplied by the local power company. The three wires originate from a utility pole and consist of a neutral wire, which is connected to earth ground, and two "hot" wires. Each of the hot lines supplies 120 V rms to the residential circuits; the two lines are 180° out of phase, for reasons that will become apparent during the course of this discussion. The phasor line voltages, shown in Figure 13.57, are usually referred to by means of a subscript convention derived from the color of the insulation on the different wires: *W* for white (neutral), *B* for black (hot), and *R* for red (hot). This convention is adhered to uniformly.

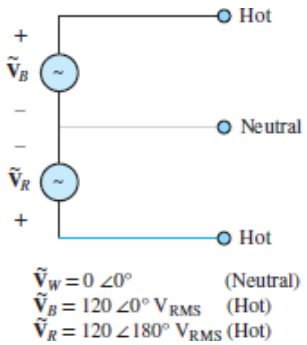


Figure 13.57 Line voltage convention for residential circuits

The voltages across the hot lines are given by

$$\tilde{\mathbf{V}}_B - \tilde{\mathbf{V}}_R = \tilde{\mathbf{V}}_{BR} = \tilde{\mathbf{V}}_B - (-\tilde{\mathbf{V}}_B) = 2\tilde{\mathbf{V}}_B = 240\angle 0° \tag{13.63}$$

Appliances such as electric stoves, air conditioners, and heaters are powered by the 240 V rms arrangement. On the other hand, lighting and all the electric outlets in the house used for small appliances are powered by a single 120 V rms line.

The use of 240 V rms service for appliances that require a substantial amount of power to operate is dictated by power transfer considerations. Consider the two circuits shown in Figure 13.58. In delivering the necessary power to a load, a lower line loss will be incurred with the 240 V rms wiring since the power loss in the lines (the $\boldsymbol{I^2R}$ **loss**, as it is commonly referred to) is directly related to the current required by the load. In an effort to minimize line losses, the size of the wires is increased for the lower-voltage case. This typically reduces the wire resistance by a factor of 2. In the top circuit, assuming $R_S/2 = 0.01\ \Omega$, the current required by the 10-kW load is approximately 83.3 A while in the bottom circuit, with $R_S = 0.02\ \Omega$, it is approximately one-half as much (41.7 A). (You should be able to verify that the approximate I^2R losses are 69.4 W in the top circuit and 34.7 W in the bottom circuit.) Limiting the I^2R losses is important from the viewpoint of Page 784efficiency, besides reducing the amount of heat generated in the wiring for safety considerations. Figure 13.59 shows some typical wiring configurations for a home. Note that several circuits are wired and fused separately.

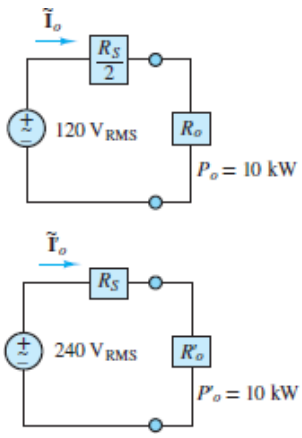


Figure 13.58 Line losses in 120- and 240-VAC circuits

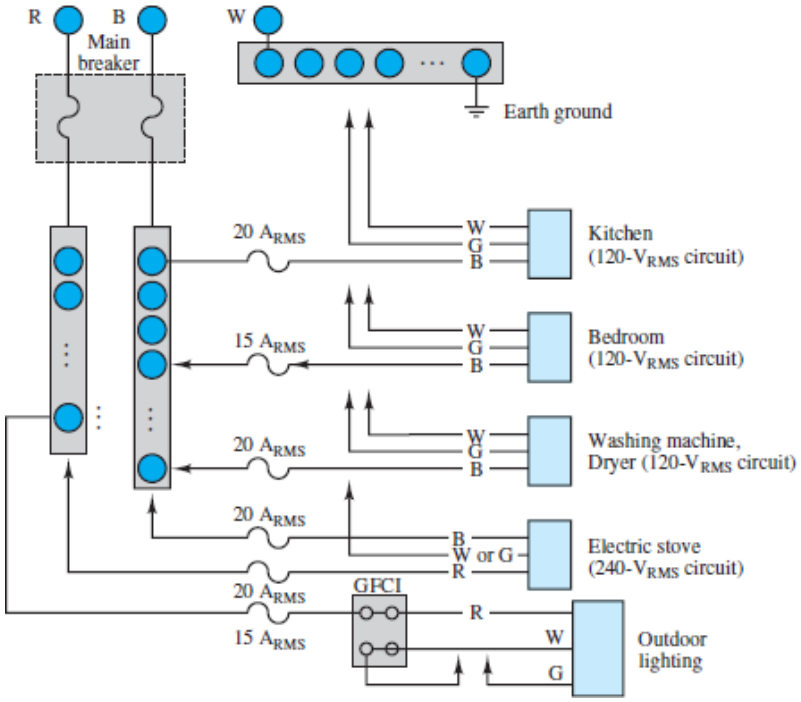


Figure 13.59 A typical residential wiring arrangement

Today, most homes have three wire connections to their outlets. The outlets appear as sketched in Figure 13.60. Then why are both the ground and neutral connections needed in an outlet? The answer to this question is *safety:* The ground connection is used to connect the chassis of the appliance to earth ground. Without this provision, the appliance chassis could be at any potential with respect to ground, possibly even at the hot wire's potential if a segment of the hot wire were to lose some insulation and come in contact with the inside of the chassis! Poorly grounded appliances can thus be a significant hazard. Figure 13.61 illustrates schematically how even though the chassis is intended to be insulated from the electric circuit, an unintended connection (represented by the dashed line) may occur, for example, because of corrosion or a loose mechanical connection. A path to ground might be provided by the body of a person touching the chassis with a hand. In the figure, such an undesired ground loop current is indicated by I_G. In this case, the ground current I_G would pass directly through the body to ground and could be harmful.

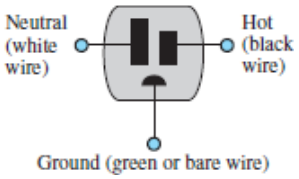


Figure 13.60 A three-wire outlet

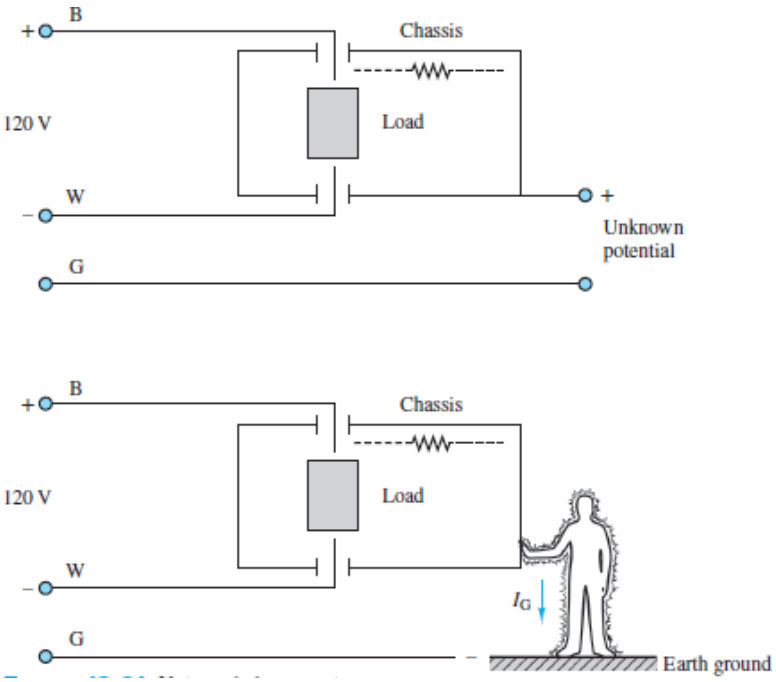


Figure 13.61 Unintended connection

In some cases the danger posed by such undesired ground loops can be great, leading to death by electric shock. Figure 13.62 describes the effects of electric currents on an average male when the point of contact is dry skin. Particularly hazardous conditions are liable to occur whenever the natural resistance to current provided by the skin breaks down, as would happen in the presence of water. Thus, the danger presented to humans by unsafe electric circuits is very much dependent on the particular conditions—whenever water or moisture is present, the natural electrical resistance Page 785of dry skin, or of dry shoe soles, decreases dramatically, and even relatively low voltages can lead to fatal currents. Proper grounding procedures, such as those required by the National Electrical Code, help prevent fatalities due to electric shock. The **ground fault circuit interrupter**, labeled **GFCI** in Figure 13.59, is a special safety circuit used primarily with outdoor circuits and in bathrooms, where the risk of death by electric shock is greatest. Its application is best described by an example.

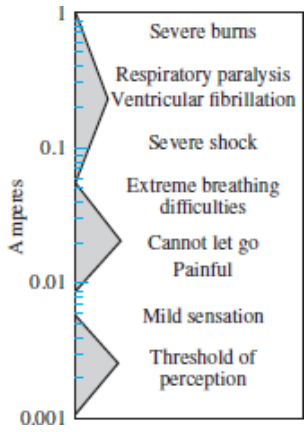


Figure 13.62 Physiological effects of electric currents

Consider the case of an outdoor pool surrounded by a metal fence, which uses an existing light pole for a post, as shown in Figure 13.63. The light pole and the Page 786metal fence can be considered as forming a chassis. If the fence were not properly grounded all the way around the pool and if the light fixture were poorly insulated from the pole, a path to ground could easily be created by an unaware swimmer reaching, say, for the metal gate. A GFCI provides protection from potentially lethal ground loops, such as this one, by sensing both the hot-wire (B) and the neutral (W) currents. If the difference between the hot-wire current I_B and the neutral current I_W is more than a few milliamperes, then the GFCI disconnects the circuit nearly instantaneously. Any significant difference between the hot and neutral (return-path) currents means that a second path to ground has been created (by the unfortunate swimmer, in this example) and a potentially dangerous condition has arisen. Figure 13.64 illustrates the idea. GFCIs are typically resettable circuit breakers, so that one does not need to replace a fuse every time the GFCI circuit is enabled.

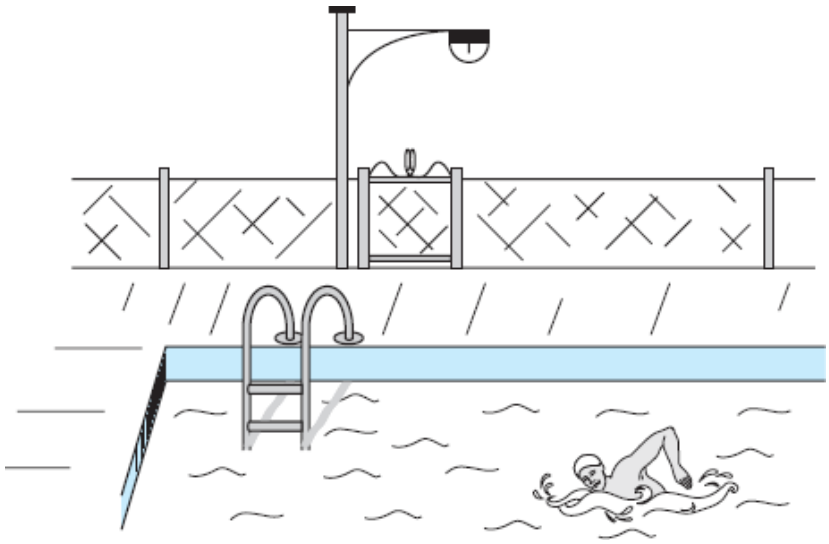


Figure 13.63 Outdoor pool

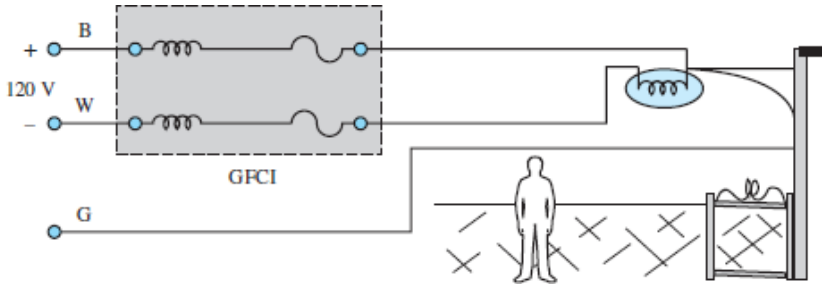


Figure 13.64 Use of a GFCI in a potentially hazardous setting

CHECK YOUR UNDERSTANDING

Use the circuits of Figure 13.58 to show that the I^2R losses will be higher for a 120-V service appliance than a 240-V service appliance if both have the same power usage rating.

Answer: The 120 V rms circuit has double the losses of the 240 V rms circuit for the same power rating.

13.7 POWER GENERATION AND DISTRIBUTION

We now conclude the discussion of power systems with a brief description of the various elements of a power system. Electric power originates from a variety of sources; in Chapter 15, electric generators are introduced as a means of producing electric power from a variety of energy conversion processes. In general, electric power may be obtained from hydroelectric, thermoelectric, geothermal, wind, solar, and nuclear sources. The choice of a given source is typically dictated by the power requirement for the given application, and by economic and environmental factors. In this section, the structure of an AC power network, from the power-generating station to the residential circuits discussed in Section 13.6, is briefly outlined.

A typical generator will produce electric power at 18 kV rms, as shown in the diagram of Figure 13.65. To minimize losses along the conductors, the output of the generators is processed through a step-up transformer to achieve line voltages Page 787of hundreds of kilovolts (345 kV rms, in Figure 13.65). Without this transformation, the majority of the power generated would be lost in the **transmission lines** that carry the electric current from the power station.

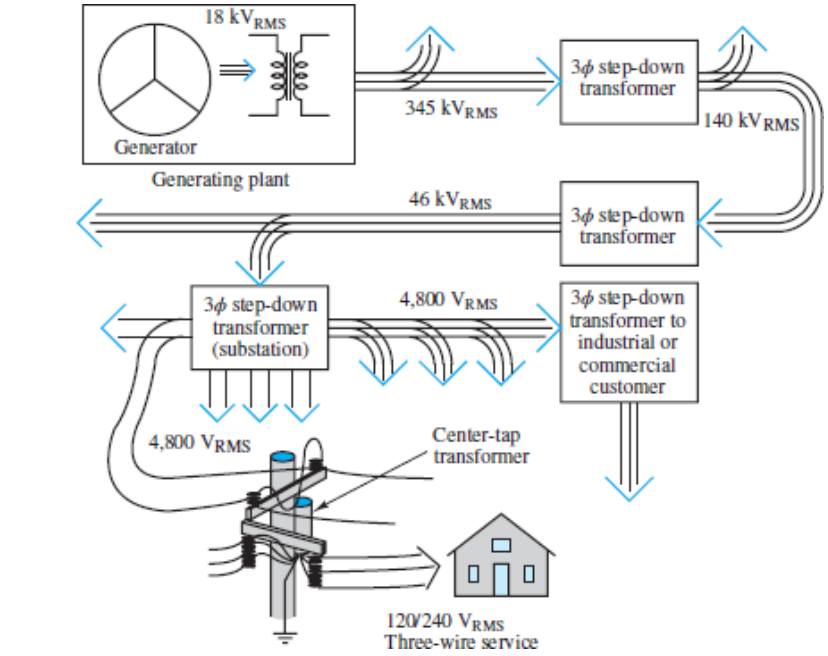


Figure 13.65 Structure of an AC power distribution network

The local electric company operates a power-generating plant that is capable of supplying several hundred megavolt-amperes (MVA) on a three-phase basis. For this reason, the power company uses a three-phase step-up transformer at the generation plant to increase the line voltage to around 345 kV rms. One can immediately see

that at the rated power of the generator (in megavolt-amperes) there will be a significant reduction of current beyond the step-up transformer.

Beyond the generation plant, an electric power network distributes energy to several **substations**. This network is usually referred to as the **power grid**. At the substations, the voltage is stepped down to a lower level (10 to 150 kV rms, typically). Some very large loads (e.g., an industrial plant) may be served directly from the power grid although most loads are supplied by individual substations in the power grid. At the local substations (one of which you may have seen in your own neighborhood), the voltage is stepped down further by a three-phase step-down transformer to 4,800 V. These substations distribute the energy to residential and industrial customers. To further reduce the line voltage to levels that are safe for residential use, step-down transformers are mounted on utility poles. These drop the voltage to the 120/240-V three-wire single-phase residential service discussed in Section 13.6. Industrial and commercial customers receive 460- and/or 208-V three-phase service.

Conclusion

Chapter 13 introduces the essential elements that permit the analysis of AC power systems. AC power is essential to all industrial activities and to the conveniences we are accustomed to in residential life. Virtually all engineers will be exposed to AC power systems in their Page 788careers, and the material presented in this chapter provides all the necessary tools to understand the analysis of AC power circuits. Upon completing this chapter, you should have mastered the following learning objectives:

1. *Understand the meaning of instantaneous and average power, master AC power notation, and compute average power for AC circuits. Compute the power factor of a complex load.* The power dissipated by a load in an AC circuit consists of the sum of an average and a fluctuating component. In practice, the average power is the quantity of interest.
2. *Learn complex power notation; compute apparent, real, and reactive power for complex loads. Draw the power triangle, and compute the capacitor size required to perform power factor correction on a load*. AC power can best be analyzed with the aid of complex notation. Complex power S is defined as the product of the phasor load voltage and the complex conjugate of the load current. The real part of S is the real power actually consumed by a load (that for which the user is charged); the imaginary part of S is called the reactive power and corresponds to energy stored in the circuit—it cannot be directly used for practical purposes. Reactive power is quantified by a quantity called the *power factor*, and it can be minimized through a procedure called *power factor correction*.

3. *Analyze the ideal transformer; compute primary and secondary currents and voltages and turns ratios. Calculate reflected sources and impedances across ideal transformers. Understand maximum power transfer*. Transformers find many applications in electrical engineering. One of the most common is in power transmission and distribution, where the electric power generated at electric power plants is stepped "up" and "down" before and after transmission, to improve the overall efficiency of electric power distribution.
4. *Learn three-phase AC power notation; compute load currents and voltages for balanced wye and delta loads*. AC power is generated and distributed in three-phase form. Residential services are typically single-phase (making use of only one branch of the three-phase lines) while industrial applications are often served directly by three-phase power.
5. *Understand the basic principles of residential electrical wiring, of electrical safety, and of the generation and distribution of AC power.*

HOMEWORK PROBLEMS

Section 13.1: Instantaneous and Average Power

13.1 The heating element in a soldering iron has a resistance of 20Ω. Find the average power dissipated in the soldering iron if it is connected to a voltage source of 90 V rms.

13.2 A coffeemaker has a rated power of 1,000 W at 240 V rms. Find the resistance of the heating element.

13.3 A current source $i(t)$ is connected to a 50-Ω resistor. Find the average power delivered to the resistor, given that $i(t)$ is

a. $7\cos 100t$ A

b. $7\cos(100t - 30°)$ A

c. $7\cos 100t - 3\cos(100t - 60°)$ A

d. $7\cos 100t - 3$A

13.4 Find the rms value of each of the following periodic currents:

a. $\cos 200t + 3\cos 200t$

b. $\cos 10t + 2\sin 10t$

c. $\cos 50t + 1$

d. $\cos 30t + \cos(30t + \pi/6)$

13.5 A current of 2.5 A through a neon light advertisement is supplied by a 115 V rms voltage source. The current lags the voltage by 30°. Find the impedance of the light, the real power dissipated by it, and its power factor.

13.6 Compute the average power dissipated by the load seen by the voltage source in Figure P13.6. Let $\omega = 377$ rad/s, $\bar{\mathbf{V}}_s = 50\angle 0$, $R = 10\ \Omega$, $L = 0.08$ H, and $C = 200\mu$F.

Page 789

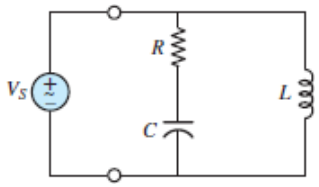


Figure P13.6

13.7 A drilling machine is driven by a single-phase induction machine connected to a 110 V rms supply. Assume that the machining operation requires 1 kW, that the tool machine has 90 percent efficiency, and that the supply current is 14 A rms with a power factor of 0.8. Find the AC machine efficiency.

13.8 Given the waveform of a voltage source shown in Figure P13.8, find:

a. The steady DC voltage that would cause the same heating effect across a resistance.

b. The average current supplied to a 10-Ω resistor connected across the voltage source.

c. The average power supplied to a 1-Ω resistor connected across the voltage source.

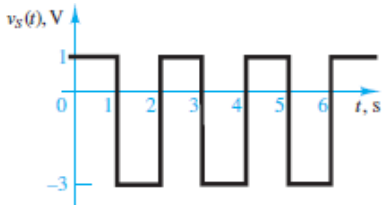


Figure P13.8

13.9 A current source $i(t)$ is connected to a 100-Ω resistor. Find the average power delivered to the resistor, given that $i(t)$ is:

a. $4\cos(100t)$ A

b. $4\cos(100t - 50°)$ A

c. $4\cos(100t - 3)\cos(100t - 50°)$ A

d. $4\cos(100t - 3)$ A

13.10 Find the rms value of each of the following periodic currents:

$\cos(377t) + \cos(377t)$ A

$\cos(2t) + \sin(2t)$ A

$\cos(377t) + 1$ A

$\cos(2t) + \cos(2t + 135°)$ A

$\cos(2t) + \cos(3t)$ A

Section 13.2: Complex Power

13.11 A current of 10 A rms results when a single-phase circuit is placed across a 220 V rms source. The current lags the voltage by 60°. Find the power dissipated by the circuit and the power factor.

13.12 A network is supplied by a 120 V rms, 60-Hz voltage source. An ammeter and a wattmeter indicate that 12 A rms is drawn from the source and 800 W are consumed by the network. Determine:

a. The network power factor.

b. The network phase angle.

c. The network impedance.

d. The equivalent resistance and reactance of the network.

13.13 For the following numeric values, determine the average power, P, the reactive power, Q, and the complex power, S, of the circuit shown in Figure P13.13. *Note:* Phasor quantities are rms.

a. $v_S(t) = 650\cos(377t)$ V
$i_o(t) = 20\cos(377t - 10°)$ A

b. $\tilde{\mathbf{V}}_S = 460\angle 0°$ V rms
$\tilde{\mathbf{I}}_o = 14.14\angle -45°$ A rms

c. $\tilde{\mathbf{V}}_S = 100\angle 0°$ V rms
$\tilde{\mathbf{I}}_o = 8.6\angle -86°$ A rms

d. $\tilde{\mathbf{V}}_S = 208\angle -30°$ V rms
$\tilde{\mathbf{I}}_o = 2.3\angle -63°$ A rms

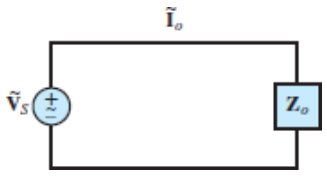


Figure P13.13

13.14 For the circuit of Figure P13.13, determine the power factor for the load $\mathbf{Z}_o$ and determine whether it is leading or lagging for the following conditions:

a. $v_S(t) = 679\cos(\omega t + 15°)$ V
$i_o(t) = 20\cos(\omega t + 47°)$ A

b. $v_S(t) = 163\cos(\omega t + 15°)$ V
$i_o(t) = 20\cos(\omega t - 22°)$ A

c. $v_S(t) = 294\cos(\omega t)$ V
$i_o(t) = 1.7\cos(\omega t + 175°)$ A

d. $\mathbf{Z}_o = (48 + j16)\ \Omega$

13.15 For the circuit of Figure P13.13, determine whether the load is capacitive or inductive, assuming:

Page 790

a. pf = 0.87 (leading)

b. pf = 0.42 (leading)

c. $v_S(t) = 42\cos(\omega t)$ V
$i_L(t) = 4.2\sin(\omega t)$ A

d. $v_S(t) = 10.4\cos(\omega t - 12°)$ V
$i_L(t) = 0.4\cos(\omega t - 12°)$ A

13.16 For the circuit shown in Figure P13.16, assume C = 265 μF, L = 25.55 mH, and R = 10 Ω. Find the instantaneous real and reactive power if:

a. $v_S(t) = 120\cos(377t)$ V (i.e., the frequency is 60 Hz)

b. $v_S(t) = 650\cos(314t)$ V (i.e., the frequency is 50 Hz)

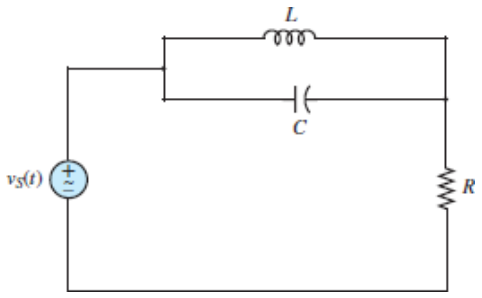


Figure P13.16

13.17 A load impedance, $\mathbf{Z}_o$ = 10 + j3 Ω, is connected to a source with line resistance equal to 1 Ω, as shown in Figure P13.17. Calculate the following values:

a. The average power delivered to the load.

b. The average power absorbed by the line.

c. The apparent power supplied by the generator.

d. The power factor of the load.

e. The power factor of line plus load.

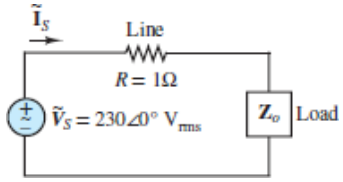


Figure P13.17

Section 13.3: Power Factor Correction

13.18 A single-phase motor draws 220 W at a power factor of 0.8 lagging when connected across a 240 V rms, 60-Hz source. A capacitor is connected in parallel with the load to produce a unity power factor. Determine the required capacitance.

13.19 The networks seen by the voltage sources in Figure P13.19 have unity power factor. Determine C_P and C_S.

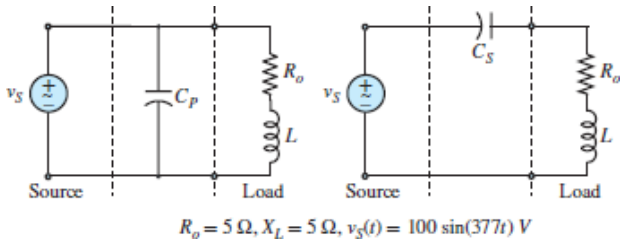


Figure P13.19

13.20 A 1,000-W electric motor is connected to a 120 V_{rms}, 60-Hz source. The power factor seen by the source is 0.8, lagging. To correct the pf to 0.95 lagging, a capacitor is placed in parallel with the motor. Calculate the current drawn from the source with and without the capacitor connected. Determine the value of the capacitor required to make the correction.

13.21 The motor inside a blender can be modeled as a resistance in series with an inductance, as shown in Figure P13.21. The wall socket source is modeled as

an ideal 120 V rms voltage source in series with a 2-Ω output resistance. Assume the source frequency is $\omega = 377$ rad/s.

a. What is the power factor of the motor?

b. What is the power factor seen by the voltage source?

c. What is the average power, P_{AV}, consumed by the motor?

d. What value of capacitor when placed in parallel with the motor will change the power factor seen by the voltage source to 0.9 lagging?

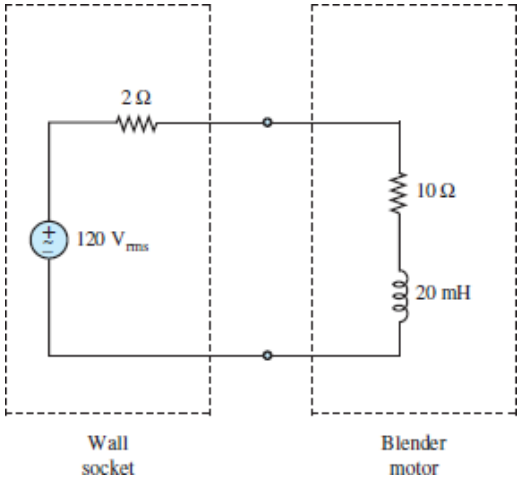


Figure P13.21

Page 791

13.22 For the circuit shown in Figure P13.22, find:

a. The Thévenin equivalent network seen by the load.

b. The power dissipated by the load resistor.

c. The load impedance that would result in maximum power transfer to the load.

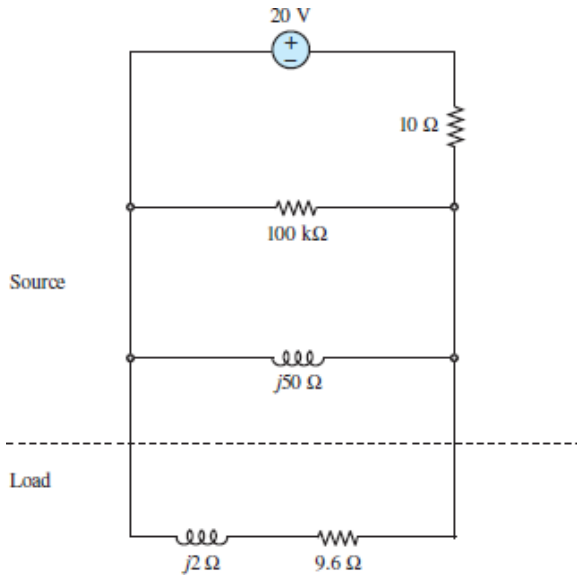


Figure P13.22

13.23 For the following numerical values, determine the capacitance to be placed in parallel with the load $\mathbf{Z}_o$ shown in Figure P13.13 that will result in a unity power factor seen by the voltage source. Assume $\omega = 377$ rad/s.

a. $\tilde{\mathbf{V}}_s = 300\angle 0$ V rms, $\tilde{\mathbf{I}}_o = 80\angle(-0.15\pi)$ A rms

b. $\tilde{\mathbf{V}}_s = 100\angle 0$ V rms, $\tilde{\mathbf{I}}_o = 30\angle(-\pi/4)$ A rms

c. $\tilde{\mathbf{V}}_s = 12\angle(-\pi/4)$ V rms, $\tilde{\mathbf{I}}_o = 3\angle(-\pi/2)$ A rms

13.24 For the circuit of Figure P13.13, determine the power factor of the load for each case listed below. Is it leading or lagging?

a. $v_S(t) = 50\cos(\omega t)$ V
$i_o(t) = 20\sin(\omega t + 1.2)$ A

b. $v_S(t) = 110\cos(\omega t + 0.1)$ V
$i_o(t) = 10\cos(\omega t - 0.1)$ A

c. $\mathbf{Z}_o = (20 + j5)\ \Omega$

d. $\mathbf{Z}_o = (20 - j5)\ \Omega$

13.25 For the circuit of Figure P13.13, determine whether the load $\mathbf{Z}_o$ is capacitive or inductive, if:

a. its power factor is pf = 0.76 lagging.

b. its power factor is pf = 0.5 (leading).

c. $v_s(t) = 10\cos(\omega t)$ V, $i_o(t) = \cos(\omega t)$ A.

d. $v_s(t) = 100\cos(\omega t)$ V, $i_o(t) = 12\cos(\omega t + \pi/4)$ A.

13.26 Find the real and reactive power supplied by the voltage source shown in Figure P13.26 for ω = 5 rad/s and ω = 15 rad/s. Let v_S = 15 cos (ωt) V, R = 5Ω, C = 0.1 F, L_1 = 1 H, L_2 = 2 H.

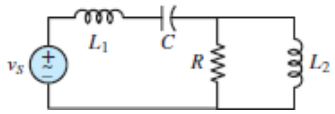


Figure P13.26

13.27 In Figure P13.27, assume $\tilde{\mathbf{V}}_{S1} = 10\angle -\pi/4$ V rms, $\tilde{\mathbf{V}}_{S2} = 12\angle 0.8$ V rms, $R_1 = 2\,\Omega$, $R_2 = 3\,\Omega$, $X_L = 4\,\Omega$, and $X_C = -4\,\Omega$. Find:

a. The amplitude of the current supplied by each source.

b. The total real power supplied by each source.

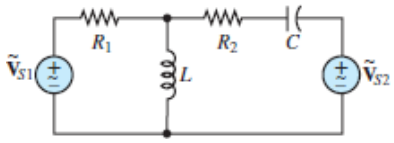


Figure P13.27

13.28 For the circuit shown in Figure P13.28, assume $f = 60$ Hz, $\tilde{\mathbf{V}}_S = 90\angle 0$ V rms, $R = 25\,\Omega$, $X_L = 70\,\Omega$, and $X_C = -8\,\Omega$. Calculate:

a. The capacitance C and the inductance L.

b. The power factor seen by the voltage source.

c. The new capacitance required to correct that power factor to unity.

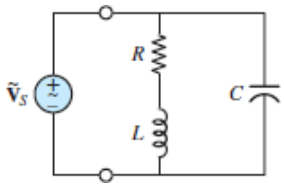


Figure P13.28

Page 792

13.29 The load $\mathbf{Z}_o$ shown in Figure P13.29 consists of a 20-Ω resistor in series with a 0.01-H inductor. Assuming $f = 60$ Hz, $R = 0.5\,\Omega$, $\tilde{\mathbf{V}}_S = 100\angle 0$ V rms. Calculate:

a. The apparent power supplied by the voltage source.

b. The apparent power delivered to the load.

c. The power factor of the load.

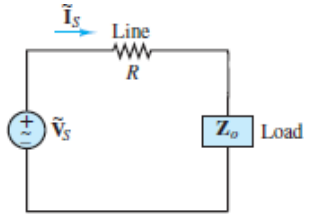


Figure P13.29

13.30 Calculate the real and reactive power of the load between terminals *a* and *b* in Figure P13.30. Assume $f = 60$ Hz, $\tilde{V}_S = 70\angle 0$ V rms, $R_S = 2\,\Omega$, $R_o = 18\ \Omega$, and $X_L = 5\Omega$.

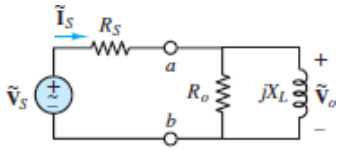


Figure P13.30

13.31 Calculate the apparent power, real power, and reactive power supplied by the voltage source shown in Figure P13.31. Draw the power triangle. Assume $f = 60$ Hz, $\tilde{V}_S = 70\angle 0$ V rms, $R = 18\ \Omega$, $C = 50\ \mu$F, and $L = 0.001$ H.

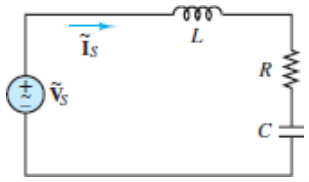


Figure P13.31

13.32 Refer to Problem 13.31 and determine the capacitance needed in parallel with the voltage source to correct the power factor seen by the source to 0.95. Draw the power triangle.

13.33 A single-phase motor is modeled as a resistor R in series with an inductor L as shown in Figure P13.33. The capacitor corrects the power factor between terminals *a* and *b* to unity. Assume the meters shown are ideal and $f = 50$ Hz, $V = 220$ V rms, $I = 20$ A rms, and $I_1 = 25$ A rms. Find the capacitor value.

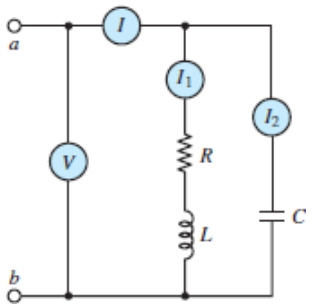


Figure P13.33

13.34 Suppose that the electricity in your home has gone out on a hot, humid summer day and the power company will not be able to fix the problem for several days. The freezer in the basement contains $300 worth of food that you cannot afford to let spoil. You would also like to keep one window air conditioner running, as well as run the refrigerator in your kitchen. When these appliances are on, they draw the following currents (all values are rms):

Air conditioner:	9.6 A rms @ 120 V rms
	pf = 0.90 lagging
Freezer:	4.2 A rms @ 120 V rms
	pf = 0.87 lagging
Refrigerator:	3.5 A rms @ 120 V rms
	pf = 0.80 lagging

In the worst-case scenario, how much power must an emergency generator supply?

13.35 The French TGV high-speed train absorbs 11 MW at 300 km/h (186 mi/h). The power supply module shown in Figure P13.35 consists of two 25-kV rms single-phase AC power stations connected at the same overhead line, one at each end of the module. For the return circuits, the rail is used. The train is also designed to operate at a low speed with 1.5-kV DC in railway stations or under the old electrification lines. The natural (average) power factor in the AC operation is 0.8. Assume that the equivalent specific resistance of the overhead line is 0.2 Ω/km and that the rail resistance can be neglected. Find:

a. A simple circuit model for the system.

b. The locomotive's current in the condition of a 10 percent voltage drop.

c. The reactive power supplied by the power stations.

d. The supplied real power, overhead line losses, and maximum distance between two power stations supplied in the condition of a 10 percent

voltage drop when the train is located at the half-distance between the stations.

Page 793

e. Overhead line losses in the condition of a 10 percent voltage drop when the train is located at the half-distance between the stations, assuming pf = 1. (The French TGV is designed with a state-of-the-art power compensation system.)

f. The maximum distance between the two power stations supplied in the condition of a 10 percent voltage drop when the train is located at the half-distance between the stations, assuming the DC (1.5-kV) operation at one-quarter power.

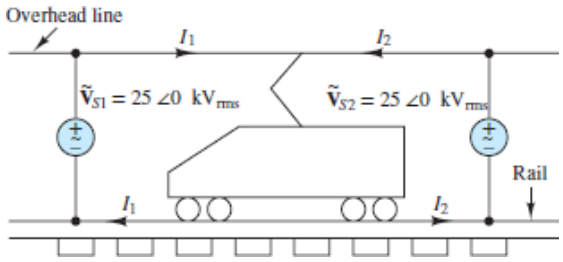


Figure P13.35

13.36 An industrial assembly hall is continuously lit by one hundred 40-W mercury vapor lamps in parallel and supplied by a 120 V rms, 60-Hz source. The power factor seen by the source is 0.65, which is so low that a 25 percent penalty is applied at billing. If the average price of 1 kWh is \$0.05 and the average cost of a capacitor is \$50 per mF, compute how long it will take before the billing penalty equals the cost of the capacitor needed to correct the power factor to 0.85.

13.37 Refer to Problem 13.36 and assume that each lamp is now available with a compensating capacitor in parallel with the original lamp. Find:

a. The compensating capacitor value for unity power factor seen by the source.

b. The maximum number of additional lamps that can be installed without exceeding the original current supplied by the source when using uncompensated lamps.

13.38 The voltage and current supplied by a source to a load are:

$$\tilde{\mathbf{V}}_s = 7\angle 0.873 \text{ V rms} \qquad \tilde{\mathbf{I}}_s = 13\angle(-0.349) \text{ A rms}$$

Determine:

a. The real power consumed as work and dissipated as heat in the load.

b. The reactive power stored in the load.

c. The impedance angle of the load and its power factor.

13.39 Determine the real power dissipated and the reactive power stored in each of the impedances shown in Figure P13.39. Assume:

$$\tilde{\mathbf{V}}_{s1} = \frac{170}{\sqrt{2}}\angle 0 \text{ V rms}$$

$$\tilde{\mathbf{V}}_{s2} = \frac{170}{\sqrt{2}}\angle \pi/2 \text{ V rms}$$

$$\omega = 377 \text{ rad/s}$$

$$\mathbf{Z}_1 = 0.7\angle\frac{\pi}{6}\ \Omega$$

$$\mathbf{Z}_2 = 1.5\angle 0.105\ \Omega$$

$$\mathbf{Z}_3 = 0.3 + j0.4\ \Omega$$

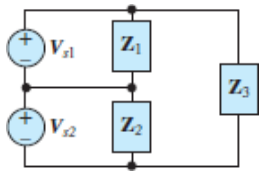


Figure P13.39

13.40 The following are supplied by a source to a load:

$$\tilde{\mathbf{V}}_s = 170\angle(-0.157) \text{ V rms} \qquad \tilde{\mathbf{I}}_s = 13\angle 0.28 \text{ A rms}$$

Determine:

a. The real power consumed as work and dissipated as heat in the load.

b. The reactive power stored in the load.

c. The impedance angle of the load and its power factor.

Section 13.4: Transformers

13.41 A center-tapped transformer has the schematic representation shown in Figure P13.41. The primary-side voltage is stepped down to two secondary-side voltages. Assume that each secondary supplies a 7-kW resistive load and that the primary is connected to 100 V rms. Find:

a. The primary power.

b. The primary current.

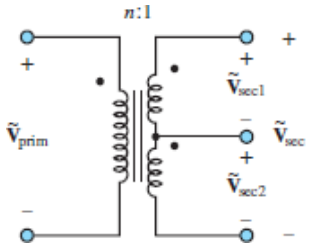


Figure P13.41

Page 794

13.42 A center-tapped transformer has the schematic representation shown in Figure P13.41. The primary-side voltage is stepped down to a secondary-side voltage $\tilde{\mathbf{V}}_{sec}$ by a ratio of n:1. On the secondary side, $\tilde{\mathbf{V}}_{sec1} = \tilde{\mathbf{V}}_{sec2} = \frac{1}{2}\tilde{\mathbf{V}}_{sec}$.

a. $\tilde{\mathbf{V}}_{prim} = 220\angle 0°$ V rms and $n = 11$, find $\tilde{\mathbf{V}}_{sec}$, $\tilde{\mathbf{V}}_{sec1}$, and $\tilde{\mathbf{V}}_{sec2}$.

b. What must n be if $\tilde{\mathbf{V}}_{prim} = 110\angle 0°$ V rms and we desire $|\tilde{\mathbf{V}}_{sec2}|$ to be 5 V rms?

13.43 For the circuit shown in Figure P13.43, assume that $\tilde{\mathbf{V}}_g = 80\angle 0$ V rms, $R_g = 2\ \Omega$, and $R_o = 12\ \Omega$. and R_o = 12Ω. Assume an ideal transformer. Find:

a. The equivalent resistance seen by the voltage source.

b. The power P_{source} supplied by the voltage source.

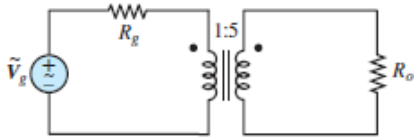


Figure P13.43

13.44 Refer to Problem 13.43 and find:

a. The power P_{load} consumed by R_o.

b. The installation efficiency P_{load}/P_{source}.

c. The load R_o that results in maximum power transfer to the load.

13.45 An ideal transformer is rated to deliver 460 kVA at 380 V rms to a customer, as shown in Figure P13.45.

a. How much current can the transformer supply to the customer?

b. If the customer's load is purely resistive (i.e., if pf = 1), what is the maximum power that the customer can receive?

c. If the customer's power factor is 0.8 lagging, what is the maximum usable power the customer can receive?

d. What is the maximum power if the pf is 0.7 lagging?

e. If the customer requires 300 kW to operate, what is the minimum power factor with the given size transformer?

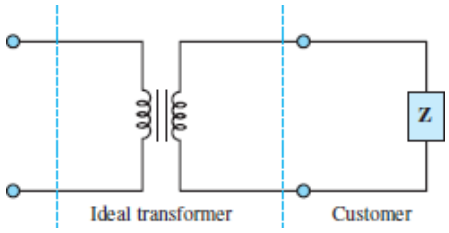


Figure P13.45

13.46 For the ideal transformer shown in Figure P13.46, assume $v_{in}(t) = 240\cos(377t)$ V, $R_{in} = 50\ \Omega$, $R_o = 20\ \Omega$, and the step-down turns ratio is set by $n = 3$. Determine:

a. The primary current i_{in}.

b. The secondary voltage v_o.

c. The secondary power $P_o = i_o^2 R_o = v_o^2/R_o$.

d. The installation efficiency P_{in}/P_o, where $P_{in} = i_{in}^2 R_{in}$.

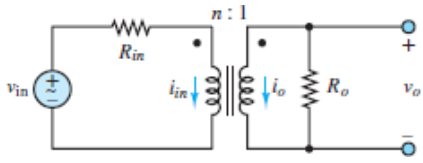


Figure P13.46

13.47 For Figure P13.47, assume the transformer is ideal. Find the step-down turns ratio $M = n$ that provides maximum power transfer to R_o. Let $R_{in} = 1{,}200\ \Omega$, $R_o = 100\ \Omega$, and $v_{in}(t) = V_{pk}\cos(\omega t)$.

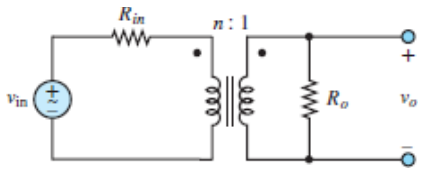


Figure P13.47

13.48 Consider the 8-Ω resistor shown in Figure P13.48 to be the load. Assume $\tilde{\mathbf{V}}_g = 110\angle 0\,\text{V rms}$ and a variable turns ratio n. What value of n results in maximum power (a) dissipated by the load? (b) supplied by the voltage source? What value of n results in maximum power transfer efficiency from source to load?

Page 795

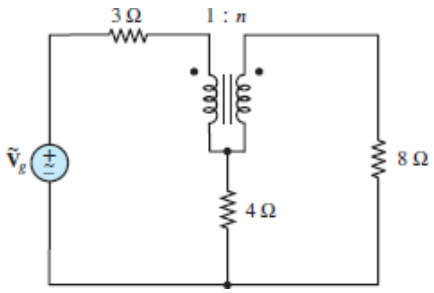


Figure P13.48

13.49 Assume the transformer shown in Figure P13.49 delivers 70 A rms at 90 V rms to a resistive load. What is the power transfer efficiency between voltage source and load? Let $R_s = 2\ \Omega, X_{C_1} = -10\ \Omega, X_{C_2} = -5\ \Omega$.

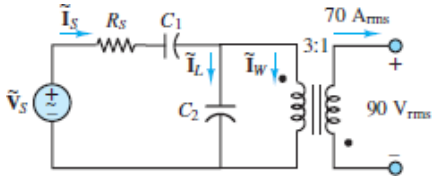


Figure P13.49

13.50 A method for determining the equivalent network of a nonideal transformer consists of two tests: the open-circuit test and the short-circuit test. The open-circuit test, shown in Figure P13.50(a), is usually done by applying rated voltage to the primary side of the transformer while leaving the secondary side open.

The current into the primary side is measured, as is the power dissipated. The short-circuit test, shown in Figure P13.50(b), is performed by increasing the primary voltage until rated current is going into the transformer while the secondary side is short-circuited. The current into the transformer, the applied voltage, and the power dissipated are measured.

The equivalent circuit of a transformer is shown in Figure P13.50(c), where r_w and L_w represent the winding resistance and inductance, respectively, and r_c and L_c represent the losses in the core of the transformer and the inductance of the core. The ideal transformer is also included in the model.

With the open-circuit test, we may assume that $\bar{\mathbf{I}}_{\text{primary}} = \bar{\mathbf{I}}_{\text{secondary}} = 0$. Then all the current that is measured is directed through the parallel combination of r_c and L_c. We also assume that $|r_c || j\omega L_c|$ is much greater than $r_w + j\omega L_w$. Using these assumptions and the open-circuit test data, we can find the resistance r_c and the inductance L_c.

In the short-circuit test, we assume that $\bar{\mathbf{V}}_{\text{secondary}}$ is zero, so that the voltage on the primary side of the ideal transformer is also zero, causing no current through the $r_c \,/\!/\, L_c$ parallel combination. Using this assumption with the short-circuit test data, we are able to find the resistance r_w and inductance L_w.

The following test data was measured by the meters indicated in Figure P13.50(a) and (b):

Open-circuit test:	$\bar{\mathbf{V}} = 241$ V rms
	$\bar{\mathbf{I}} = 0.95$ A rms
	$P = 32$ W
Short-circuit test:	$\bar{\mathbf{V}} = 5$ V rms
	$\bar{\mathbf{I}} = 5.25$ A rms
	$P = 26$ W

Both tests were made at ω = 377 rad/s. Use the data to determine the equivalent network of the nonideal transformer.

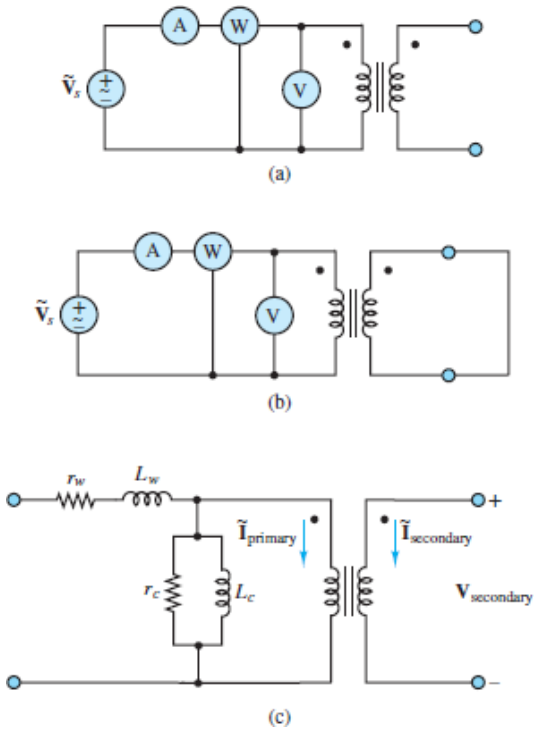


Figure P13.50

13.51 Use the methods outlined in Problem 13.50 and the following data to find the equivalent network of a nonideal transformer.

Open-circuit test:	$\tilde{\mathbf{V}} = 4{,}600$ V rms
	$\tilde{\mathbf{I}} = 0.7$ A rms
	$P = 200$ W
Short-circuit test:	$P = 50$ W
	$\tilde{\mathbf{V}} = 5.2$ V rms

The transformer is rated at 460 kVA, and tests are performed at 60 Hz.

Page 796

13.52 A method of thermal treatment for a steel pipe is to heat the pipe by the Joule effect, caused when a current is directed through the pipe. In most cases, a low-voltage, high-current transformer is used to deliver the current through the pipe. In this problem, we consider a single-phase transformer at 220 V rms, which delivers 1.2 V rms. Because of the pipe's resistance variation with temperature, a secondary voltage regulation is needed in the range of 10 percent, as shown in Figure P13.52. The voltage regulation is obtained with five different slots in the primary winding (high-voltage regulation). Assuming that the secondary coil has two turns, find the number of turns for each slot.

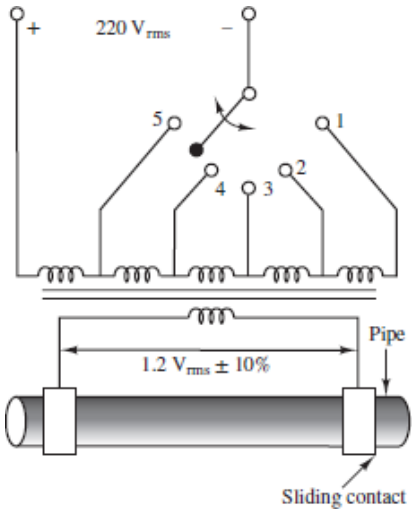


Figure P13.52

13.53 Refer to Problem 13.52 and assume a pipe resistance of $2 \times 10^{-4}\Omega$ and a secondary resistance (wire leads + slide contacts) of $5 \times 10^{-5}\Omega$. The primary current is 28.8 A rms and the power factor seen by the 220 V rms source is 0.91. Find:

a. The slot number.

b. The secondary reactance.

c. The power transfer efficiency.

13.54 A single-phase transformer used for street lighting (high-pressure sodium discharge lamps) converts 6 kV rms to 230 V rms with an efficiency of 0.95. Assuming the power factor seen by the high voltage source is 0.8 and the primary apparent power is 30 kVA, find:

a. The secondary current.

b. The transformer turns ratio N.

13.55 The transformer shown in Figure P13.55 has several sets of windings on the secondary side. The windings have the following turns ratios:

a. : $N = 1/15$

b. : $N = 1/4$

c. : $N = 1/12$

d. : $N = 1/18$

If $\tilde{\mathbf{V}}_{\text{primary}} = 120\angle 0^\circ$ V rms, find and draw the connections that will allow you to produce the following secondary voltages:

a. $24.67\angle 0°$Vrms

b. $36.67\angle 0°$Vrms

c. $18\angle 0°$Vrms

d. $54.67\angle 180°$Vrms

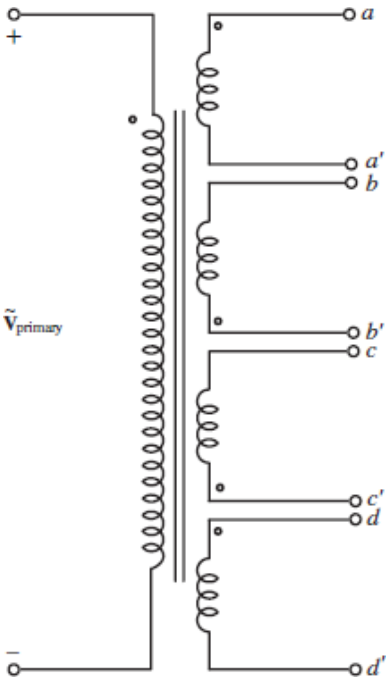


Figure P13.55

13.56 The circuit in Figure P13.56 shows the use of ideal transformers for impedance matching. You have a limited choice of turns ratios among available transformers. Suppose you can find transformers with turns ratios of 2:1, 7:2, 120:1, 3:2, and 6:1. If $\mathbf{Z}_o$ is $475\angle -25°\Omega$ and $\mathbf{Z}ab$ must be $267\angle -25°$, find the combination of transformers that will provide this Page 797impedance. (You may assume that polarities are easily reversed on these transformers.)

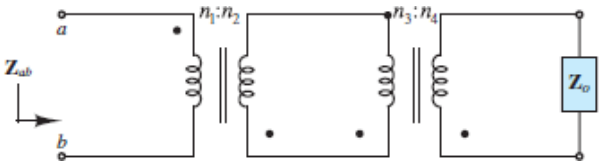


Figure P13.56

13.57 Before cable TV was generally available, TV networks broadcast their signals wirelessly. Large antennas were often installed atop residential homes to improve the reception of these signals. The impedance of the wire connecting the roof antenna to the TV set was typically 300 Ω, as shown in Figure P13.57(a). However, a typical TV had a 75-Ω impedance connection, as shown in Figure P13.57(b). To achieve maximum power transfer from the antenna to the television set, an ideal transformer was placed between the antenna and the TV, as shown in Figure P13.57(c). What is the turns ratio, $N = 1/n$, needed to obtain maximum power transfer?

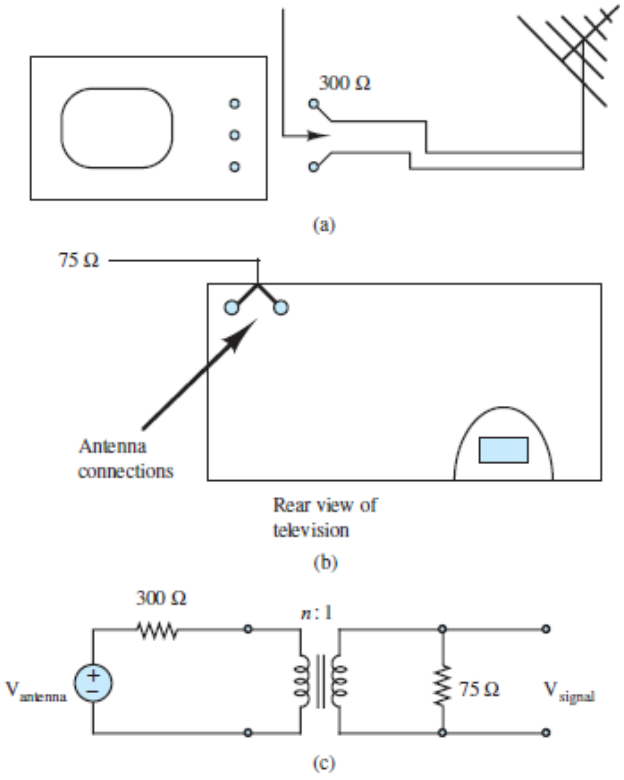


Figure P13.57

Section 13.5: Three-Phase Power

13.58 The magnitude of the phase voltage of a balanced three-phase wye system is 208 V rms. Express each phase and line voltage in both polar and rectangular coordinates.

13.59 The phase currents in a four-wire wye-connected load, such as that shown in Figure 13.49, are:

$$\tilde{\mathbf{I}}_{an} = 22\angle 0 \text{ A rms} \quad \tilde{\mathbf{I}}_{bn} = 10\angle \frac{2\pi}{3} \text{ A rms} \quad \tilde{\mathbf{I}}_{cn} = 15\angle \frac{\pi}{4}$$

Determine the current in the neutral wire.

13.60 Each voltage source shown in Figure P13.60 has a relative phase difference of $2\pi/3$.

a. Find $\tilde{\mathbf{V}}_{RW}$, $\tilde{\mathbf{V}}_{WB}$, and $\tilde{\mathbf{V}}_{BR}$, where $\tilde{\mathbf{V}}_{RW} = \tilde{\mathbf{V}}_R - \tilde{\mathbf{V}}_W$, $\tilde{\mathbf{V}}_{WB} = \tilde{\mathbf{V}}_W - \tilde{\mathbf{V}}_B$, and $\tilde{\mathbf{V}}_{BR} = \tilde{\mathbf{V}}_B - \tilde{\mathbf{V}}_R$.

b. Compare the results of part a with the calculations:

$$\tilde{\mathbf{V}}_{RW} = \tilde{\mathbf{V}}_R\sqrt{3}\angle(-\pi/6)$$
$$\tilde{\mathbf{V}}_{WB} = \tilde{\mathbf{V}}_W\sqrt{3}\angle(-\pi/6)$$
$$\tilde{\mathbf{V}}_{BR} = \tilde{\mathbf{V}}_B\sqrt{3}\angle(-\pi/6)$$

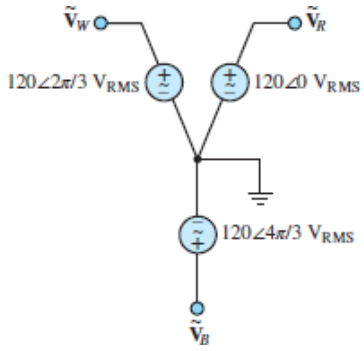


Figure P13.60

13.61 For the three-phase network shown in Figure P13.61, find the current in each wire and the real power consumed by the wye network. Let $\tilde{\mathbf{V}}_R = 110\angle 0$ V rms, $\tilde{\mathbf{V}}_W = 110\angle 2\pi/3$ V rms, and $\tilde{\mathbf{V}}_B = 110\angle 4\pi/3$ V rms. $R = 50\,\Omega$, $L = 120$ mH, $C = 133\,\mu$F, $f = 60$ Hz.

Page 798

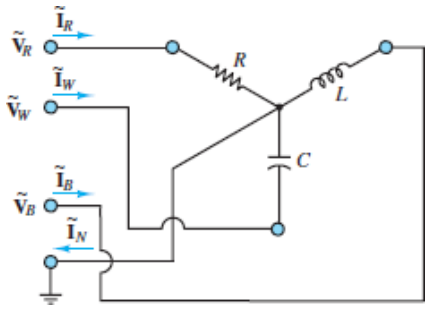


Figure P13.61

13.62 For the three-phase network shown in Figure P13.62, find the current in each wire and the real power consumed by the wye network. Let

$\tilde{\mathbf{V}}_R = 170\angle 0$ V rms, $\tilde{\mathbf{V}}_W = 170\angle 2\pi/3$ V rms, and $\tilde{\mathbf{V}}_B = 170\angle 4\pi/3$ V rms.

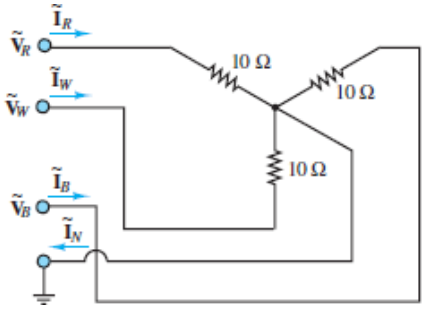


Figure P13.62

13.63 A three-phase steel-treatment electric oven has a phase resistance of 10 Ω and is connected at three-phase 380 V rms AC. Compute

a. The current through the resistors in wye and delta connections.

b. The power of the oven in wye and delta connections.

13.64 A naval in-board synchronous generator has an apparent power of 50 kVA and supplies a three-phase network of 380 V rms. Compute the phase currents, the real power, and the reactive power if:

a. The power factor is 0.85.

b. The power factor is 1.

13.65 The three-phase circuit shown in Figure P13.65 has a balanced wye source but an unbalanced wye load.

$$v_{s1} = 170\cos(\omega t)\ \text{V}$$
$$v_{s2} = 170\cos(\omega t + 2\pi/3)\ \text{V}$$
$$v_{s3} = 170\cos(\omega t - 2\pi/3)\ \text{V}$$
$$f = 60\ \text{Hz} \qquad \mathbf{Z}_1 = 0.5\angle 20^\circ\ \Omega$$
$$\mathbf{Z}_2 = 0.35\angle 0^\circ\ \Omega \qquad \mathbf{Z}_3 = 1.7\angle(-90^\circ)\ \Omega$$

Determine the current through $\mathbf{Z}_1$, using the following methods:

a. Mesh analysis.

b. Superposition.

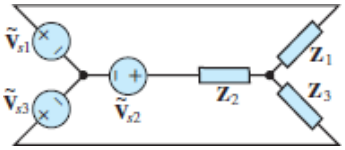


Figure P13.65

13.66 Determine the current through R shown in Figure P13.66. Assume: $\tilde{\mathbf{V}}_1 = 150\angle 0\,\text{V rms}$, $\tilde{\mathbf{V}}_2 = 150\angle 2\pi/3\,\text{V rms}$, $\tilde{\mathbf{V}}_3 = 150\angle 4\pi/3\,\text{V rms}$, $f = 300$ Hz, $R = 80\,\Omega$, $C = 0.3\,\mu$F, and $L = 80$ mH.

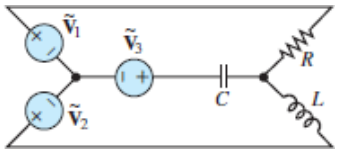


Figure P13.66

13.67 The circuit of Figure P13.67 has a balanced three-phase wye source but an unbalanced delta load. Determine the current through each impedance.

$v_1(t) = 170\cos(\omega t)$ V
$v_2(t) = 170\cos(\omega t + 2\pi/3)$ V
$v_3(t) = 170\cos(\omega t - 2\pi/3)$ V
$f = 60$ Hz $\quad \mathbf{Z}_1 = 3\angle 0\,\Omega$
$\mathbf{Z}_2 = 7\angle \pi/2\,\Omega \quad \mathbf{Z}_3 = -j11\,\Omega$

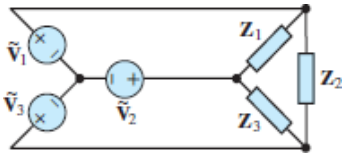


Figure P13.67

Page 799

13.68 If we model each winding of a three-phase motor like the circuit shown in Figure P13.68(a) and connect the windings as shown in Figure P13.68(b), we have the three-phase circuit shown in Figure P13.68(c). The motor can be constructed so that $R_1 = R_2 = R_3$ and $L_1 = L_2 = L_3$, as is the usual case. If we connect the motor as shown in Figure P13.68(c), find the currents $\bar{\mathbf{I}}_R$, $\bar{\mathbf{I}}_W$, $\bar{\mathbf{I}}_B$, and $\bar{\mathbf{I}}_N$, assuming that the resistances are 40 Ω each and each inductance is 5 mH. The frequency of each source is 60 Hz.

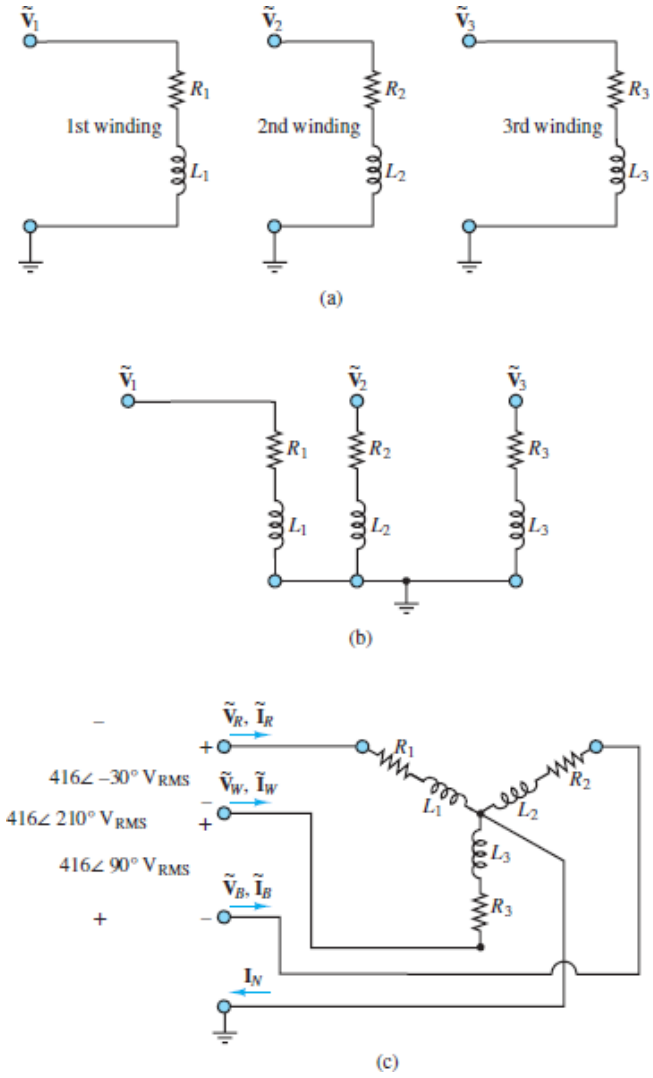


Figure P13.68

13.69 With reference to the motor of Problem 13.67,

a. How much power (in watts) is delivered to the motor?

b. What is the motor's power factor?

c. Why is it common in industrial practice *not* to connect the ground lead to motors of this type?

13.70 In general, a three-phase induction motor is designed for wye connection operation. However, for short-time operation, a delta connection can be used at the nominal wye voltage. Find the ratio between the power delivered to the same motor in the wye and delta connections.

13.71 A residential four-wire system supplies power at 240 V rms to the following single-phase appliances: On the first phase, there are ten 60-W bulbs. On the second phase, there is a 1-kW vacuum cleaner with a power factor of 0.9. On the third phase, there are ten 23-W compact fluorescent lamps with a power factor of 0.61. Find:

a. The current in the neutral wire.

b. The real, reactive, and apparent power for each phase.

13.72 The electric power company is concerned with the loading of its transformers. Since it is responsible for a large number of customers, it must be certain that it can supply the demands of *all* customers. The power company's transformers will deliver rated kVA to the secondary load. However, if the demand increased to a point where greater than rated current were required, the secondary voltage would have to drop below rated value. Also, the current would increase, and with it the I^2R losses (due to winding resistance), possibly causing the transformer to overheat. Unreasonable current demand could be caused, for example, by excessively low power factors at the load.

The customer, on the other hand, is not greatly concerned with an inefficient power factor, provided that sufficient power reaches the load. To make the customer more aware of power factor considerations, the power company may install a penalty on the customer's bill. A typical penalty–power factor chart is shown in Table 13.3. Power factors below 0.7 are not permitted. A 25 percent penalty will be applied to any billing after two consecutive months in which the customer's power factor has remained below 0.7.

Table 13.3

Power factor	Penalty
0.850 and higher	None
0.8 to 0.849	1%
0.75 to 0.799	2%
0.7 to 0.749	3%

Courtesy of Detroit Edison.

The wye-wye circuit shown in Figure P13.72 is representative of a three-phase motor load.

a. Find the total power supplied to the motor.

b. Find the power converted to mechanical energy if the motor is 80 percent efficient.

c. Find the power factor.

d. Does the company risk facing a power factor penalty on its next bill if all the motors in the factory are similar to this one?

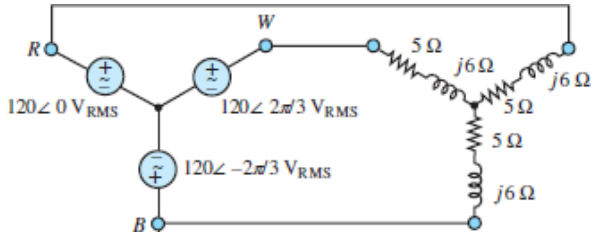


Figure P13.72

13.73 To correct the power factor problems of the motor in Problem 13.72, the company has decided to install capacitors as shown in Figure P13.73.

a. What capacitance must be installed to achieve a unity power factor if the line frequency is 60 Hz?

b. Repeat part a if the power factor is to be 0.85 lagging.

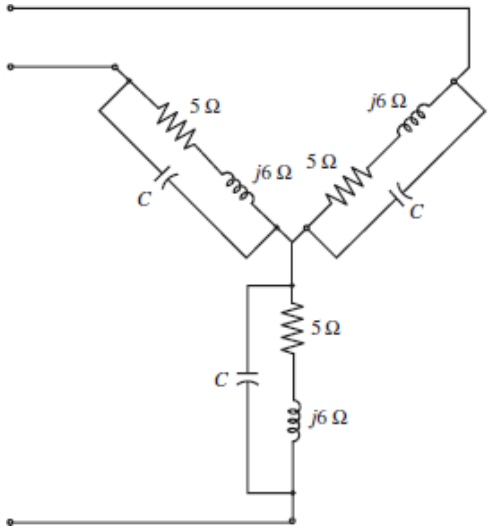


Figure P13.73

13.74 Find the apparent power and the real power delivered to the load in the Y-Δ circuit shown in Figure P13.74. What is the power factor?

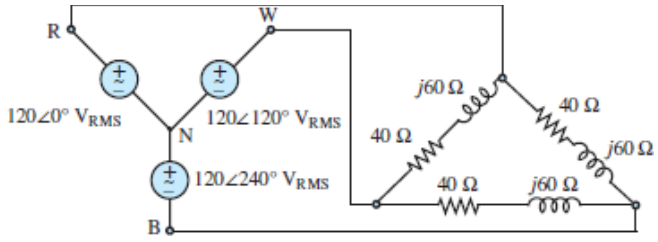


Figure P13.74

13.75 The circuit shown in Figure P13.75 is a Y-Δ-Y connected three-phase circuit. The primaries of the transformers are wye-connected, the secondaries are delta-connected, and the load is wye-connected. Find the currents $\tilde{\mathbf{I}}_{RP}$, $\tilde{\mathbf{I}}_{WP}$, $\tilde{\mathbf{I}}_{BP}$, $\tilde{\mathbf{I}}_{A}$, $\tilde{\mathbf{I}}_{B}$, and $\tilde{\mathbf{I}}_{C}$.

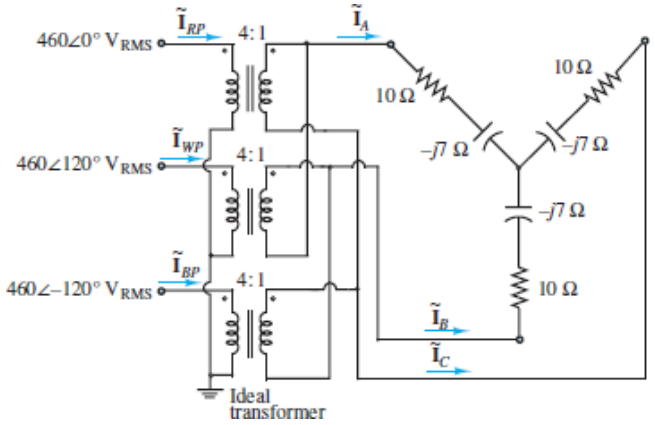


Figure P13.75

13.76 A three-phase motor is modeled by the wye-connected circuit shown in Figure P13.76. At $t = t_1$, a line fuse is blown (modeled by the switch). Find the line currents $\tilde{\mathbf{I}}_{R}$, $\tilde{\mathbf{I}}_{W}$, and $\tilde{\mathbf{I}}_{B}$ and the power dissipated by the motor in the following conditions:

a. $t \ll t_1$

b. $t \gg t_1$

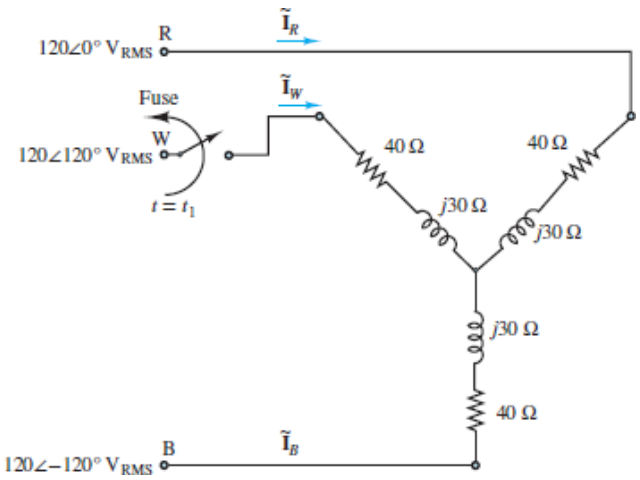


Figure P13.76

13.77 For the circuit shown in Figure P13.77, find the currents $\tilde{\mathbf{I}}_A, \tilde{\mathbf{I}}_B, \tilde{\mathbf{I}}_C$ and $\tilde{\mathbf{I}}_N$, and the real power dissipated by the load.

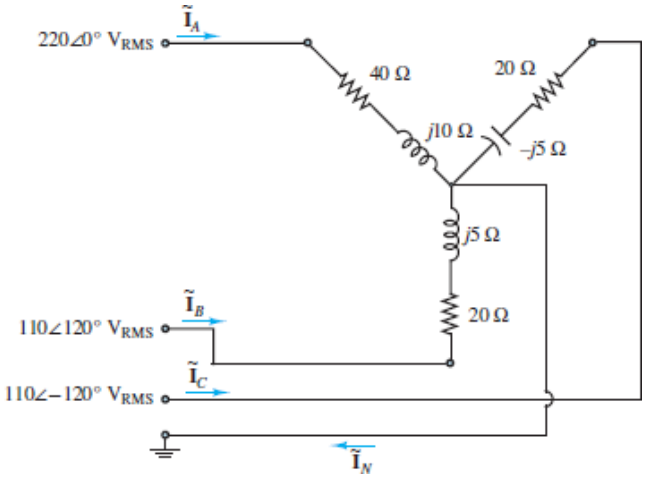


Figure P13.7

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

C H A P T E R

14

PRINCIPLES OF ELECTROMECHANICS

The objective of this chapter is to introduce the fundamental notions of electromechanical energy conversion, leading to an understanding of the operation of various electromechanical transducers. The chapter also serves as an introduction to the material on electric machines to be presented in Chapter 15. The foundations for the material introduced in this chapter may be found in the circuit analysis chapters (1 through 6).

The subject of electromechanical energy conversion is one that should be of particular interest to the *non–electrical* engineer, because it forms one of the important points of contact between electrical engineering and other engineering disciplines. Electromechanical transducers are commonly used in the design of industrial and aerospace control systems and in biomedical applications, and they form the basis of many common appliances. In the course of our exploration of electromechanics, we illustrate the operation of practical devices, such as loudspeakers, relays, solenoids, and sensors for the measurement of position and velocity.

Learning Objectives

Students will learn to...

1. Review the basic principles of electricity and magnetism. *Section 14.1*.
2. Use the concepts of reluctance and magnetic circuit equivalents to compute magnetic flux and currents in simple magnetic structures. *Section 14.2*.
3. Understand the properties of magnetic materials and their effects on magnetic circuit models. *Section 14.3*.
4. Use magnetic circuit models to analyze transformers. *Section 14.4*.
5. Model and analyze force generation in electromagnetomechanical systems. Analyze moving-iron transducers (electromagnets, solenoids, relays) and moving-coil transducers (electrodynamic shakers, loudspeakers, and seismic transducers). *Section 14.5*.

14.1 ELECTRICITY AND MAGNETISM

The notion that the phenomena of electricity and magnetism are interconnected was first proposed in the early 1800s by H. C. Oersted, a Danish physicist. Oersted showed that an electric current produces magnetic effects (more specifically, a magnetic field). Soon after, the French scientist André Marie Ampère expressed this relationship by means of a precise formulation known as *Ampère's law.* A few years later, the English scientist Faraday illustrated how the converse of Ampère's law also holds true, that is, that a magnetic field can generate an electric field; in short, *Faraday's law* states that a changing magnetic field gives rise to a voltage.

As is explained in the next few sections, the magnetic field forms a necessary connection between electrical and mechanical energy. Ampère's and Faraday's laws formally illustrate the relationship between electric and magnetic fields, but it should already be evident from your own individual experience that the magnetic field can also convert magnetic energy to mechanical energy (e.g., by lifting a piece of iron with a magnet). In effect, the devices we commonly refer to as *electromechanical* should more properly be referred to as electro*magneto*mechanical, since they almost invariably operate through a conversion from electrical to mechanical energy (or vice versa) by means of a magnetic field. Chapters 14 and 15 are concerned with the use of electricity and magnetic materials for the purpose of converting electrical to mechanical energy, and back.

The Magnetic Field and Faraday's Law

The quantities used to quantify the strength of a magnetic field are the **magnetic flux** ϕ, in units of **webers** (Wb); and the **magnetic flux density B**, in units of webers per square meter (Wb/m^2), or **teslas** (T). The latter quantity and the associated **magnetic field intensity H** (in units of amperes per meter, or A/m) are vectors.[1] Thus, the density of the magnetic flux and its intensity are in general described in Page 805vector form, in terms of the components present in each spatial direction (e.g., on the *x, y,* and *z* axes). In discussing magnetic flux density and field intensity in this chapter and Chapter 15, we almost always assume that the field is a *scalar field*—that is, that it lies in a single spatial direction. This will simplify many explanations.

It is customary to represent the magnetic field by means of the familiar *lines of force* (a concept also due to Faraday); we visualize the strength of a magnetic field by observing the density of these lines in space. You probably know from a previous course in physics that such lines are closed in a magnetic field; that is, they form continuous loops exiting at a magnetic north pole (by definition) and entering at a magnetic south pole. The relative strengths of the magnetic fields generated by two magnets could be depicted as shown in Figure 14.1.

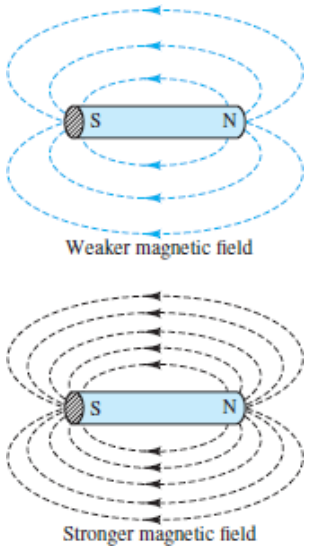


Figure 14.1 Lines of force in a magnetic field

Magnetic fields are generated by electric charge in motion, and their effect is measured by the force they exert on a moving charge. As you may recall from previous physics courses, the vector force **f** exerted on a charge of q moving at velocity **u** in the presence of a magnetic field with flux density **B** is given by

$$\mathbf{f} = q\mathbf{u} \times \mathbf{B} \tag{14.1}$$

where the symbol × denotes the (vector) cross product. If the charge is moving at a velocity **u** in a direction that makes an angle θ with the magnetic field, then the magnitude of the force is given by

$$f = quB \sin\theta \tag{14.2}$$

and the direction of this force is at right angles with the plane formed by the vectors **B** and **u.** This relationship is depicted in Figure 14.2.

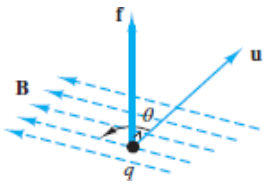


Figure 14.2 Charge moving in a constant magnetic field

The magnetic flux ϕ is then defined as the integral of the flux density over some surface area. For the simplified (but often useful) case of magnetic flux lines perpendicular to a cross-sectional area A, the flux is given by:

$$\phi = \int_A B\,dA \tag{14.3}$$

in webers, where the subscript A indicates that the integral is evaluated over surface A. Furthermore, if the flux were to be uniform over the cross-sectional area A (a useful simplification), the preceding integral is approximated by:

LO

$$\boxed{\phi = B \cdot A} \tag{14.4}$$

Figure 14.3 illustrates this idea, by showing hypothetical magnetic flux lines traversing a surface, delimited in the figure by a thin conducting wire.

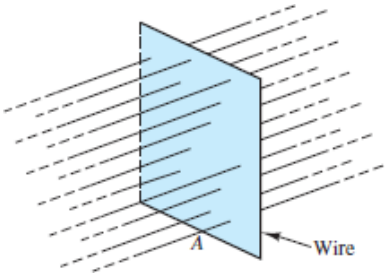
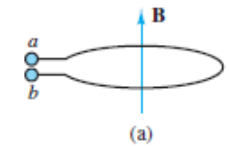


Figure 14.3 Magnetic flux lines crossing a surface bounded by a thin conducting wire.

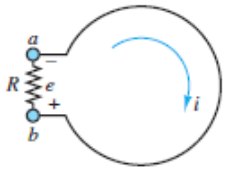
Faraday's law states that if the imaginary surface A were bounded by a conductor—for example, the thin wire of Figure 14.3—then a *changing* magnetic field would induce a voltage, and therefore a current, in the conductor. More precisely, Faraday's law states that a time-varying flux causes an induced **electromotive force**, or **emf**, e as follows:

$$e = -\frac{d\phi}{dt} \tag{14.5}$$

A little discussion is necessary at this point to explain the meaning of the minus sign in equation 14.5. Consider the one-turn coil of Figure 14.4, which forms Page 806a circular cross-sectional area, in the presence of a magnetic field with flux density **B** oriented in a direction perpendicular to the plane of the coil. If the magnetic field, and therefore the flux within the coil, is constant, no voltage will exist across terminals a and b; if, however, the flux were increasing and terminals a and b were connected—for example, by means of a resistor, as indicated in Figure 14.4(b)—current would be generated in the coil such that *the magnetic flux generated by the current would oppose the increasing flux*. Thus, the flux induced by such a current would be in the direction opposite to that of the original flux density vector **B**. This principle is known as **Lenz's law**. The reaction flux would then point downward in Figure 14.4(a), or into the page in Figure 14.4(b). Now, by virtue of the **right-hand rule**, this reaction flux would induce a current clockwise in Figure 14.4(b), that is, a current out of terminal b and into terminal a. The resulting voltage across the hypothetical resistor R would then be negative. If, on the other hand, the original flux were decreasing, current would be induced in the coil so as to reestablish the initial flux; but this would mean that the current would have to generate a flux in the upward direction in Figure 14.4(a) [or out of the page in Figure 14.4(b)]. Thus, the resulting voltage would change sign.



(a)



Current generating a magnetic flux opposing the increase in flux due to **B**

(b)

Figure 14.4 Flux direction

The polarity of the induced voltage can usually be determined from physical considerations; therefore the minus sign in equation 14.5 can be left out. We use this convention throughout the chapter.

In practical applications, the size of the voltages induced by the changing magnetic field can be significantly increased if the conducting wire is coiled so as to multiply the area crossed by the magnetic flux lines many times over. For an N-turn coil with cross-sectional area A, for example, we have the emf

$$e = N\frac{d\phi}{dt} \qquad \text{(minus sign is understood)} \tag{14.6}$$

CHECK YOUR UNDERSTANDING

A coil having 100 turns is immersed in a magnetic field that is varying uniformly from 80 to 30 mWb in 2 s. Find the induced voltage in the coil.

Answer: $e = -2.5$ V

Page 807

Figure 14.5 shows an N-turn coil *linking* a certain amount of magnetic flux; you can see that if N is very large and the coil is tightly wound (as is usually the case in the construction of practical devices), it is not unreasonable to presume that each turn of the coil links the same flux. It is convenient, in practice, to define the **flux linkage** λ as

$$\lambda = N\phi \tag{14.7}$$

so that

LO

$$\boxed{e = \frac{d\lambda}{dt}} \tag{14.8}$$

Note that equation 14.8, relating the derivative of the flux linkage to the induced emf, is analogous to the equation describing current as the derivative of charge:

$$i = \frac{dq}{dt} \tag{14.9}$$

In other words, flux linkage can be viewed as the dual of charge in circuit analysis provided that we are aware of the simplifying assumptions just stated in the preceding paragraphs, namely, a uniform magnetic field perpendicular to the area delimited by a tightly wound coil. These assumptions are not at all unreasonable when applied to the inductor coils commonly employed in electric circuits.

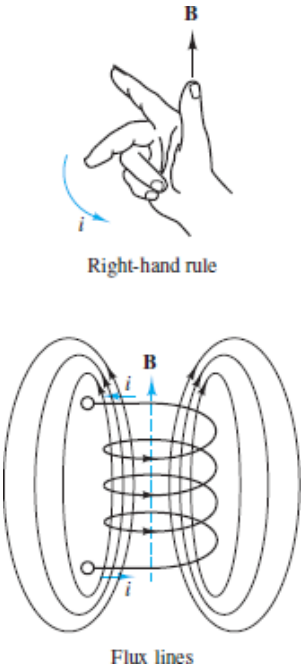


Figure 14.5 Concept of flux linkage

What, then, are the physical mechanisms that can cause magnetic flux to change, and therefore to induce an electromotive force? Two such mechanisms are possible. The first consists of physically moving a permanent magnet in the vicinity of a coil, for example, so as to create a time-varying flux. The second requires a time-varying current to produce a time-varying magnetic field. The latter method is more practical in many circumstances, since it does not require the use of permanent magnets and allows variation of field strength by varying the applied current; however, the former method is conceptually simpler to visualize. The voltages induced by a moving magnetic field are called **motion voltages**; those generated by a time-varying magnetic field are termed **transformer voltages**. We are interested in both in this chapter, for different applications.

In the analysis of linear circuits, as in Chapter 3, it is assumed that the relationship between flux linkage and current is linear:

$$\lambda = Li \tag{14.10}$$

so that the effect of a time-varying current is to induce a transformer voltage across an inductor coil, according to the expression

$$v = L\frac{di}{dt} \tag{14.11}$$

This is, in fact, the defining equation for the ideal **self-inductance** L. In addition to self-inductance, however, it is important to consider the **magnetic coupling** that can occur between neighboring circuits. Self-inductance measures the voltage induced in a circuit by the magnetic field generated by a current flowing in the same circuit. It is also possible that a second circuit in the vicinity of the first may experience an induced voltage as a consequence of the magnetic field generated in the first circuit. As explained in Section 14.4, this principle underlies the operation of all transformers.

Page 808

Self-Inductance and Mutual Inductance

Figure 14.6 depicts a pair of coils, one of which, L_1, is excited by a current i_1 and therefore develops a magnetic field and a resulting induced voltage v_1. The second coil, L_2, is not energized by a current, but links some of the flux generated by current i_1 around L_1 because of its close proximity to the first coil. The magnetic coupling between the coils established by virtue of their proximity is described by a quantity called **mutual inductance** and defined by the symbol M. The mutual inductance is defined by the equation

$$v_2 = M\frac{di_1}{dt} \tag{14.12}$$

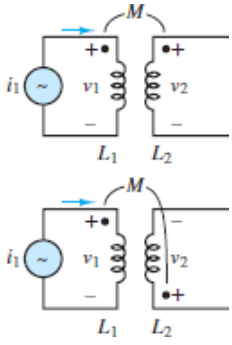


Figure 14.6 Mutual inductance

The dots shown in the two drawings indicate the polarity of the coupling between the coils. If the dots are at the same end of the coils, the voltage induced in coil 2 by a current in coil 1 has the same polarity as the voltage induced by the same current in coil 1; otherwise, the voltages are in opposition, as shown in the lower part of Figure 14.6. Thus, the presence of such dots indicates that magnetic coupling is present between two coils. It should also be pointed out that if a current (and therefore a magnetic field) were present in the second coil, an additional voltage would be induced across coil 1. The voltage induced across a coil is, in general, equal to the sum of the voltages induced by self-inductance and mutual inductance.

In practical electromagnetic circuits, the self-inductance of a circuit is not necessarily constant; in particular, the inductance parameter L is not constant, in general, but depends on the strength of the magnetic field intensity, so that it will not be possible to use such a simple relationship as $v = L\, di/dt$, with L constant. If we revisit the definition of the transformer voltage

$$e = N\frac{d\phi}{dt} \tag{14.13}$$

we see that in an inductor coil, the inductance is given by

$$L = \frac{N\phi}{i} = \frac{\lambda}{i} \tag{14.14}$$

This expression implies that the relationship between current and flux in a magnetic structure is linear if the inductance L is constant (the inductance being the slope of the line). In fact, the properties of ferromagnetic materials are such that the flux–current relationship is nonlinear, so that the simple linear inductance parameter used in electric circuit analysis is not adequate to represent the behavior of the magnetic circuits of this chapter. In any practical situation, the relationship between the flux linkage λ and the current is nonlinear, and might be described by a curve similar to that shown in Figure 14.7. Whenever the i-λ curve is not a straight line, it is more convenient to analyze the magnetic system in terms of energy calculations, since the corresponding circuit equation would be nonlinear.

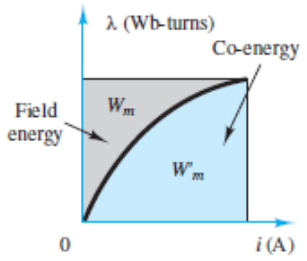


Figure 14.7 Relationship between flux linkage, current, energy, and co-energy

In a magnetic system, the energy stored in the magnetic field is equal to the integral of the instantaneous power, which is the product of voltage and current, just as in a conventional electric circuit:

$$W_m = \int ei \, dt' \tag{14.15}$$

Page 809

However, in this case, the voltage corresponds to the induced emf, according to Faraday's law,

$$e = \frac{d\lambda}{dt} = N\frac{d\phi}{dt} \tag{14.16}$$

and is therefore related to the rate of change of the magnetic flux. The energy stored in the magnetic field could therefore be expressed in terms of the current by:

$$W_m = \int ei \, dt' = \int \frac{d\lambda}{dt} i \, dt' = \int i \, d\lambda' \tag{14.17}$$

It should be straightforward to recognize that this energy is equal to the area above the λ-i curve of Figure 14.7. From the same figure, it is also possible to define a fictitious (but useful) quantity called **co-energy**, equal to the area under the curve and identified by the symbol W'_m. . From the figure, it is also possible to see that the co-energy can be expressed in terms of the stored energy by:

$$W'_m = i\lambda - W_m \tag{14.18}$$

Example 14.1 illustrates the calculation of energy, co-energy, and induced voltage, using the concepts developed in these paragraphs.

The calculation of the energy stored in the magnetic field around a magnetic structure will be particularly useful later in the chapter when the discussion turns to practical electromechanical transducers and it will be necessary to actually compute the forces generated in magnetic structures.

EXAMPLE 14.1 Energy and Co-Energy Calculation for an Inductor

Problem

Compute the energy, co-energy, and incremental linear inductance for an iron-core inductor with a given λ-*i* relationship. Also compute the voltage across the terminals, given the current through the coil.

Solution

Known Quantities: λ-*i* relationship; nominal value of λ; coil resistance; coil current.

Find: W_m; W'_m; L_Δ; v.

Schematics, Diagrams, Circuits, and Given Data: $i = (\lambda + 0.5\lambda^2)$ A; $\lambda_0 = 0.5$ V-s; $R = 1\ \Omega$; $i(t) = 0.625 + 0.01 \sin(400t)$.

Assumptions: Assume that the magnetic equation can be linearized, and use the linear model in all circuit calculations.

Analysis:

1. *Calculation of energy and co-energy.* From equation 14.17, we calculate the energy as follows.

$$W_m = \int_0^{\lambda} i(\lambda')\,d\lambda' = \frac{\lambda^2}{2} + \frac{\lambda^3}{6}$$

Page 810

The above expression is valid in general; in our case, the inductor is operating at a nominal flux linkage $\lambda_0 = 0.5$ V-s, and we can therefore evaluate the energy to be

$$W_m(\lambda = \lambda_0) = \left(\frac{\lambda^2}{2} + \frac{\lambda^3}{6}\right)\Big|_{\lambda=0.5} = 0.1458 \text{ J}$$

Thus, after equation 14.18, the co-energy is given by

$$W'_m = i\lambda - W_m$$

where

$$i = \lambda + 0.5\lambda^2 = 0.625 \text{ A}$$

and

$$W'_m = i\lambda - W_m = (0.625)(0.5) - (0.1458) = 0.1667 \text{ J}$$

2. *Calculation of incremental inductance.* If we know the nominal value of flux linkage (i.e., the operating point), we can calculate a linear inductance $L\Delta$, valid around values of λ close to the operating point λ_0. This incremental inductance is defined by the expression

$$L_\Delta = \left(\frac{di}{d\lambda}\right)^{-1}\Big|_{\lambda=\lambda_0}$$

and can be computed to be

$$L_\Delta = \left(\frac{di}{d\lambda}\right)^{-1}\Big|_{\lambda=\lambda_0} = (1+\lambda)^{-1}\Big|_{\lambda=\lambda_0} = \frac{1}{1+\lambda}\Big|_{\lambda=0.5} = 0.667 \text{ H}$$

The above expressions can be used to analyze the circuit behavior of the inductor when the flux linkage is around 0.5 V-s, or, equivalently, when the current through the inductor is around 0.625 A.

3. *Circuit analysis using linearized model of inductor.* We can use the incremental linear inductance calculated above to compute the voltage across the inductor in the presence of a current $i(t) = 0.625 + 0.01 \sin(400t)$. Using the basic circuit definition of an inductor with series resistance R, the voltage across the inductor is given by

$$\begin{aligned} v &= iR + L_\Delta \frac{di}{dt} = [0.625 + 0.01 \sin(400t)] \times 1 + 0.667 \times 4 \cos(400t) \\ &= 0.625 + 0.01 \sin(400t) + 2.668 \cos(400t) \\ &= 0.625 + 2.668 \sin(400t + 89.8°) \quad \text{V} \end{aligned}$$

Comments: The linear approximation in this case is not a bad one: the small sinusoidal current is oscillating around a much larger average current. In this type of situation, it is reasonable to assume that the inductor behaves linearly. This example explains why the linear inductor model introduced in Chapter 3 is an acceptable approximation in most circuit analysis problems.

CHECK YOUR UNDERSTANDING

The relation between the flux linkages and the current for a magnetic material is given by $\lambda = 6i/(2i + 1)$ Wb-turns. Determine the energy stored in the magnetic field for $\lambda = 2$ Wb-turns.

Answer: $W_m = 0.648$ J

FOCUS ON MEASUREMENTS

Linear Variable Differential Transformer

The linear variable differential transformer (LVDT) is a displacement transducer based on the mutual inductance concept just discussed. Figure 14.8 represents an LVDT as a primary coil subject to AC excitation (v_{ex}) and of a pair of identical secondary coils, which are connected so that:

$$v_{out} = v_1 - v_2$$

The ferromagnetic core between the primary and secondary coils can be displaced in proportion to some external motion x and determines the magnetic coupling between primary and secondary coils. Intuitively, as the core is displaced upward, greater coupling will occur between the primary coil and the top secondary coil, thus inducing a greater voltage in the top secondary coil. Hence, $v_{out} > 0$ for positive displacements. The converse is true for negative displacements. More formally, if the primary coil has resistance R_p and self-inductance L_p, we can write

$$iR_p + L_p \frac{di}{dt} = v_{ex}$$

and the voltages induced in the secondary coils are given by

$$v_1 = M_1 \frac{di}{dt}$$

$$v_2 = M_2 \frac{di}{dt}$$

so that

$$v_{out} = (M_1 - M_2) \frac{di}{dt}$$

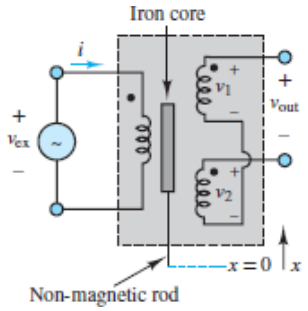


Figure 14.8 Linear variable differential transformer

where M_1 and M_2 are the mutual inductances between the primary and the respective secondary coils. It should be apparent that each of the mutual inductances is dependent on the position of the iron core. For example, with the core at the *null position,* $M_1 = M_2$ and $v_{out} = 0$. The LVDT is typically designed so that $M_1 - M_2$ is linearly related to the displacement of the core x.

Page 812

Because the excitation is by necessity an AC signal (why?), the output voltage is actually given by the difference of two sinusoidal voltages at the same frequency and is therefore itself a sinusoid, whose amplitude and phase depend on the displacement x. Thus, v_{out} is an *amplitude-modulated* (AM) signal, similar to the one discussed in the Focus on Measurements box, "Capacitive Displacement Transducer and Microphone," in Chapter 3. To recover a signal proportional to the actual displacement, it is therefore necessary to use a demodulator circuit, such as the one discussed in the Focus on Measurements box, "Peak Detector Circuit for Capacitive Displacement Transducer," in Chapter 8.

Ampère's Law

As explained in the previous section, Faraday's law is one of two fundamental laws relating electricity to magnetism. The second relationship, which forms a counterpart to Faraday's law, is **Ampère's law**. Qualitatively, Ampère's law states that the magnetic field intensity **H** in the vicinity of a conductor is related to the current carried by the conductor; thus Ampère's law establishes a dual relationship with Faraday's law.

In the previous section, the magnetic field is described by its flux density **B** and flux ϕ. To explain Ampère's law and the behavior of magnetic materials, we define the magnetic field intensity **H** as:

$$\mathbf{B} = \mu\mathbf{H} = \mu_r\mu_0\mathbf{H} \qquad \text{Wb/m}^2 \text{ or T} \tag{14.19}$$

where μ is a scalar constant for a particular physical medium (at least, for the applications we consider here) and is called the **permeability** of the medium. The permeability of a material can be factored as the product of the permeability of free space $\mu_0 = 4\pi \times 10^{-7}$ H/m, and the relative permeability μ_r, which varies greatly according to the medium. For example, for air and for most electrical conductors and insulators, μ_r is equal to 1. For ferromagnetic materials, μ_r can take values ranging from 10^3 to 10^6. The size of μ_r represents a measure of the magnetic properties of the material. A consequence of Ampère's law is that the larger the value of μ, the smaller the current required to produce a large flux density in an electromagnetic structure. Consequently, many electromechanical devices make use of ferromagnetic materials, called iron cores, to enhance their magnetic properties. Table 14.1 gives approximate values of μ_r for some common materials.

Table 14.1 Relative permeabilities for common materials

Material	μ_r
Air	1
Water	1
Copper	1
Nickel	100–600
Ferrite (manganese-zinc)	640
Steel	100
Electrical steel	4,000
Iron	5,000
Permalloy	8,000
Mu-metal	20,000
Nanoperm	80,000
Metglas	1,000,000

The reason for introducing the magnetic field intensity is that it is independent of the properties of the materials employed in the construction of magnetic circuits.

Thus, a given magnetic field intensity **H** will give rise to different flux densities in different materials. It will therefore be useful to define *sources* of magnetic energy in terms of the magnetic field intensity, so that different magnetic structures and materials can then be evaluated or compared for a given source. As stated earlier, both the magnetic flux density and the field intensity are vector quantities; however, for ease of analysis, scalar fields will be chosen by appropriately selecting the orientation of the fields, wherever possible.

Page 813

Ampère's law states that the integral of the vector magnetic field intensity **H** around a closed path is equal to the total current linked by the closed path i:

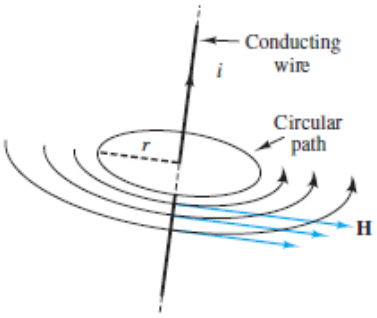
$$\oint \mathbf{H} \cdot d\mathbf{l} = \sum i \tag{14.20}$$

where $d\mathbf{l}$ is an increment in the direction of the closed path. If at every point along the path the magnetic field is parallel to the path, we can use scalar quantities to write:

$$\int H \, dl = \sum i \tag{14.21}$$

Figure 14.9 illustrates the case of a wire carrying a current i and of a circular path of radius r surrounding the wire. In this simple case, you can see that the magnetic field intensity **H** is determined by the familiar right-hand rule. This rule states that if the direction of current i points in the direction of the thumb of one's right hand, the resulting magnetic field encircles the conductor in the direction in which the other four fingers would encircle it. Thus, in the case of Figure 14.9, the closed-path integral becomes equal to $H \cdot 2\pi r$, since the path and the magnetic field are in the same direction, and therefore the magnitude of the magnetic field intensity is given by

$$H = \frac{i}{2\pi r} \tag{14.22}$$



By the right-hand rule, the current i generates a magnetic field intensity **H** in the direction shown.

Figure 14.9 Illustration of Ampère's law

CHECK YOUR UNDERSTANDING

The magnitude of **H** at a radius of 0.5 m from a long linear conductor is 1 A-m^{-1}. Find the current in the wire.

Answer: $I = \pi$ A

Now, the magnetic field intensity **H** is unaffected by the material surrounding the conductor, but the flux density **B** depends on the material properties. The density of flux lines around the conductor would be far greater in the presence of a magnetic material than if the conductor were surrounded by air. The field generated by a Page 814single conducting wire is relatively weak; however, if the wire is a tightly wound coil with many turns, the strength of the magnetic field is increased greatly. For a coil with N turns, one can verify visually that the lines of force associated with the magnetic field link all the turns of the conducting coil, so that we have effectively increased the current linked by the flux lines N-fold. The product $N \cdot i$ is a useful quantity in electromagnetic circuits and is called the **magnetomotive force**,[2] $\mathcal{F}$ (or **mmf**), in analogy with the electromotive force.

$$\mathcal{F} = Ni \quad \text{A-turns} \quad \text{Magnetomotive force} \tag{14.23}$$

Figure 14.10 illustrates the magnetic flux lines in the vicinity of a coil. The magnetic field generated by the coil can be made to generate a much greater flux density if the coil encloses a magnetic material. The most common ferromagnetic materials are steel and iron; in addition to these, many alloys and oxides of iron—as well as nickel—and some artificial ceramic materials called **ferrites** exhibit magnetic properties. In recent years, rare earth magnets have found increasing use, especially in the design of high-performance electric motors. The two most common rare earth

materials are neodymium and samarium (lanthanides), which are used in compounds that include transition metals, such as iron, nickel, and cobalt. Such magnets can produce magnetic fields of strength two to three times greater than ferrites. Winding a coil around a ferromagnetic material accomplishes two useful tasks at once: It forces the magnetic flux to be concentrated within the coil and—if the shape of the magnetic material is appropriate—completely confines the flux within the magnetic material, Page 815thus forcing the closed path for the flux lines to be almost entirely enclosed within the ferromagnetic material. Typical arrangements are the iron-core inductor and the toroid of Figure 14.11. The flux densities for these inductors are given by

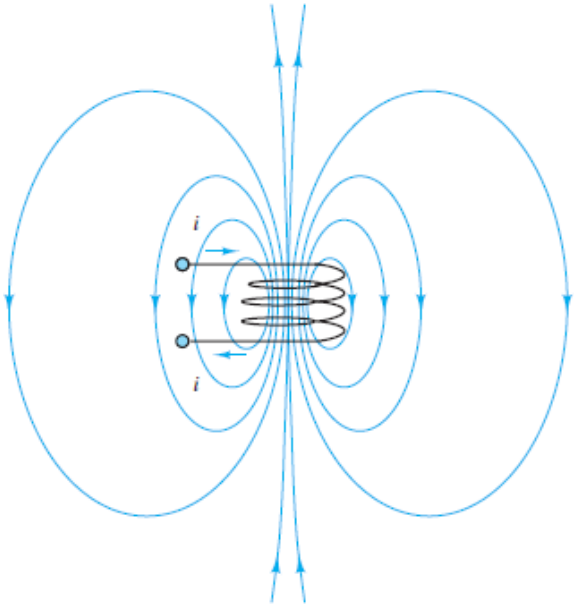


Figure 14.10 Magnetic flux lines in the vicinity of a current-carrying coil

$$B = \frac{\mu N i}{l} \qquad \text{Flux density for tightly wound circular coil} \tag{14.24}$$

$$B = \frac{\mu N i}{2\pi r_2} \qquad \text{Flux density for toroidal coil} \tag{14.25}$$

In equation 14.24, l represents the length of the coil wire; Figure 14.11 defines the parameter r_2 in equation 14.25.

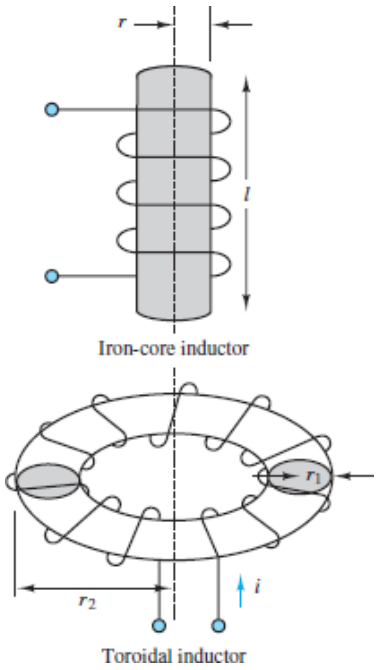


Figure 14.11 Practical inductors

Intuitively, the presence of a high-permeability material near a source of magnetic flux causes the flux to preferentially concentrate in the high-μ material, rather than in air, much as a conducting path concentrates the current produced by an electric field in an electric circuit. Figure 14.12 depicts an example of a simple electromagnetic structure which forms the basis of the practical transformer.

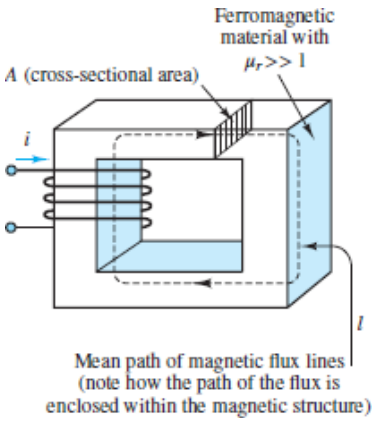


Figure 14.12 A simple electromagnetic structure

Table 14.2 summarizes the variables introduced thus far in the discussion of electricity and magnetism.

Table 14.2 Magnetic variables and units

Variable	Symbol	Units
Current	I	A
Magnetic flux density	B	$Wb/m^2 = T$
Magnetic flux	ϕ	Wb
Magnetic field intensity	H	A/m
Electromotive force	e	V
Magnetomotive force	$\mathcal{F}$	A-turns
Flux linkage	λ	Wb-turns

14.2 MAGNETIC CIRCUITS

It is possible to analyze the operation of electromagnetic devices such as the one depicted in Figure 14.12 by means of magnetic equivalent circuits, similar in many respects to the equivalent electric circuits of earlier chapters. Before we can present this technique, however, we need to make a few simplifying approximations. The first of these approximations assumes that there exists a **mean path** for the magnetic flux and that the corresponding mean flux density is approximately constant over the cross-sectional area of the magnetic structure. Using equation 14.4, we see that a coil wound around a core with cross-sectional area A will have flux density

$$B = \frac{\phi}{A} \tag{14.26}$$

where A is assumed to be perpendicular to the direction of the flux lines. Figure 14.12 illustrates such a mean path and the cross-sectional area A. Knowing the flux density, we obtain the field intensity:

$$H = \frac{B}{\mu} = \frac{\phi}{A\mu} \tag{14.27}$$

But then, knowing the field intensity, we can relate the mmf of the coil $\mathcal{F}$ to the product of the magnetic field intensity H and the length of the magnetic (mean) Page 816path l; we can use equations 14.24 and 14.19 to derive

$$\mathcal{F} = N \cdot i = H \cdot l \tag{14.28}$$

In summary, the mmf is equal to the magnetic flux times the length of the magnetic path, divided by the permeability of the material times the cross-sectional area:

$$\mathcal{F} = \phi \frac{l}{\mu A} \tag{14.29}$$

A review of this formula reveals that the magnetomotive force $\mathcal{F}$ may be viewed as being analogous to the voltage source in a series electric circuit, and that the flux ϕ is then equivalent to the electric current in a series circuit and the term $l/\mu A$ to the "magnetic resistance" of one leg of the magnetic circuit. You will note that the term $l/\mu A$ is very similar to the term describing the resistance of a cylindrical conductor of length l and cross-sectional area A, where the permeability μ is analogous to the conductivity σ. The term $l/\mu A$ occurs frequently enough to be assigned the name of **reluctance** and the symbol $\mathcal{R}$. It is also important to recognize the *relationship between the reluctance of a magnetic structure and its inductance.* This can be derived easily starting from equation 14.14:

$$L = \frac{\lambda}{i} = \frac{N\phi}{i} = \frac{N}{i}\frac{Ni}{\mathcal{R}} = \frac{N^2}{\mathcal{R}} \quad \text{H} \tag{14.30}$$

In summary, when an N-turn coil carrying a current i is wound around a magnetic core such as the one indicated in Figure 14.12, the mmf $\mathcal{F}$ generated by the coil produces a flux ϕ that is *mostly* concentrated within the core and is assumed to be uniform across the cross section. Within this simplified picture, then, the analysis of a magnetic circuit is analogous to that of resistive electric circuits. This analogy is illustrated in Table 14.3 and in the examples in this section.

Table 14.3 Analogy between electric and magnetic circuits

Electrical quantity	Magnetic quantity
Electrical field intensity E, V/m	Magnetic field intensity H, A-turns/m
Voltage v, V	Magnetomotive force $\mathcal{F}$, A-turns
Current i, A	Magnetic flux ϕ, Wb
Current density J, A/m^2	Magnetic flux density B, Wb/m^2
Resistance R, Ω	Reluctance $\mathcal{R} = l/\mu A$, A-turns/Wb
Conductivity σ, 1/Ω-m	Permeability μ, Wb/A-m

The usefulness of the magnetic circuit analogy can be emphasized by analyzing a magnetic core similar to that of Figure 14.12, but with a slightly modified geometry. Figure 14.13 depicts the magnetic structure and its equivalent-circuit analogy. In the

figure, we see that the mmf $\mathcal{F} = Ni$ excites the magnetic circuit, which is composed of four legs: two of mean path length l_1 and cross-sectional area $A_1 = d_1 w$, and the other two of mean length l_2 and cross-sectional area $A_2 = d_2 w$. Thus, the reluctance Page 817encountered by the flux in its path around the magnetic core is given by the quantity $\mathcal{R}_{\text{series}}$, with

$$\mathcal{R}_{\text{series}} = 2\mathcal{R}_1 + 2\mathcal{R}_2$$

and

$$\mathcal{R}_1 = \frac{l_1}{\mu A_1} \qquad \mathcal{R}_2 = \frac{l_2}{\mu A_2}$$

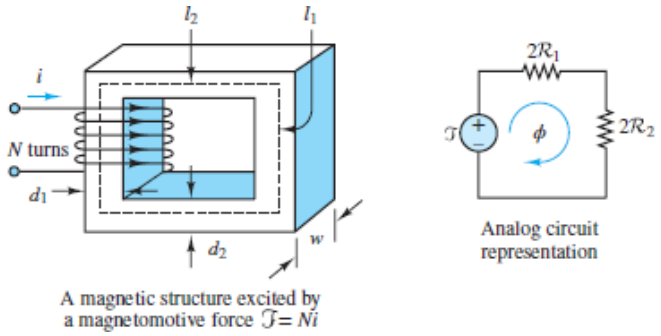


Figure 14.13 Analogy between magnetic and electric circuits

It is important at this stage to review the assumptions and simplifications made in analyzing the magnetic structure of Figure 14.13:

1. All the magnetic flux is linked by all the turns of the coil.
2. The flux is confined exclusively within the magnetic core.
3. The density of the flux is uniform across the cross-sectional area of the core.

You can probably see intuitively that the first of these assumptions might not hold true near the ends of the coil, but that it is more reasonable if the coil is tightly wound. The second assumption is equivalent to stating that the relative permeability of the core is infinitely higher than that of air (presuming that this is the medium surrounding the core); if this were the case, the flux would indeed be confined within the core. It is worthwhile to note that we make a similar assumption when we treat

wires in electric circuits as perfect conductors: The conductivity of copper is substantially greater than that of free space, by a factor of approximately 10^{15}. In the case of magnetic materials, however, even for the best alloys, we have a relative permeability only on the order of 10^3 to 10^5. Thus, an approximation that is quite appropriate for electric circuits is not nearly as good in the case of magnetic circuits. The flux in a structure, such as those of Figures 14.12 and 14.13, not confined within the core is usually referred to as **leakage flux**. Finally, the assumption that the flux is uniform across the core cannot hold for a finite-permeability medium, but it is very helpful in giving an approximate *mean* behavior of the magnetic circuit.

Page 818

The magnetic circuit analogy is therefore far from exact. However, short of employing the tools of electromagnetic field theory and of vector calculus, or advanced numerical simulation software, it is the most convenient tool at the engineer's disposal for the analysis of magnetic structures. In the remainder of this chapter, the approximate analysis based on the electric circuit analogy is used to obtain approximate solutions to problems involving a variety of useful magnetic circuits. Among these are the loudspeaker, solenoids, automotive fuel injectors, and sensors for the measurement of linear and angular velocity and position.

EXAMPLE 14.2 Analysis of Magnetic Structure and Equivalent Magnetic Circuit

Problem

Calculate the flux, flux density, and field intensity on the magnetic structure of Figure 14.14.

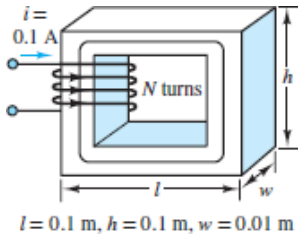


Figure 14.14 Figure for Example 14.2.

Solution

Known Quantities: Relative permeability; number of coil turns; coil current; structure geometry.

Find: ϕ; B; H.

Schematics, Diagrams, Circuits, and Given Data: $\mu_r = 1{,}000$; $N = 500$ turns; $i = 0.1$ A. The cross-sectional area is $A = w^2 = (0.01)^2 = 0.0001$ m^2. The magnetic circuit geometry is defined in Figures 14.14 and 14.15.

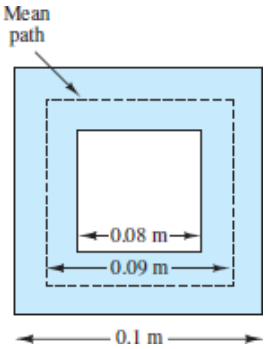


Figure 14.15 Cross section of magnetic structure for Example 14.2.

Assumptions: All magnetic flux is linked by the coil; the flux is confined to the magnetic core; the flux density is uniform.

Analysis:

1. *Calculation of magnetomotive force.* From equation 14.28, we calculate the magnetomotive force:

$$\mathcal{F} = \text{mmf} = Ni = (500 \text{ turns})(0.1 \text{ A}) = 50 \text{ A-turns}$$

2. *Calculation of mean path.* Next, we estimate the mean path of the magnetic flux. On the basis of the assumptions, we can calculate a mean path that runs through the geometric center of the magnetic structure, as shown in Figure 14.15. The path length is

$$l_c = 4 \times 0.09 \text{ m} = 0.36 \text{ m}$$

3. *Calculation of reluctance.* Knowing the magnetic path length and cross-sectional area, we can calculate the reluctance of the circuit:

$$\mathcal{R} = \frac{l_c}{\mu A} = \frac{l_c}{\mu_r \mu_0 A} = \frac{0.36}{1{,}000 \times 4\pi \times 10^{-7} \times 0.0001}$$
$$= 2.865 \times 10^6 \text{ A-turns/Wb}$$

The corresponding equivalent magnetic circuit is shown in Figure 14.16.

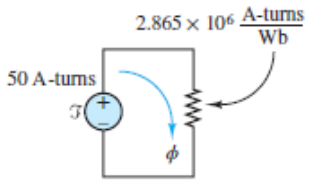


Figure 14.16 Equivalent magnetic circuit for Example 14.2.

Page 819

4. *Calculation of magnetic flux, flux density, and field intensity.* On the basis of the assumptions, we can now calculate the magnetic flux

$$\phi = \frac{\mathcal{F}}{\mathcal{R}} = \frac{50 \text{ A-turns}}{2.865 \times 10^6 \text{ A-turns/Wb}} = 1.75 \times 10^{-5} \text{ Wb}$$

the flux density

$$B = \frac{\phi}{A} = \frac{\phi}{w^2} = \frac{1.75 \times 10^{-5} \text{ Wb}}{0.0001 \text{ m}^2} = 0.175 \text{ Wb/m}^2$$

and the magnetic field intensity

$$H = \frac{B}{\mu} = \frac{B}{\mu_r \mu_0} = \frac{0.175 \text{ Wb/m}^2}{1{,}000 \times 4\pi \times 10^{-7} \text{ H/m}} = 139 \text{ A-turns/m}$$

Comments: This example illustrates all the basic calculations that pertain to magnetic structures. Remember that the assumptions stated in this example (and earlier in the chapter) simplify the problem and make its approximate numerical solution possible in a few simple steps. In reality, flux leakage, fringing, and uneven distribution of flux across the structure would require the solution of three-dimensional equations using finite-element methods. These methods are not discussed in this book, but are necessary for practical engineering designs.

The usefulness of these approximate methods is that you can, for example, quickly calculate the approximate magnitude of the current required to generate a given magnetic flux or flux density. These calculations can be used to determine electromagnetic energy and magnetic forces in practical structures.

The methodology described in this example is summarized in the following Focus on Problem Solving box.

CHECK YOUR UNDERSTANDING

Determine the equivalent reluctance of the structure of Figure 14.17 as seen by the "source" if μ_r for the structure is 1,000, l = 5 cm, and all the legs are 1 cm on a side.

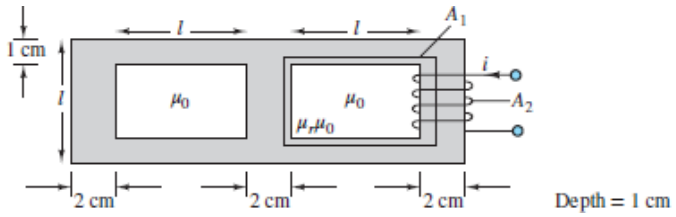


Figure 14.17 Magnetic structure with two loops.

Answer: Assuming a mean path 1 cm from the edges of the structure,
$\mathcal{R}_{eq} = 1.41 \times 10^6$ A-turns/Wb

Page 820

FOCUS ON PROBLEM SOLVING

MAGNETIC STRUCTURES AND EQUIVALENT MAGNETIC CIRCUITS

Direct Problem

Given—The structure geometry and the coil parameters.

Calculate—The magnetic flux in the structure.

1. Compute the mmf.
2. Determine the length and cross section of the magnetic path for each continu *leg* or section of the path.
3. Calculate the equivalent reluctance of the *leg*.
4. Generate the equivalent magnetic circuit diagram, and calculate the equivalent reluctance.
5. Calculate the flux, flux density, and magnetic field intensity, as needed.

Inverse Problem

Given—The desired flux or flux density and structure geometry.

Calculate—The necessary coil current and number of turns.

1. Calculate the total equivalent reluctance of the structure from the desired flux
2. Generate the equivalent magnetic circuit diagram.
3. Determine the mmf required to establish the required flux.
4. Choose the coil current and number of turns required to establish the des mmf.

Consider the analysis of the same simple magnetic structure when an **air gap** is present. Air gaps are very common in magnetic structures; in rotating machines, for example, air gaps are necessary to allow for free rotation of the inner core of the machine. The magnetic circuit of Figure 14.18(a) differs from the circuit analyzed in Example 14.2 simply because of the presence of an air gap; the effect of the gap is to break the continuity of the high-permeability path for the flux, adding a high-reluctance component to the equivalent circuit. The situation is analogous to adding a very large series resistance to a series electric circuit. It should be evident from Figure 14.18(a) that the basic concept of reluctance still applies, although now two different permeabilities must be taken into account.

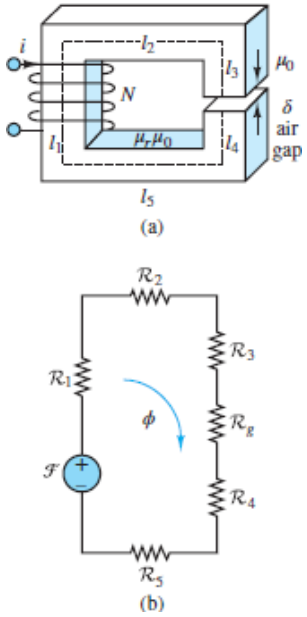


Figure 14.18 (a) Magnetic circuit with air gap; (b) its equivalent magnetic circuit

The equivalent circuit for the structure of Figure 14.18(a) may be drawn as shown in Figure 14.18(b), where $\mathcal{R}_n$ is the reluctance of path l_n , for $n = 1, 2, \ldots, 5$, Page 821and $\mathcal{R}_g$ is the reluctance of the air gap. The reluctances can be expressed as follows, if we assume that the magnetic structure has a uniform cross-sectional area A:

$$\mathcal{R}_1 = \frac{l_1}{\mu_r \mu_0 A} \qquad \mathcal{R}_2 = \frac{l_2}{\mu_r \mu_0 A} \qquad \mathcal{R}_3 = \frac{l_3}{\mu_r \mu_0 A}$$
$$\mathcal{R}_4 = \frac{l_4}{\mu_r \mu_0 A} \qquad \mathcal{R}_5 = \frac{l_5}{\mu_r \mu_0 A} \qquad \mathcal{R}_g = \frac{\delta}{\mu_0 A_g} \tag{14.31}$$

Note that in computing $\mathcal{R}_g$, the length of the gap is given by δ and the permeability is given by μ_0, as expected, but A_g is different from the cross-sectional area A of the structure. This is so because the flux lines exhibit a phenomenon known as **fringing** as they cross an air gap. The flux lines actually *bow out* of the gap defined by the cross section A, not being contained by the high-permeability material any longer. Thus, it is customary to define an area A_g that is greater than A, to account for this phenomenon. Example 14.3 describes in greater detail the procedure for finding A_g and also discusses the phenomenon of fringing.

EXAMPLE 14.3 Magnetic Structure With Air Gaps

Problem

Compute the equivalent reluctance of the magnetic circuit of Figure 14.19 and the flux density established in the bottom bar of the structure.

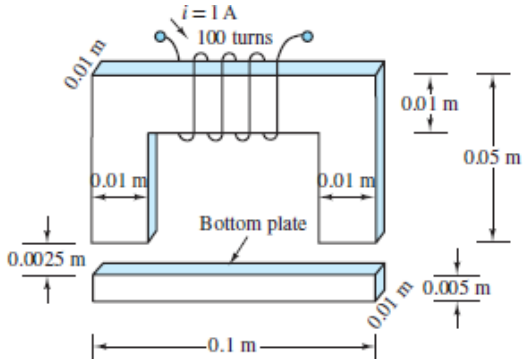


Figure 14.19 Electromagnetic structure with air gaps

Solution

Known Quantities: Relative permeability; number of coil turns; coil current; structure geometry.

Find: $\mathcal{R}_{eq}$; B_{bar}.

Schematics, Diagrams, Circuits, and Given Data: $\mu_r = 10{,}000$; $N = 100$ turns; $i = 1$ A.

Assumptions: All magnetic flux is linked by the coil; the flux is confined to the magnetic core; the flux density is uniform.

Page 822

Analysis:

1. *Calculation of magnetomotive force.* From equation 14.28, we calculate the magnetomotive force:

$$\mathcal{F} = \text{mmf} = Ni = (100 \text{ turns})(1 \text{ A}) = 100 \text{ A-turns}$$

2. *Calculation of mean path.* Figure 14.20 depicts the geometry. The path length is

$$l_c = l_1 + l_2 + l_3 + l_4 + l_5 + l_6 + l_g + l_g$$

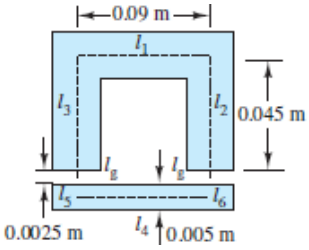


Figure 14.20 Geometry of magnetic structure with air gap

However, the path must be broken into three legs: the upside-down U-shaped element, the air gaps, and the bar. We cannot treat these three parts as one because the relative permeability of the magnetic material is very different from that of the air gap. Thus, we define the following three paths, neglecting the very small (half bar thickness) lengths l_5 and l_6:

$$l_U = l_1 + l_2 + l_3 \qquad l_{bar} = l_4 + l_5 + l_6 \approx l_4 \qquad l_{gap} = l_g + l_g$$

where

$$l_U = 0.18 \text{ m} \qquad l_{bar} = 0.09 \text{ m} \qquad l_{gap} = 0.005 \text{ m}$$

Next, we compute the cross-sectional area. For the magnetic structure, we calculate the U-shaped element cross section to be $A_U = w^2 = (0.01)^2 = 0.0001$ m^2 and the cross section of the bar to be $A_{bar} = (0.01 \times 0.005) = 0.0005$ m^2. For the air gap, we will make an empirical adjustment to account for the phenomenon of *fringing*, that is, to account for the tendency of the magnetic flux lines to bow out of the magnetic path, as illustrated in Figure 14.21. A rule of thumb used to account for fringing is to add the length of the gap to each dimension of the actual cross-sectional area. Thus

$$A_{gap} = (0.01 \text{ m} + l_g)^2 = (0.0125)^2 = 0.15625 \times 10^{-3} \text{m}^2$$

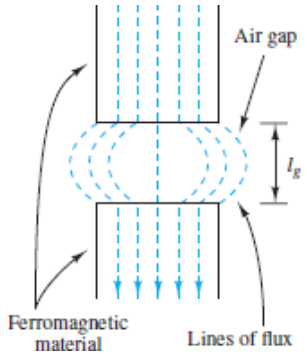


Figure 14.21 Fringing effects in air gap

3. *Calculation of reluctance.* Knowing the magnetic path length and cross-sectional area, we can calculate the reluctance of each leg of the circuit:

$$\mathcal{R}_U = \frac{l_U}{\mu_U A_U} = \frac{l_U}{\mu_r \mu_0 A_U} = \frac{0.18}{10{,}000 \times 4\pi \times 10^{-7} \times 0.0001}$$
$$= 1.43 \times 10^5 \text{ A-turns/Wb}$$

$$\mathcal{R}_{bar} = \frac{l_{bar}}{\mu_{bar} A_{bar}} = \frac{l_{bar}}{\mu_r \mu_0 A_{bar}} = \frac{0.09}{10{,}000 \times 4\pi \times 10^{-7} \times 0.0005}$$
$$= 143.2 \times 10^3 \text{ A-turns/Wb}$$

$$\mathcal{R}_{gap} = \frac{l_{gap}}{\mu_{gap} A_{gap}} = \frac{l_{gap}}{\mu_0 A_{gap}} = \frac{0.005}{4\pi \times 10^{-7} \times 0.156 \times 10^{-3}} = 25.5 \times 10^6 \text{ A-turns/Wb}$$

Note that the reluctance of the air gap is dominant with respect to that of the magnetic structure, in spite of the small dimension of the gap. This is so because the relative permeability of the air gap is much smaller than that of the magnetic material.

The equivalent reluctance of the structure is

$$\mathcal{R}_{eq} = \mathcal{R}_U + \mathcal{R}_{bar} + \mathcal{R}_{gap} = 1.43 \times 10^5 + 143.2 \times 10^3 + 2.55 \times 10^7$$
$$= 25.8 \times 10^6 \text{ A-turns/Wb}$$

Thus,

$$\mathcal{R}_{eq} \approx \mathcal{R}_{gap}$$

Since the gap reluctance is two orders of magnitude greater than the reluctance of the magnetic structure, it is reasonable to neglect the magnetic structure reluctance and work only with the gap reluctance in calculating the magnetic flux.

Page 823

4. *Calculation of magnetic flux and flux density in the bar.* From the result of the preceding subsection, we calculate the flux

$$\phi = \frac{\mathcal{F}}{\mathcal{R}_{eq}} \approx \frac{\mathcal{F}}{\mathcal{R}_{gap}} = \frac{100 \text{ A-turns}}{2.55 \times 10^7 \text{ A-turns/Wb}} = 3.9 \times 10^{-6} \text{ Wb}$$

and the flux density in the bar

$$B_{bar} = \frac{\phi}{A} = \frac{3.92 \times 10^{-6} \text{ Wb}}{0.00005 \text{ m}^2} = 78.5 \times 10^{-3} \text{ Wb/m}^2$$

Comments: It is very common to neglect the reluctance of the magnetic material sections in these approximate calculations. We shall make this assumption very frequently in the remainder of the chapter.

CHECK YOUR UNDERSTANDING

Find the equivalent reluctance of the magnetic circuit shown in Figure 14.22 if μ_r of the structure is infinite, $\delta = 2$ mm, and the physical cross section of the core is 1 cm^2. Do not neglect fringing.

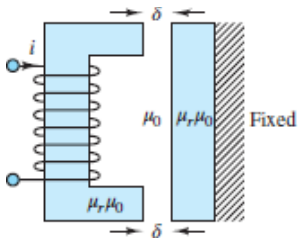


Figure 14.22 Magnetic circuit

Answer: $\mathcal{R}_{eq} = 22 \times 10^6$ A-turns/Wb

EXAMPLE 14.4 Magnetic Structure of Electric Motor

Problem

Figure 14.23 depicts the configuration of an electric motor. The electric motor consists of a *stator* and a *rotor*. Compute the air gap flux and flux density. Neglect fringing.

Solution

Known Quantities: Relative permeability; number of coil turns; coil current; structure geometry.

Find: ϕ_{gap}; B_{gap}.

Schematics, Diagrams, Circuits, and Given Data: $\mu_r \rightarrow \infty$; N = 1,000 turns; i = 10 A; l_{gap} = 0.01 m; A_{gap} = 0.1 m^2. The magnetic circuit geometry is defined in Figure 14.23.

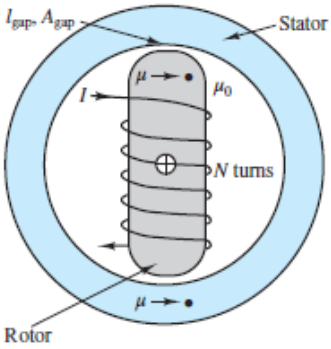


Figure 14.23 Cross-sectional view of synchronous motor

Assumptions: All magnetic flux is linked by the coil; the flux is confined to the magnetic core; the flux density is uniform. The reluctance of the magnetic structure is negligible.

Analysis:

1. *Calculation of magnetomotive force.* From equation 14.28, we calculate the magnetomotive force:

$$\mathcal{F} = \text{mmf} = Ni = (1{,}000 \text{ turns})(10 \text{ A}) = 10{,}000 \text{ A-turns}$$

2. *Calculation of reluctance.* Knowing the magnetic path length and cross-sectional area, we can calculate the equivalent reluctance of the two gaps:

$$\mathcal{R}_{gap} = \frac{l_{gap}}{\mu_{gap} A_{gap}} = \frac{l_{gap}}{\mu_0 A_{gap}} = \frac{0.01}{4\pi \times 10^{-7} \times 0.1} = 7.96 \times 10^4 \text{ A-turns/Wb}$$

$$\mathcal{R}_{eq} = 2\mathcal{R}_{gap} = 1.59 \times 10^5 \text{ A-turns/Wb}$$

3. *Calculation of magnetic flux and flux density.* From the results of steps 1 and 2, we calculate the flux

$$\phi = \frac{\mathcal{F}}{\mathcal{R}_{eq}} = \frac{10{,}000 \text{ A-turns}}{1.59 \times 10^5 \text{ A-turns/Wb}} = 0.0628 \text{ Wb}$$

and the flux density

$$B_{bar} = \frac{\phi}{A} = \frac{0.0628 \text{ Wb}}{0.1 \text{ m}^2} = 0.628 \text{ Wb/m}^2$$

Comments: Note that the flux and flux density in this structure are significantly larger than those in Example 14.3 because of the larger mmf and larger gap area of this magnetic structure.

The subject of electric motors is formally approached in Chapter 15.

EXAMPLE 14.5 Equivalent Circuit of Magnetic Structure With Multiple Air Gaps

Problem

Figure 14.24 depicts the configuration of a magnetic structure with two air gaps. Determine the equivalent circuit of the structure.

Page 825

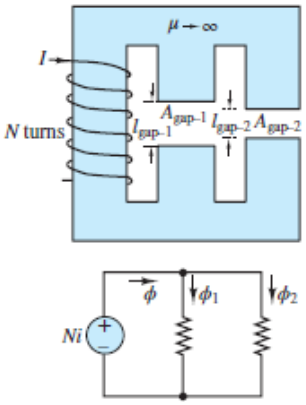


Figure 14.24 Magnetic structure with two air gaps

Solution

Known Quantities: Structure geometry.

Find: Equivalent-circuit diagram.

Assumptions: All magnetic flux is linked by the coil; the flux is confined to the magnetic core; the flux density is uniform. The reluctance of the magnetic structure is negligible.

Analysis:

1. *Calculation of magnetomotive force.*

$$\mathcal{F} = \text{mmf} = Ni$$

2. *Calculation of reluctance.* Knowing the magnetic path length and cross-sectional area, we can calculate the equivalent reluctance of the two gaps:

$$\mathcal{R}_{\text{gap}-1} = \frac{l_{\text{gap}-1}}{\mu_{\text{gap}-1} A_{\text{gap}-1}} = \frac{l_{\text{gap}-1}}{\mu_0 A_{\text{gap}-1}}$$

$$\mathcal{R}_{\text{gap}-2} = \frac{l_{\text{gap}-2}}{\mu_{\text{gap}-2} A_{\text{gap}-2}} = \frac{l_{\text{gap}-2}}{\mu_0 A_{\text{gap}-2}}$$

3. *Calculation of magnetic flux and flux density.* Note that the flux must now divide between the two legs, and that a different air-gap flux will exist in each leg. Thus

$$\phi_1 = \frac{Ni}{\mathcal{R}_{\text{gap}-1}} = \frac{Ni\mu_0 A_{\text{gap}-1}}{l_{\text{gap}-1}}$$

$$\phi_2 = \frac{Ni}{\mathcal{R}_{\text{gap}-2}} = \frac{Ni\mu_0 A_{\text{gap}-2}}{l_{\text{gap}-2}}$$

and the total flux generated by the coil is $\phi = \phi_1 + \phi_2$.

The equivalent circuit is shown in the bottom half of Figure 14.24.

Comments: Note that the two legs of the structure act as resistors in a parallel circuit.

CHECK YOUR UNDERSTANDING

Find the equivalent magnetic circuit of the structure of Figure 14.25 if μ_r is infinite. Give expressions for each of the circuit values if the physical cross-sectional area of each of the legs is given by

$$A = l \times w$$

Do not neglect fringing.

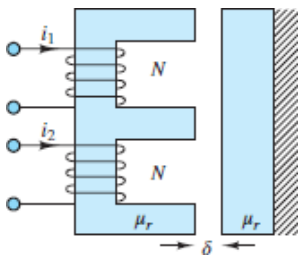


Figure 14.25 Magnetic circuit

Answer: $\mathcal{R}_g = \mathcal{R}_1 = \mathcal{R}_2 = \mathcal{R}_3 = \delta/\mu_0(l + \delta)(w + \delta)$; $\mathcal{F}_1 = Ni_1$; $\mathcal{F}_2 = Ni_2$

Page 826

EXAMPLE 14.6 Inductance, Stored Energy, and Induced Voltage

Problem

1. Find the inductance and the magnetic energy stored in the structure of Figure 14.18(a). The structure is identical to that of Example 14.2 except for the air gap. Ignore fringing.
2. Assume that the flux density in the air gap varies sinusoidally as $B(t) = B_0 sin(\omega t)$. Determine the induced voltage across the coil e.

Solution

Known Quantities: Relative permeability; number of coil turns; coil current; structure geometry; flux density in air gap.

Find: L; W_m; e.

Schematics, Diagrams, Circuits, and Given Data: $\mu_r \rightarrow \infty$; $N = 500$ turns; $i = 0.1$ A. The magnetic circuit geometry is defined in Figures 14.14 and 14.15. The air gap has $l_g = 0.002$ m. $B_0 = 0.6$ Wb/m^2.

Assumptions: All magnetic flux is linked by the coil; the flux is confined to the magnetic core; the flux density is uniform. The reluctance of the magnetic structure is negligible.

Analysis:

1. Use equation 14.30 to calculate the inductance of the magnetic structure.

$$L = \frac{N^2}{\mathcal{R}}$$

   To calculate the reluctance, assume that the reluctance of the structure is negligible.

$$\mathcal{R}_{gap} = \frac{l_{gap}}{\mu_{gap} A_{gap}} = \frac{l_{gap}}{\mu_0 A_{gap}} = \frac{0.002}{4\pi \times 10^{-7} \times 0.0001} = 1.59 \times 10^7 \text{ A-turns/Wb}$$

and

$$L = \frac{N^2}{\mathcal{R}} = \frac{500^2}{1.59 \times 10^7} = 0.157 \text{ H}$$

Finally, calculate the stored magnetic energy as follows:

$$W_m = \frac{1}{2} L i^2 = \frac{1}{2} \times (0.157 \text{ H}) \times (0.1 \text{ A})^2 = 0.785 \times 10^{-3} \text{ J}$$

2. To calculate the induced voltage due to a time-varying magnetic flux at the frequency of 60 Hz(377 rad/s), we use equation 14.16:

$$e = \frac{d\lambda}{dt} = N\frac{d\phi}{dt} = NA\frac{dB}{dt} = NAB_0 \omega \cos(\omega t)$$
$$= 500 \times 0.0001 \times 0.6 \times 377 \cos(377t) = 11.31 \cos(377t) \quad \text{V}$$

Comments: The voltage induced across a coil in an electromagnetic transducer is a very important quantity called the *back electromotive force*, or back emf.

Page 827

FOCUS ON MEASUREMENTS

Magnetic Reluctance Position Sensor

A simple magnetic structure, very similar to those examined in the previous examples, finds very common application in the variable-reluctance position sensor, which, in turn, finds widespread application in a variety of configurations for the measurement of linear and angular velocity. Figure 14.26 depicts one particular configuration that is used in many applications. In this structure, a permanent magnet with a coil of wire wound around it forms the sensor; a steel disk (typically connected to a rotating shaft) has a number of tabs that pass between the pole pieces of the sensor. The area of the tab is assumed equal to the area of the cross section of the pole pieces and is equal to a^2. The reason for the name *variable-reluctance sensor* is that the reluctance of the magnetic structure is variable, depending on whether a ferromagnetic tab lies between the pole pieces of the magnet.

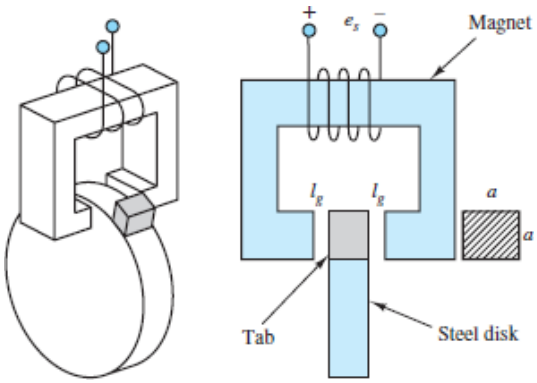


Figure 14.26 Variable-reluctance position sensor

The principle of operation of the sensor is that an electromotive force e_S is induced across the coil by the change in magnetic flux caused by the passage of the tab between the pole pieces when the disk is in motion. As the tab enters the volume between the pole pieces, the flux will increase, because of the lower reluctance of the configuration, until it reaches a maximum when the tab is centered between the poles of the magnet. Figure 14.27 depicts the approximate shape of the resulting voltage, which, according to Faraday's law, is given by

$$e_S = -\frac{d\phi}{dt}$$

The rate of change of flux is dictated by the geometry of the tab and of the pole pieces and by the speed of rotation of the disk. It is important to note that, since the flux is changing only if the disk is rotating, this sensor cannot detect the static position of the disk.

One common application of this concept is in the measurement of the speed of rotation of rotating machines, including electric motors and internal combustion engines. In these applications, use is made of a *60-tooth wheel*, which permits the conversion of the speed rotation directly to units of revolutions per minute. The output of a variable-reluctance position sensor magnetically coupled to a rotating disk equipped Page 828with 60 tabs (teeth) is processed through a comparator or Schmitt trigger circuit (see Chapter 7). The voltage waveform generated by the sensor is nearly sinusoidal when the teeth are closely spaced, and it is characterized by one sinusoidal cycle for each tooth on the disk. If a negative zero-crossing detector (see Chapter 7) is employed, the trigger circuit will generate a pulse corresponding to the passage of each tooth, as shown in Figure 14.28. If the time between any two pulses is measured by means of a high-frequency clock, the speed of the engine can be directly determined in units of revolutions per minute (r/min) by means of a digital counter (see Chapter 12).

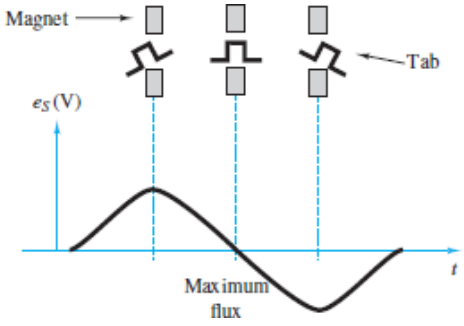


Figure 14.27 Variable-reluctance position sensor waveform

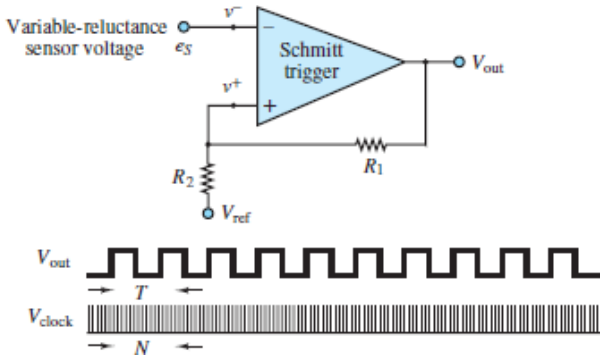


Figure 14.28 Signal processing for a 60-tooth wheel rpm sensor

FOCUS ON MEASUREMENTS

Voltage Calculation in Magnetic Reluctance Position Sensor

Problem:

This example illustrates the calculation of the voltage induced in a magnetic reluctance sensor by a rotating toothed wheel. In particular, we will find an approximate expression for the reluctance and the induced voltage for the position sensor shown in Figure 14.29, Page 829and we will show that the induced voltage is speed dependent. It will be assumed that the reluctance of the core and fringing at the air gaps are both negligible.

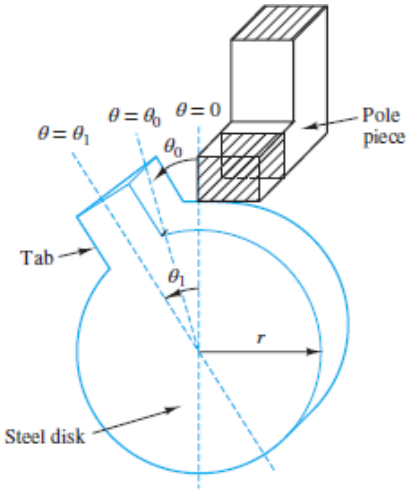


Figure 14.29 Reluctance sensor for measurement of angular position

Solution:

From the geometry shown in the preceding Focus on Measurements box, the equivalent reluctance of the magnetic structure is twice that of one gap, since the permeability of the tab and the magnetic structure are assumed infinite (i.e., they have negligible reluctance). When the tab and the poles are aligned, the angle θ is zero, as shown in Figure 14.29, and the area of the air gap is maximum. For angles greater than $2\theta_0$, the magnetic length of the air gaps is so large that the magnetic field may reasonably be taken as zero.

To model the reluctance of the gaps, we assume the following simplified expression, where the area of overlap of the tab with the magnetic poles is assumed proportional to the angular displacement:

$$\mathcal{R} = \frac{2l_g}{\mu_0 A} = \frac{2l_g}{\mu_0 ar(\theta_1 - \theta)} \qquad \text{for } 0 < \theta < \theta_1$$

Naturally, this is an approximation; however, the approximation captures the essential idea of this transducer, namely, that the reluctance will decrease with increasing overlap area until it reaches a minimum, and then the reluctance will increase as the overlap area decreases. For $\theta = \theta_1$, that is, with the tab outside the magnetic pole pieces, we have $\mathcal{R}_{max} \to \infty$. For $\theta = 0$, that is, with the tab perfectly aligned with the pole pieces, we have $\mathcal{R}_{min} = 2l_g/\mu_0 ar\theta_1$. The flux ϕ may therefore be computed as follows:

$$\phi = \frac{Ni}{\mathcal{R}} = \frac{Ni\mu_0 ar(\theta_1 - \theta)}{2l_g}$$

Page 830

The induced voltage e_S is found by

$$e_S = -\frac{d\phi}{dt} = -\frac{d\phi}{d\theta}\frac{d\theta}{dt} = -\frac{Ni\mu_0 ar}{2l_g}\omega$$

where $\omega = d\theta/dt$ is the rotational speed of the steel disk. It should be evident that the induced voltage is speed dependent. For a = 1 cm, r = 10 cm, l_g = 0.1 cm, N = 1,000 turns, i = 10 mA, $\theta_1 = 6° \approx 0.1$ rad, and ω = 400 rad/s (approximately 3,800 r/min), we have

$$\mathcal{R}_{\max} = \frac{2 \times 0.1 \times 10^{-2}}{4\pi \times 10^{-7} \times 1 \times 10^{-2} \times 10 \times 10^{-2} \times 0.1}$$
$$= 1.59 \times 10^{7} \text{ A-turns/Wb}$$

$$e_{S\,\text{peak}} = \frac{1{,}000 \times 10 \times 10^{-3} \times 4\pi \times 10^{-7} \times 1 \times 10^{-2} \times 10^{-1}}{2 \times 0.1 \times 10^{-2}} \times 400$$
$$= 2.5 \text{ mV}$$

That is, the peak amplitude of e_S will be 2.5 mV.

14.3 MAGNETIC MATERIALS AND *B-H* CURVES

In the analysis of magnetic circuits presented in the previous sections, the relative permeability μ_r was treated as a constant. In fact, the relationship between the magnetic flux density **B** and the associated field intensity **H** is:

$$\mathbf{B} = \mu\mathbf{H} \tag{14.32}$$

and is characterized by the fact that the relative permeability of magnetic materials is not a constant but is a function of the magnetic field intensity. In effect, all magnetic materials exhibit a phenomenon called **saturation**, whereby the flux density increases in proportion to the field intensity until it cannot do so any longer. Figure 14.30 illustrates the general behavior of all magnetic materials. You will note that since the *B-H* curve shown in the figure is nonlinear, the value of μ (which is the slope of the curve) depends on the intensity of the magnetic field.

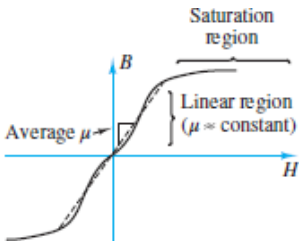


Figure 14.30 Permeability and magnetic saturation effects

To understand the reasons for the saturation of a magnetic material, we need to briefly review the mechanism of magnetization. The basic idea behind magnetic materials is that the spin of electrons constitutes motion of charge, and therefore leads to magnetic effects, as explained in the introductory section of this chapter. In

most materials, the electron spins cancel out, on the whole, and no net effect remains. In ferromagnetic materials, on the other hand, atoms can align so that the electron spins cause a net magnetic effect. In such materials, there exist small regions with strong magnetic properties, called **magnetic domains**, the effects of Page 831which are neutralized in unmagnetized material by other, similar regions that are oriented differently, in a random pattern. When the material is magnetized, the magnetic domains tend to align with one another, to a degree that is determined by the intensity of the applied magnetic field.

In effect, a large number of miniature magnets within the material are aligned (*polarized*) by the applied magnetic field. As the field increases, more and more domains become aligned. When all the domains have become aligned, any further increase in magnetic field intensity does not yield an increase in flux density beyond the increase that would be caused in a nonmagnetic material. Thus, the relative permeability μ_r approaches 1 in the saturation region. It should be apparent that an exact value of μ_r cannot be determined; the value of μ_r used in the earlier examples is to be interpreted as an average permeability, for intermediate values of flux density. For example, commercial magnetic steels saturate at flux densities of a few teslas.

There are two more features that cause magnetic materials to further deviate from the ideal model of the linear *B-H* relationship: **eddy currents** and **hysteresis**. The first phenomenon consists of currents that are caused by any time-varying flux in the core material. As you know, a time-varying flux will induce a voltage, and therefore a current. When this happens inside the magnetic core, the induced voltage will cause *eddy* currents (the terminology should be self-explanatory) in the core, which depend on the resistivity of the core. Figure 14.31 illustrates the phenomenon of eddy currents. The effect of these currents is to dissipate energy in the form of heat. Eddy currents are reduced by selecting high-resistivity core materials, or by *laminating* the core, introducing tiny, discontinuous air gaps between core layers (see Figure 14.31). Lamination of the core reduces eddy currents greatly without affecting the magnetic properties of the core.

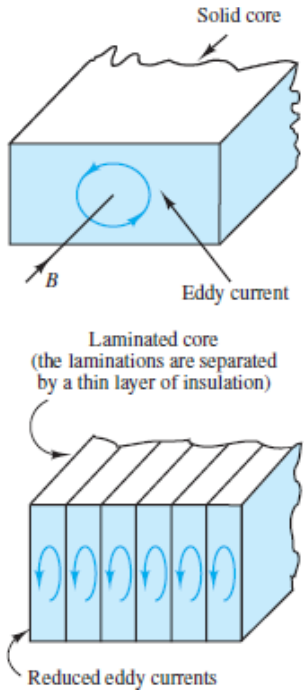


Figure 14.31 Eddy currents in magnetic structures

Hysteresis is another loss mechanism in magnetic materials; it displays a rather complex behavior, related to the magnetization properties of a material. The curve of Figure 14.32 reveals that the *B*-*H* curve for a magnetic material during magnetization (as *H* is increased) is displaced with respect to the curve that is measured when the material is demagnetized. To understand the hysteresis process, consider a core that has been energized for some time, with a field intensity of H_1 A-turns/m. As the current required to sustain the mmf corresponding to H_1 is decreased, we follow the hysteresis curve from the point α to the point β. When the mmf is exactly zero, the material displays the **remanent** (or **residual**) **magnetization** B_r. To bring the flux density to zero, we must further decrease the mmf (i.e., produce a negative current) until the field intensity reaches the value $-H_0$ (point γ on the curve). As the mmf is made more negative, the curve eventually reaches the point α'. If the excitation current to the coil is now increased, the magnetization curve will follow the path $\alpha' = \beta' = \gamma' = \alpha$, eventually returning to the original point in the *B*-*H* plane, but via a different path.

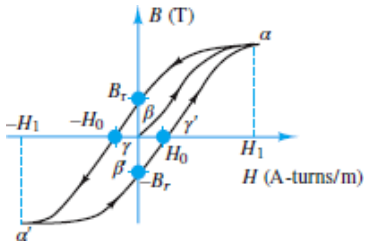


Figure 14.32 Hysteresis in magnetization curves

The result of this process, by which an *excess mmf* is required to magnetize or demagnetize the material, is a net energy loss. It is difficult to evaluate this loss; however, it can be shown that it is related to the area between the curves of Figure 14.32. Experimental techniques exist that measure these losses.

Figure 14.33 depicts magnetization curves for three very common ferromagnetic materials: cast iron, cast steel, and sheet steel.

Page 832

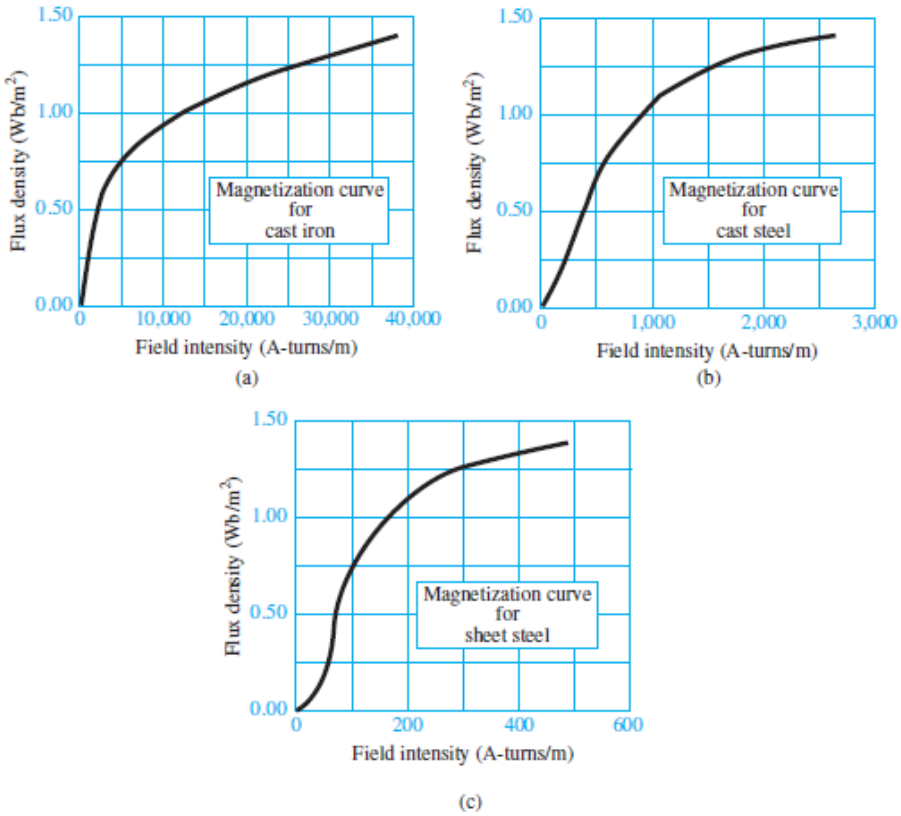


Figure 14.33 Magnetization curves for (a) cast iron, (b) cast steel, and (c) sheet steel

14.4 TRANSFORMERS

One of the more common magnetic structures in everyday applications is the transformer. The ideal transformer was introduced in Chapter 13 as a device that can step an AC voltage up or down by a fixed ratio, with a corresponding decrease or increase in current. The structure of a simple magnetic transformer is shown in Figure 14.34, which illustrates that a transformer is very similar to the magnetic circuits described earlier in this chapter. Coil L_1 represents the input side of the transformer, while coil L_2 is the output coil; both coils are wound around the same magnetic structure, which we show here to be similar to the "square doughnut" of the earlier examples.

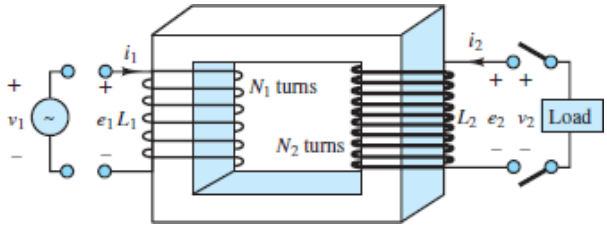


Figure 14.34 Structure of a transformer

Page 833

The ideal transformer is defined by the same set of assumptions made earlier. The flux is confined to the core, the flux links all turns of both coils, and the permeability of the core is infinite. The last assumption is equivalent to stating that an arbitrarily small mmf is sufficient to establish a flux in the core. In addition, we assume that the ideal transformer coils offer negligible resistance to current.

A time-varying voltage applied to the primary side of the transformer results in a corresponding time-varying current in L_1. This current acts as an mmf and causes a time-varying flux in the structure. This flux will induce an emf across the secondary coil! Without the need for a direct electrical connection, the transformer can couple a source voltage across to the primary winding to the secondary winding, which is connected to a load; the coupling occurs by means of the magnetic field acting on both coils. Thus, a transformer operates by converting electric energy to magnetic, and then back to electric. The following derivation illustrates this viewpoint in the ideal case (no loss of energy) and compares the result with the definition of the ideal transformer in Chapter 13.

If a time-varying voltage source is connected to the input side, then by virtue of Faraday's law, a corresponding time-varying flux $d\phi/dt$ is established in coil L_1:

$$e_1 = N_1 \frac{d\phi}{dt} = v_1 \tag{14.33}$$

But since the flux thus produced also links coil L_2, an emf is induced across the output coil as well:

$$e_2 = N_2 \frac{d\phi}{dt} = v_2 \tag{14.34}$$

This induced emf can be measured as the voltage v_2 at the output terminals, and one can readily see that the ratio of the open-circuit output voltage to input-terminal voltage is

$$\frac{v_2}{v_1} = \frac{N_2}{N_1} = N \tag{14.35}$$

A load current i_2, and its corresponding mmf $\mathcal{F}_2 = N_2 i_2$, is produced when a load is connected to the output terminals in Figure 14.34. The mmf would cause the flux in the core to change; however, this is not possible since a change in ϕ would cause a corresponding change in the voltage induced across the input coil. But this voltage is determined (fixed) by the source v_1 so that the input coil is forced to generate a **counter-mmf** to oppose the mmf of the output coil drawing a current i_1 from the source v_1 such that:

$$i_1 N_1 = i_2 N_2 \tag{14.36}$$

or

$$\frac{i_2}{i_1} = \frac{N_1}{N_2} = \alpha = \frac{1}{N} \tag{14.37}$$

where α is the ratio of primary to secondary turns (the transformer ratio) and N_1 and N_2 are the primary and secondary turns, respectively. If there were any net difference between the input and output mmf, the flux balance required by the input voltage source would not be satisfied. Thus, the two magnetomotive forces must be equal. Page 834As you can easily verify, these results are the same as in Chapter 13; in particular, the ideal transformer does not dissipate any power, since

$$v_1 i_1 = v_2 i_2 \tag{14.38}$$

Note the distinction we have made between the induced voltages (emf's) e and the terminal voltages v. In general, these are not the same.

The results obtained for the ideal case do not completely represent the physical nature of transformers. A number of loss mechanisms need to be included in a practical transformer model, to account for the effects of leakage flux, for various magnetic core losses (e.g., hysteresis), and for the unavoidable resistance of the wires that form the coils.

Commercial transformer ratings are usually given on the **nameplate**, which indicates the following normal operating conditions:

- Primary-to-secondary voltage ratio
- Design frequency of operation
- (Apparent) rated output power

For example, a typical nameplate might read 480:240 V, 60 Hz, 2 kVA. The voltage ratio can be used to determine the turns ratio, while the rated output power represents the continuous power level that can be sustained without overheating. It is important that this power be rated as the apparent power in kilovoltamperes, rather than real power in kilowatts, since a load with low power factor would still draw current and therefore operate near rated power. Another important performance characteristic of a practical transformer is its **power efficiency**, defined by:

$$\text{Power efficiency } \eta = \frac{\text{Output power}}{\text{Input power}} \tag{14.39}$$

EXAMPLE 14.7 Transformer Nameplate

Problem

Determine the turns ratio and the rated currents of a transformer from nameplate data.

Solution

Known Quantities: Nameplate data.

Find: $\alpha = N_1/N_2$; I_1; I_2.

Schematics, Diagrams, Circuits, and Given Data: Nameplate data: 120 V/480 V; 48 kVA; 60 Hz.

Assumptions: Assume an ideal transformer.

Analysis: The first element in the nameplate data is a pair of voltages, indicating the primary and secondary voltages for which the transformer is rated. The ratio α is found as follows:

Page 835

$$\alpha = \frac{N_1}{N_2} = \frac{480}{120} = 4$$

To find the primary and secondary currents, we use the kilovoltampere rating (apparent power) of the transformer:

$$I_1 = \frac{|S|}{V_1} = \frac{48 \text{ kVA}}{480 \text{ V}} = 100 \text{ A} \qquad I_2 = \frac{|S|}{V_2} = \frac{48 \text{ kVA}}{120 \text{ V}} = 400 \text{ A}$$

Comments: In computing the rated currents, we have assumed that no losses take place in the transformer; in fact, there will be losses due to coil resistance and magnetic core effects. These losses result in heating of the transformer and limit its rated performance.

CHECK YOUR UNDERSTANDING

The high-voltage side of a transformer has 500 turns, and the low-voltage side has 100 turns. When the transformer is connected as a step-down transformer, the load current is 12 A. Calculate: (a) the turns ratio α; and (b) the primary current. Then, (c) Calculate the turns ratio if the transformer is used as a step-up transformer.

The output of a transformer under certain conditions is 12 kW. The copper losses are 189 W, and the core losses are 52 W. Calculate the efficiency of this transformer.

Answer: (a) $\alpha = 5$; (b) $I_1 = I_2/\alpha = 2.4$ A; (c) $\alpha = 0.2$; $\eta = 98$ percent

LO

EXAMPLE 14.8 Impedance Transformer

Problem

Find the equivalent load impedance seen by the voltage source (i.e., reflected from secondary to primary) for the transformer of Figure 14.35.

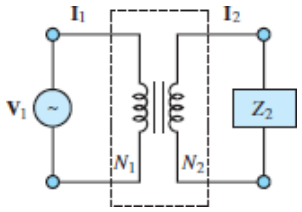


Figure 14.35 Ideal transformer

Solution

Known Quantities: Transformer turns ratio α.

Find: Reflected impedance Z'_2.

Assumptions: Assume an ideal transformer.

Analysis: By definition, the load impedance is equal to the ratio of secondary phasor voltage and current:

$$Z_2 = \frac{\mathbf{V}_2}{\mathbf{I}_2}$$

To find the reflected impedance, we can express the above ratio in terms of the primary voltage and current:

$$Z_2 = \frac{\mathbf{V}_2}{\mathbf{I}_2} = \frac{\mathbf{V}_1/\alpha}{\alpha \mathbf{I}_1} = \frac{1}{\alpha^2}\frac{\mathbf{V}_1}{\mathbf{I}_1}$$

Page 836

where the ratio $\mathbf{V}_1/\mathbf{I}_1$ is the impedance seen by the source at the primary coil, that is, the *reflected load impedance* seen by the primary (source) side of the circuit. Thus, we can write the load impedance Z_2 in terms of the primary circuit voltage and current; we call this the *reflected impedance* Z'_2:

$$Z_2 = \frac{1}{\alpha^2}\frac{\mathbf{V}_1}{\mathbf{I}_1} = \frac{1}{\alpha^2}Z_1 = \frac{1}{\alpha^2}Z'_2$$

Thus, $Z'_2 = \alpha^2 Z^2$. Figure 14.36 depicts the equivalent circuit with the load impedance reflected back to the primary.

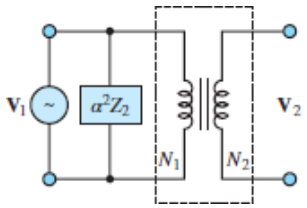


Figure 14.36 Equivalent reflected circuit for impedance transformer.

Comments: The equivalent reflected circuit calculations are convenient because all circuit elements can be referred to a single set of variables (i.e., only primary or secondary voltages and currents).

CHECK YOUR UNDERSTANDING

The output impedance of a servo amplifier is 250 Ω. The servomotor that the amplifier must drive has an impedance of 2.5 Ω. Calculate the turns ratio of the transformer required to match these impedances.

Answer: $a = 10$

14.5 ELECTROMECHANICAL ENERGY CONVERSION

From the material developed thus far, it should be apparent that electromagnetomechanical devices are capable of converting mechanical forces and displacements to electromagnetic energy, and that the converse is also possible. The objective of this section is to formalize the basic principles of energy conversion in electromagnetomechanical systems, and to illustrate its usefulness and potential for application by presenting several examples of **energy transducers**. A transducer is a device that can convert electric to mechanical energy (in this case, it is often called an **actuator**), or vice versa (in which case it is called a **sensor**).

Several physical mechanisms permit conversion of electric to mechanical energy and back, including the **piezoelectric effect**,[3] consisting of the generation of a change in electric field in the presence of strain in certain crystals (e.g., quartz), and **electrostriction** and **magnetostriction**, in which changes in the dimension of certain materials lead to a change in their electrical (or magnetic) properties. This chapter is

concerned only with transducers in which electric energy is converted to mechanical energy through the coupling of a magnetic field. It is important to note that all rotating machines (motors and generators) fit the basic definition of electromechanical transducers we have just given.

Page 837

Forces in Magnetic Structures

Mechanical forces can be converted to electric signals, and vice versa, by means of the coupling provided by energy stored in the magnetic field. In this subsection, we discuss the computation of mechanical forces and of the corresponding electromagnetic quantities of interest; these calculations are of great practical importance in the design and application of electromechanical actuators. For example, a problem of interest is the computation of the current required to generate a given force in an electromechanical structure. This is the kind of application that is likely to be encountered by the engineer in the selection of an electromechanical device for a given task.

As already seen in this chapter, an electromechanical system includes an electrical system, interacting through a magnetic field. Figure 14.37 illustrates the coupling between the electrical and mechanical systems. In the mechanical system, energy loss can occur because of the heat developed as a consequence of *friction*, while in the electrical system, analogous losses are incurred because of *resistance.* Loss mechanisms are also present in the magnetic coupling medium, since *eddy current losses* and *hysteresis losses* are unavoidable in ferromagnetic materials. Either system can supply energy, and either system can store energy. Thus, the figure depicts the flow of energy from the electrical to the mechanical system, accounting for these various losses. The same flow could be reversed if mechanical energy were converted to electrical form.

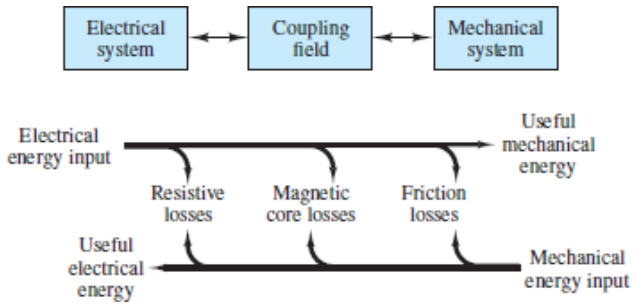


Figure 14.37 Losses in electromechanical energy conversion.

Moving-Iron Transducers

One important class of electromagnetomechanical transducers is that of **moving-iron transducers**, which include common devices such as electromagnets, solenoids, and relays. The simplest example of a moving-iron transducer is Figure 14.38, in which the U-shaped element is fixed and the bar is movable. In the following paragraphs, we shall derive a relationship between the current applied to the coil, the displacement of the movable bar, and the magnetic force acting in the air gap.

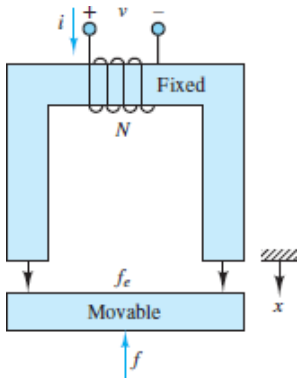


Figure 14.38 Basic electromagnet.

The principle that will be applied throughout the section is that for a mass to be displaced, some work needs to be done; this work corresponds to a change in the energy stored in the electromagnetic field, which causes the mass to be displaced. With reference to Figure 14.38, let f_e represent the magnetic force acting on the bar and x the displacement of the bar, in the direction shown. Then the net work W_m into the electromagnetic field is equal to the sum of the work done by the electric circuit plus the work done by the mechanical system. Therefore, for an Page 838incremental amount of work, we can write

$$dW_m = ei\,dt - f_e\,dx \tag{14.40}$$

where e is the electromotive force across the coil and the minus sign is due to the sign convention indicated in Figure 14.38. Recalling that the emf e is equal to the derivative of the flux linkage (equation 14.16), we can further expand equation 14.40 to obtain

$$dW_m = ei\,dt - f_e\,dx = i\frac{d\lambda}{dt}dt - f_e\,dx = i\,d\lambda - f_e\,dx \tag{14.41}$$

or

$$f_e\,dx = i\,d\lambda - dW_m \tag{14.42}$$

Now, observe that the flux in the magnetic structure of Figure 14.38 depends on two variables, which are in effect independent: the current through the coil and the displacement of the bar. Each of these variables can cause the magnetic flux to change. Similarly, the energy stored in the electromagnetic field is also dependent on both current and displacement. Thus we can rewrite equation 14.42 as follows:

$$f_e\,dx = i\left(\frac{\partial\lambda}{\partial i}di + \frac{\partial\lambda}{\partial x}dx\right) - \left(\frac{\partial W_m}{\partial i}di + \frac{\partial W_m}{\partial x}dx\right) \tag{14.43}$$

Since i and x are independent variables, we can write

$$f_e = i\frac{\partial\lambda}{\partial x} - \frac{\partial W_m}{\partial x} \quad \text{and} \quad 0 = i\frac{\partial\lambda}{\partial i} - \frac{\partial W_m}{\partial i} \tag{14.44}$$

From the first expression in equation 14.44 we obtain the relationship

$$f_e = \frac{\partial}{\partial x}(i\lambda - W_m) = \frac{\partial}{\partial x}(W'_m) \tag{14.45}$$

where W'_m is the co-energy. Observe that the force acting to *push* the bar toward the electromagnet structure is of opposite sign to f_e, and assuming that $W_m = W'_m$, we can write

$$f = -f_e = -\frac{\partial}{\partial x}(W'_m) = -\frac{\partial W_m}{\partial x} \tag{14.46}$$

Equation 14.46 includes a very important assumption: The energy is equal to the co-energy. If you refer to Figure 14.7, you will realize that in general this is not true. Energy and co-energy are equal only if the λ-i relationship is linear. Thus, the useful result of equation 14.46, stating that the magnetic force acting on the moving iron is proportional to the rate of change of stored energy with displacement, applies only for *linear magnetic structures.*

Thus, to determine the forces present in a magnetic structure, it is necessary to compute the energy stored in the magnetic field. To simplify the analysis, we assume hereafter that the structures analyzed are magnetically linear. This is, of course, only an approximation, in that it neglects a number of practical aspects of electromechanical systems (e.g., the nonlinear λ-i curves described earlier, and the core losses typical of magnetic materials), but it permits relatively simple analysis of many useful magnetic structures. Thus, although the analysis method presented in this section is only approximate, it will serve the purpose of providing a feeling Page 839for the direction and the magnitude of the forces and currents present in electromechanical devices. On the basis of a linear approximation, it can be shown that the stored energy in a magnetic structure is given by

$$W_m = \frac{\phi \mathcal{F}}{2} \tag{14.47}$$

and since the flux and the mmf are related by the expression

$$\phi = \frac{Ni}{\mathcal{R}} = \frac{\mathcal{F}}{\mathcal{R}} \tag{14.48}$$

the stored energy can be related to the reluctance of the structure according to

$$W_m = \frac{\phi^2 \mathcal{R}(x)}{2} \tag{14.49}$$

where the reluctance has been explicitly shown to be a function of displacement, as is the case in a moving-iron transducer. Finally, then, we use the following approximate expression to compute the magnetic force acting on the moving iron:

$$f = -\frac{dW_m}{dx} = -\frac{\phi^2}{2}\frac{d\mathcal{R}(x)}{dx} \quad \text{magnetic force} \tag{14.50}$$

Examples 14.9, 14.10, and 14.12 illustrate the application of this approximate technique for the computation of forces and currents (the two problems of practical engineering interest to the user of such electromechanical systems) in some common devices. The Focus on Problem Solving box outlines the solution techniques for these classes of problems.

FOCUS ON PROBLEM SOLVING

ANALYSIS OF MOVING-IRON ELECTROMECHANICAL TRANSDUCERS

Calculation of current required to generate a given fo

1. Derive an expression for the reluctance of the structure as a function of air displacement: $\mathcal{R}(x)$

2. Express the magnetic flux in the structure as a function of the mmf (i.e., o current I) and of the reluctance $\mathcal{R}(x)$:

$$\phi = \frac{\mathcal{F}(i)}{\mathcal{R}(x)}$$

3. Compute an expression for the force, using the known expressions for the and for the reluctance:

$$|f| = \frac{\phi^2}{2}\frac{d\mathcal{R}(x)}{dx}$$

Page 840

4. Solve the expression in step 3 for the unknown current i.

Calculation of force generated due to transducer geometry and mmf

Repeat steps 1 through 3 above, substituting the known current to solve for the f f.

EXAMPLE 14.9 An Electromagnet

Problem

An electromagnet is used to collect and support a solid piece of steel, as shown in Figure 14.38. Calculate the *starting current* required to lift the load and the *holding current* required to keep the load in place once it has been lifted and is attached to the magnet. Assume that the cross-sectional areas of the electromagnet, load (bar), and air gap are equal.

Solution

Known Quantities: Geometry, magnetic permeability, number of coil turns, mass, acceleration of gravity, initial position of steel bar.

Find: Current required to lift the bar; current required to hold the bar in place.

Schematics, Diagrams, Circuits, and Given Data:

$N = 500$

$\mu_0 = 4\pi \times 10^{-7}$

$\mu_r = 10^4$ (equal for electromagnet and load)

Initial distance (air gap) = 0.5 m

Magnetic path length of electromagnet = $l_1 = 0.60$ m

Magnetic path length of movable load = $l_2 = 0.30$ m

Gap cross-sectional area = 3×10^{-4} m^2

m = mass of load = 5 kg

$g = 9.8$ m/s^2

Assumptions: None.

Analysis: To compute the current we need to derive an expression for the force in the air gap. We use the equation

$$f_{\text{mech}} = \frac{\phi^2}{2}\frac{\partial \mathcal{R}(x)}{\partial x}$$

and calculate the reluctance, flux and force as follows:

$$\begin{aligned}\mathcal{R}(x) &= \mathcal{R}_{Fe} + \mathcal{R}_{\text{gap}} \\ &= \frac{2x}{\mu_0 A} + \frac{l_1 + l_2}{\mu_0 \mu_r A}\end{aligned}$$

Page 841

$$\phi = \frac{\mathcal{F}}{\mathcal{R}(x)} = \frac{Ni}{\left(\frac{2x}{\mu_0 A} + \frac{l_1+l_2}{\mu_0 \mu_r A}\right)}$$

$$\frac{\partial \mathcal{R}(x)}{\partial x} = \frac{2}{\mu_0 A} \Rightarrow f_{\text{mag}} = \frac{\phi^2}{2}\frac{\partial \mathcal{R}(x)}{\partial x} = \frac{(Ni)^2}{\left(\frac{2x}{\mu_0 A} + \frac{l_1+l_2}{\mu_0 \mu_r A}\right)^2}\frac{1}{\mu_0 A}$$

With this expression we can now calculate the current required to overcome the gravitational force when the load is 0.5 m away. The force we must overcome is mg = 49 N.

$$f_{\text{mag}} = \frac{(Ni)^2}{\left(\frac{2x}{\mu_0 A} + \frac{l_1+l_2}{\mu_0\mu_r A}\right)^2}\frac{1}{\mu_0 A} = f_{\text{gravity}}$$

$$i^2 = f_{\text{gravity}}\frac{\mu_0 A\left(\frac{2x}{\mu_0 A} + \frac{l_1+l_2}{\mu_0\mu_r A}\right)^2}{N^2} = 520 \times 10^3\,\text{A}^2 \qquad i = 721\ \text{A}$$

Finally, we calculate the holding current by letting $x = 0$:

$$f_{\text{mag}} = \frac{(Ni)^2}{\left(\frac{l_1+l_2}{\mu_0\mu_r A}\right)^2}\frac{1}{\mu_0 A} = f_{\text{gravity}}$$

$$i^2 = f_{\text{gravity}}\frac{\mu_0 A\left(\frac{l_1+l_2}{\mu_0\mu_r A}\right)^2}{N^2} = 4.21 \times 10^{-3}\,\text{A}^2$$

$$i = 64.9\ \text{mA}$$

Comments: Note how much smaller the holding current is than the lifting current.

One of the more common practical applications of the concepts discussed in this section is the solenoid. Solenoids find application in a variety of electrically controlled valves. The action of a solenoid valve is such that when it is energized, the plunger moves in such a direction as to permit the flow of a fluid through a conduit, as shown schematically in Figure 14.39.

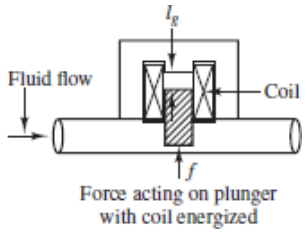


Figure 14.39 Application of the solenoid as a valve

Examples 14.10 and 14.11 illustrate the calculations involved in the determination of forces and currents in a solenoid.

EXAMPLE 14.10 A Solenoid

Problem

Figure 14.40 depicts a simplified representation of a solenoid. The restoring force for the plunger is provided by a spring.

1. Derive a general expression for the force exerted on the plunger as a function of the plunger position x.
2. Determine the mmf required to pull the plunger to its end position ($x = a$).

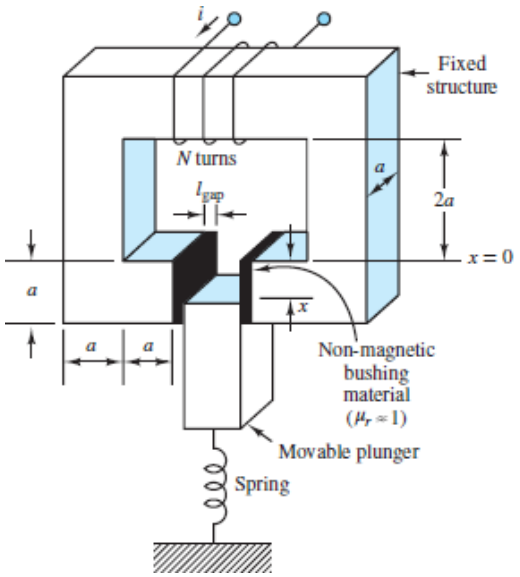


Figure 14.40 A solenoid

Solution

Known Quantities: Geometry of magnetic structure; spring constant.

Page 842

Find: f; mmf.

Schematics, Diagrams, Circuits, and Given Data: $a = 0.01$ m; $l_{29} = 0.001$ m; $k = 10$ N/m.

Assumptions: Assume that the reluctance of the iron is negligible; neglect fringing. At $x = 0$ the plunger is in the gap by an infinitesimal displacement ε.

Analysis:

1. *Force on the plunger.* To compute a general expression for the magnetic force exerted on the plunger, we need to derive an expression for the force in the air gap. Using equation 14.50, we see that we need to compute the reluctance of the structure and the magnetic flux to derive an expression for the force.

Since we are neglecting the iron reluctance, we can write the expression for the reluctance as follows. Note that the area of the gap is variable, depending on the position of the plunger, as shown in Figure 14.41.

$$\mathcal{R}_{\text{gap}}(x) = 2 \times \frac{l_{\text{gap}}}{\mu_0 A_{\text{gap}}} = \frac{2l_{\text{gap}}}{\mu_0 ax}$$

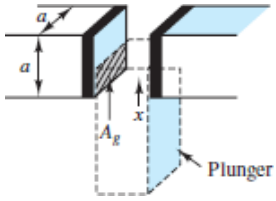


Figure 14.41 Detail of solenoid structure.

The derivative of the reluctance with respect to the displacement of the plunger can then be computed to be

$$\frac{d\mathcal{R}_{\text{gap}}(x)}{dx} = \frac{-2l_{\text{gap}}}{\mu_0 ax^2}$$

Knowing the reluctance, we can calculate the magnetic flux in the structure as a function of the coil current:

$$\phi = \frac{Ni}{\mathcal{R}(x)} = \frac{Ni\mu_0 ax}{2l_{\text{gap}}}$$

The force in the air gap is given by

$$f_{\text{gap}} = \frac{\phi^2}{2}\frac{d\mathcal{R}(x)}{dx} = \frac{(Ni\mu_0 ax)^2}{8l_{\text{gap}}^2}\frac{-2l_{\text{gap}}}{\mu_0 ax^2} = -\frac{\mu_0 a(Ni)^2}{4l_{\text{gap}}} = kX$$

Page 843

Thus, the force in the gap is proportional to the square of the current and does not vary with plunger displacement.

2. *Calculation of magnetomotive force.* To determine the required magnetomotive force, we observe that the magnetic force must overcome the mechanical (restoring) force generated by the spring. Thus, $f_{\text{gap}} = kx = ka$. For the stated values, $f_{\text{gap}} = (10 \text{ N/m}) \times (0.01 \text{ m}) = 0.1$ N, and

$$Ni = \sqrt{\frac{4l_{\text{gap}}f_{\text{gap}}}{\mu_0 a}} = \sqrt{\frac{4 \times 0.001 \times 0.1}{4\pi \times 10^{-7} \times 0.01}} = 178 \text{ A-turns}$$

The required mmf can be most effectively realized by keeping the current value relatively low and using a large number of turns.

Comments: The same mmf can be realized with an infinite number of combinations of current and number of turns; however, there are tradeoffs involved. If the current is very large (and the number of turns small), the required wire diameter will be very large. Conversely, a small current will require a small wire diameter and a large number of turns. A homework problem explores this tradeoff.

CHECK YOUR UNDERSTANDING

A solenoid is used to exert force on a spring. Estimate the position of the plunger if the number of turns in the solenoid winding is 1,000 and the current going into the winding is 40 mA. Use the same values as in Example 14.10 for all other variables.

Answer: $x = 0.5$ mm

Practical Facts About Solenoids

Solenoids can be used to produce linear or rotary motion, in either the *push* or the *pull* mode. The most common solenoid types are listed here:

1. *Single-action linear* (push or pull). Linear stroke motion, with a restoring force (e.g., from a spring), to return the solenoid to the neutral position.
2. *Double-acting linear.* Two solenoids back to back can act in either direction. The restoring force is provided by another mechanism (e.g., a spring).
3. *Mechanical latching solenoid* (bistable). An internal latching mechanism holds the solenoid in place against the load.
4. *Keep solenoid.* Fitted with a permanent magnet so that no power is needed to hold the load in the pulled-in position. Plunger is released by applying a current pulse of opposite polarity to that required to pull in the plunger.
5. *Rotary solenoid.* Constructed to permit rotary travel. Typical range is 25 to 95°. Return action via mechanical means (e.g., a spring).
6. *Reversing rotary solenoid.* Rotary motion is from one end to the other; when the solenoid is energized again, it reverses direction.

Solenoid power ratings are dependent primarily on the current required by the coil, and on the coil resistance. The I^2R is the primary power sink, and solenoids are therefore limited by the heat they can dissipate. Solenoids can operate in continuous or pulsed mode. The power rating depends on the mode of operation, and can be increased by adding *hold-in resistors* to the circuit to reduce the *holding current* required for continuous operation. The hold resistor is switched into the circuit once the *pull-in* current required to pull the plunger has been applied and the plunger has moved into place. The holding current can be significantly smaller than the pull-in current.

A common method to reduce the solenoid holding current employs a normally closed (NC) switch in parallel with a hold-in resistor. In Figure 14.42, when the pushbutton (PB) closes the circuit, full voltage is applied to the solenoid coil, bypassing the resistor through the NC switch. When the solenoid closes, the NC switch opens, connecting the resistor in series with the coil. The resistor will now limit the current to the value required to hold the solenoid in position. Note the diode "snubber" circuit to shunt the reverse current when the solenoid is deenergized.

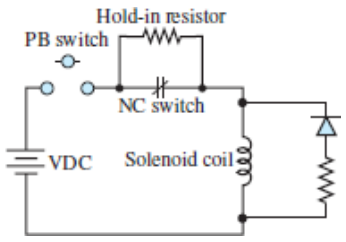


Figure 14.42 Practical solenoid circuit.

Another electromechanical device that finds common application in industrial practice is the relay. The relay is an electromechanical switch that permits the opening and closing of electrical contacts by means of an electromagnetic structure similar to those discussed earlier in this section.

A relay such as would be used to start a high-voltage single-phase motor is shown in Figure 14.43. The magnetic structure has dimensions equal to 1 cm on all sides, and the transverse dimension is 8 cm. The relay works as follows. When the pushbutton is pressed, an electric current flows through the coil and generates a field in the magnetic structure. The resulting force draws the movable part toward the fixed part, causing an electrical contact to be made. The advantage of the relay is that a relatively low-level current can be used to control the opening and closing of a

circuit Page 845that can carry large currents. In this particular example, the relay is energized by a 120-VAC contact, establishing a connection in a 240-VAC circuit. Such relay circuits are commonly employed to remotely switch large industrial loads.

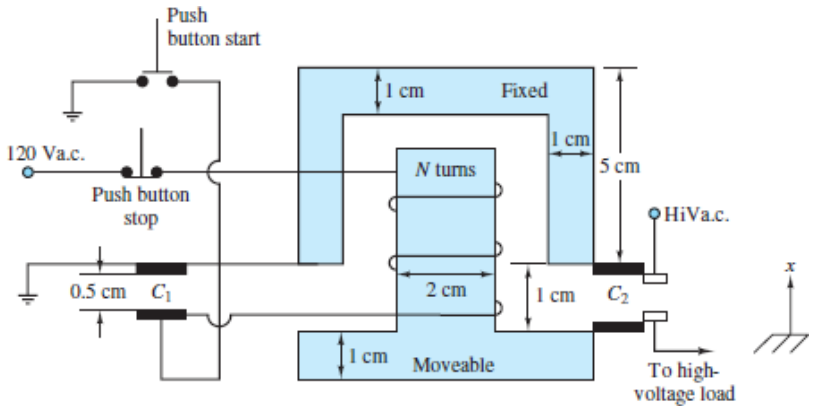
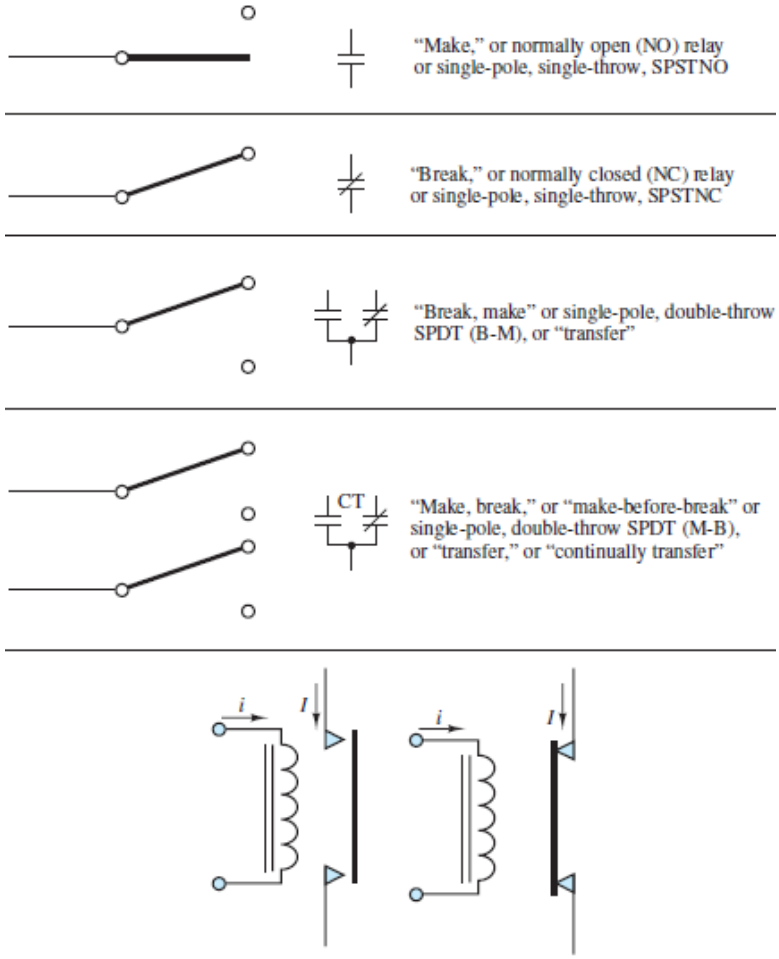


Figure 14.43 A relay

Circuit symbols for relays are shown in Figure 14.44. An example of the calculations that would typically be required in determining the mechanical and electrical characteristics of a simple relay are given in Example 14.11.



Basic operation of the electromechanical relay: The (small) coil current i causes the relay to close (or open) and enables (interrupts) the larger current I.
On the left: SPSTNO relay (magnetic field causes relay to close).
On the right: SPSTNC relay (magnetic field causes relay to open).

Figure 14.44 Circuit symbols and basic operation of relays

EXAMPLE 14.11 A Relay

Problem

Figure 14.45 depicts a simplified representation of a relay. Determine the current required for the relay to make contact (i.e., pull in the ferromagnetic plate) from a distance x.

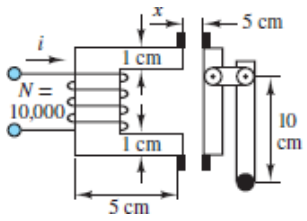


Figure 14.45 Relay circuit for Example 14.11.

Solution

Known Quantities: Relay geometry; restoring force to be overcome; distance between bar and relay contacts; number of coil turns.

Find: i.

Schematics, Diagrams, Circuits, and Given Data: $A_{gap} = (0.01\ m)^2$; $x = 0.05$ m; $f_{restore} = 5$ N; $N = 10{,}000$.

Assumptions: Assume that the reluctance of the iron is negligible; neglect fringing.

Analysis:

$$\mathcal{R}_{gap}(x) = \frac{2x}{\mu_0 A_{gap}}$$

The derivative of the reluctance with respect to the displacement of the plunger can then be computed as

$$\frac{d\mathcal{R}_{gap}(x)}{dx} = \frac{2}{\mu_0 A_{gap}}$$

Knowing the reluctance, we can calculate the magnetic flux in the structure as a function of the coil current:

$$\phi = \frac{Ni}{\mathcal{R}(x)} = \frac{Ni\mu_0 A_{gap}}{2x}$$

and the force in the air gap is given by

$$f_{gap} = \frac{\phi^2}{2}\frac{d\mathcal{R}(x)}{dx} = \frac{(Ni\mu_0 A_{gap})^2}{8x^2}\frac{2}{\mu_0 A_{gap}} = \frac{\mu_0 A_{gap}(Ni)^2}{4x^2}$$

The magnetic force must overcome a mechanical holding force of 5 N; thus,

$$f_{gap} = \frac{\mu_0 A_{gap} (Ni)^2}{4x^2} = f_{restore} = 5 \text{ N}$$

or

$$i = \frac{1}{N}\sqrt{\frac{4x^2 f_{restore}}{\mu_0 A_{gap}}} = \frac{1}{10{,}000}\sqrt{\frac{4(0.05)^2 5}{4\pi \times 10^{-7} \times 0.0001}} = \pm 2 \text{ A}$$

Comments: The current required to close the relay is much larger than that required to hold the relay closed, because the reluctance of the structure is much smaller once the gap is reduced to zero.

Moving-Coil Transducers

Another important class of electromagnetomechanical transducers is that of **moving-coil transducers**. This class of transducers includes a number of common devices, such as microphones, loudspeakers, and all electric motors and generators. The aim of this section is to explain the relationship between a fixed magnetic field, the emf across the moving coil, and the forces and motions of the moving element of the transducer.

Page 847

The basic principle of operation of electromechanical transducers is that a magnetic field exerts a force on a charge moving through it. The equation describing this effect is

$$\mathbf{f} = q\mathbf{u} \times \mathbf{B} \tag{14.51}$$

which is a vector equation, as explained earlier. To correctly interpret equation 14.51, we must recall the right-hand rule and apply it to the transducer, illustrated in Figure 14.46, depicting a structure consisting of a sliding bar which makes contact with a fixed conducting frame. Although this structure does not represent a practical actuator, it will be a useful aid in explaining the operation of moving-coil transducers such as motors and generators. In Figure 14.46, and in all similar figures in this section, a small cross represents the "tail" of an arrow pointing into the page, while a dot represents an arrow pointing out of the page; this convention will be useful in visualizing three-dimensional pictures.

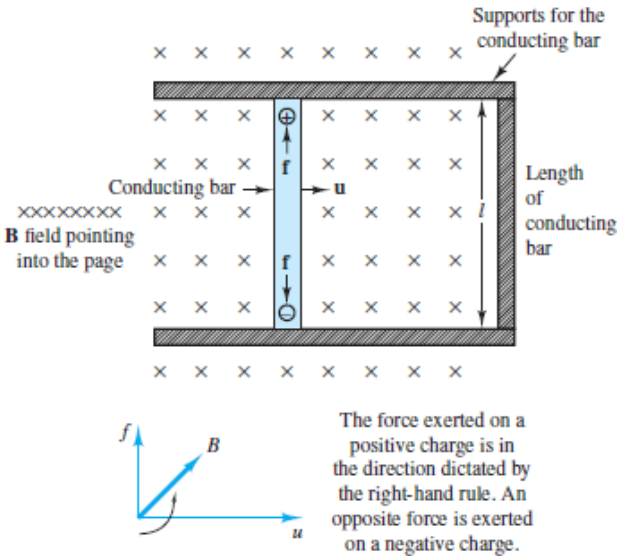


Figure 14.46 A simple electromechanical motion transducer

CHECK YOUR UNDERSTANDING

In the circuit in Figure 14.46, the conducting bar is moving with a velocity of 6 m/s. The flux density is 0.5 Wb/m^2, and l = 1.0 m. Find the magnitude of the resulting induced voltage.

Answer: 3 V

Motor Action

A moving-coil transducer can act as a motor when an externally supplied current flowing through the electrically conducting part of the transducer is converted to a force that can cause the moving part of the transducer to be displaced. Such a current would flow, for example, if the support of Figure 14.46 were made of Page 848conducting material, so that the conductor and the right-hand side of the support "rail" were to form a loop (in effect, a one-turn coil). To understand the effects of this current flow in the conductor, one must consider the fact that a charge moving at a velocity u' (along the conductor and perpendicular to the velocity of the conducting bar, as shown in Figure 14.47) corresponds to a current $i = dq/dt$ along the length l of the conductor. This fact can be explained by considering the current i along a differential element dl and writing

$$i\,dl = \frac{dq}{dt} \cdot u'\,dt \tag{14.52}$$

since the differential element dl would be traversed by the current in time dt at a velocity u'. Thus we can write

$$i\,dl = dq\,u' \tag{14.53}$$

or

$$il = q\,u' \tag{14.54}$$

for the geometry of Figure 14.47. From Section 14.1, the force developed by a charge moving in a magnetic field is, in general, given by

$$\mathbf{f} = q\mathbf{u} \times \mathbf{B} \tag{14.55}$$

For the term $q\mathbf{u}'$ we can substitute $i\mathbf{l}$, to obtain

$$\mathbf{f}' = i\mathbf{l} \times \mathbf{B} \tag{14.56}$$

Using the right-hand rule, we determine that the force $\mathbf{f}'$ generated by the current i is in the direction that would push the conducting bar to the left. The magnitude of this force is $f' = Bli$ if the magnetic field and the direction of the current are perpendicular. If they are not, then we must consider the angle γ formed by **B** and **l**; in the more general case,

$$f' = Bli \sin\gamma = Bli \text{ if } \gamma = 90° \qquad \textit{Bli} \text{ law} \tag{14.57}$$

The phenomenon we have just described is sometimes referred to as the ***Blilaw***.

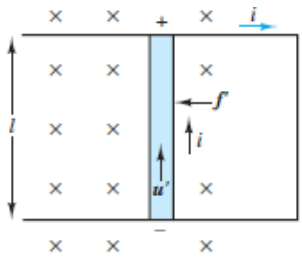


Figure 14.47 Simplified structure of moving-coil transducer.

Generator Action

The other mode of operation of a moving-coil transducer occurs when an external force causes the coil (i.e., the moving bar, in Figure 14.46) to be displaced. This external force is converted to an emf across the coil, as will be explained in the following paragraphs.

Since positive and negative charges are forced in opposite directions in the transducer of Figure 14.46, a potential difference will appear across the conducting bar; this potential difference is the electromotive force, or emf. The emf must be equal to the force exerted by the magnetic field. In short, the electric force per unit charge (or electric field) e/l must equal the magnetic force per unit charge $f/q = Bu$. Thus, the relationship

LO

$$e = Blu \qquad \textit{Blu} \text{ law} \tag{14.58}$$

Page 849

holds whenever **B**, **l**, and **u** are mutually perpendicular, as in Figure 14.48. If equation 14.58 is analyzed in greater depth, it can be seen that the product lu (length times velocity) is the area crossed per unit time by the conductor. If one visualizes the conductor as "cutting" the flux lines into the base in Figure 14.47, it can be concluded that the electromotive force is equal to the *rate at which the conductor "cuts" the magnetic lines of flux.* It will be useful for you to carefully absorb this notion of conductors cutting lines of flux, since this greatly simplifies the understanding of the material in this section and in Chapter 15.

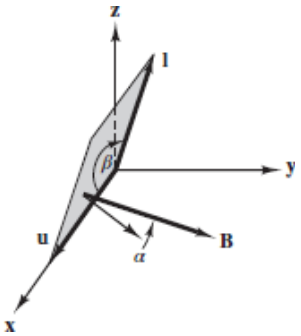


Figure 14.48 When magnetic flux, current and velocity vectors are mutually perpendicular, $e = Blu$.

In general, **B**, **l**, and **u** are not necessarily perpendicular. In this case one needs to consider the angles formed by the magnetic field with the normal to the plane containing **l** and **u**, and the angle between **l** and **u**. The former is angle α of Figure 14.48; the latter is angle β in the same figure. It should be apparent that the optimum values of α and β are 0° and 90°, respectively. Thus, most practical devices are constructed with these values of α and β. Unless otherwise noted, it will be tacitly assumed that this is the case. The***Blu*** **law** just illustrated explains how a moving conductor in a magnetic field can generate an electromotive force.

To summarize the electromechanical energy conversion that takes place in the simple device of Figure 14.46, we must note now that the presence of a current in the loop formed by the conductor and the rail requires that the conductor move to the right at a velocity u (*Blu* law), thus cutting the lines of flux and generating the emf that gives rise to current i. On the other hand, the same current causes a force f' to be exerted on the conductor (*Bli* law) in the direction opposite to the movement of the conductor. Thus, it is necessary that an *externally applied force* f_{ext} exist to cause the conductor to move to the right with a velocity u. The external force must overcome the force f'. This is the basis of electromechanical energy conversion.

An additional observation we must make at this point is that the current i flowing around a closed loop generates a magnetic field, as explained in Section 14.1. Since this additional field is generated by a one-turn coil in our illustration, it is reasonable to assume that it is negligible with respect to the field already present (perhaps established by a permanent magnet). Finally, we must consider that this coil links a certain amount of flux, which changes as the conductor moves from left to right. The area crossed by the moving conductor in time dt is

$$dA = lu\,dt \tag{14.59}$$

so that if the flux density B is uniform, the rate of change of the flux linked by the one-turn coil is

$$\frac{d\phi}{dt} = B\frac{dA}{dt} = Blu \tag{14.60}$$

In other words, the *rate of change* of the flux linked by the conducting loop is equal to the emf generated in the conductor. You should realize that this statement simply confirms Faraday's law.

It was briefly mentioned that the *Blu* and *Bli* laws indicate that, thanks to the coupling action of the magnetic field, a conversion of mechanical to electric energy —or the converse—is possible. The simple structures of Figures 14.46 and 14.47 can, again, serve as an illustration of this energy conversion process, although we have not yet indicated how these idealized structures can be converted to a practical

device. In this section we begin to introduce some physical considerations. Before we proceed any further, we should try to compute the power—electric and Page 850mechanical—that is generated (or is required) by our ideal transducer. The electric power is given by

$$P_E = ei = Blui \quad \text{W} \tag{14.61}$$

while the mechanical power required, say, to move the conductor from left to right is given by the product of force and velocity:

$$P_M = f_{\text{ext}}u = Bliu \quad \text{W} \tag{14.62}$$

The principle of conservation of energy states that in this ideal (lossless) transducer we can convert a given amount of electric energy to mechanical energy, or vice versa. We can utilize the structure of Figure 14.46 to illustrate this reversible action. If the closed path containing the moving conductor is now formed from a closed circuit containing a resistance R and a battery V_B, as shown in Figure 14.49, the externally applied force f_{ext} generates a positive current i into the battery provided that the emf is greater than V_B. When $e = Blu > V_B$, the ideal transducer acts as a *generator.* For any given set of values of B, l, R, and V_B, there will exist a velocity u for which the current i is positive. If the velocity is lower than this value—that is, if $e = Blu < V_B$—then the current i is negative, and the conductor is forced to move to the right. In this case the battery acts as a source of energy and the transducer acts as a *motor* (i.e., electric energy drives the mechanical motion).

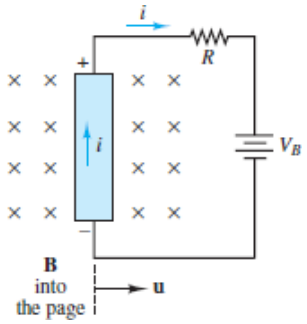


Figure 14.49 Motor and generator action in an ideal transducer

In practical transducers, we must be concerned with the inertia, friction, and elastic forces that are invariably present on the mechanical side of the transducer. Similarly, on the electrical side we must account for the inductance of the circuit, its resistance, and possibly some capacitance. Consider the structure of Figure 14.50. In the figure, the conducting bar has been placed on a surface with a coefficient of sliding friction b; it has a mass m and is attached to a fixed structure by means of a

spring with spring constant *k*. The equivalent circuit representing the coil inductance and resistance is also shown.

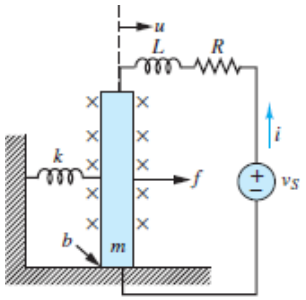


Figure 14.50 A more realistic representation of the transducer of Figure 14.49

If we recognize that $u = dx/dt$ in the figure, we can write the equation of motion for the conductor as

$$m\frac{du}{dt} + bu + \frac{1}{k}\int u\,dt = f = Bli \tag{14.63}$$

where the *Bli* term represents the driving input that causes the mass to move. The driving input in this case is provided by the electric energy source v_S; thus the transducer acts as a motor, and f is the electromechanical force acting on the mass of the conductor. On the electrical side, the circuit equation is

$$v_S - L\frac{di}{dt} - Ri = e = Blu \tag{14.64}$$

Equations 14.63 and 14.64 could then be solved by knowing the excitation voltage v_S and the physical parameters of the mechanical and electric circuits. For example, if the excitation voltage were sinusoidal, with

$$v_S(t) = V_S \cos \omega t$$

and the field density were constant

$$B = B_0$$

then we could postulate sinusoidal solutions for the transducer velocity u and current i:

$$u = U\cos(\omega t + \theta_u) \qquad i = I\cos(\omega t + \theta_i) \tag{14.65}$$

and use phasor notation to solve for the unknowns (U, I, θ_u, θ_i).

Page 851

The results obtained in the present section apply directly to transducers that are based on translational (linear) motion. These basic principles of electromechanical energy conversion and the analysis methods developed in the section are next applied to practical transducers in a few examples. A Focus on Problem Solving box outlines the analysis procedure for moving-coil transducers.

FOCUS ON PROBLEM SOLVING

ANALYSIS OF MOVING-COIL ELECTROMECHANICAL TRANSDUCERS

1. Apply KVL to write the differential equation for the electrical subsys including the back emf ($e = Blu$) term.
2. Apply Newton's second law to write the differential equation for the mechar subsystem, including the magnetic force $f = Bli$ term.
3. Use a Laplace transform on the two coupled differential equations to formul system of linear algebraic equations, and solve for the desired mechanical electrical variables.

EXAMPLE 14.12 A Loudspeaker

Problem

A loudspeaker, shown in Figure 14.51, uses a permanent magnet and a moving coil to produce the vibrational motion that generates the pressure waves we perceive as sound. Vibration of the loudspeaker is caused by changes in the input current to a coil; the coil is, in turn, coupled to a magnetic structure that can produce time-varying forces on the speaker diaphragm. A simplified model for the mechanics of the speaker is also shown in Figure 14.51. The force exerted on the coil is also

exerted on the mass of the speaker diaphragm, as shown in Figure 14.52, which depicts a free-body diagram of the forces acting on the loudspeaker diaphragm.

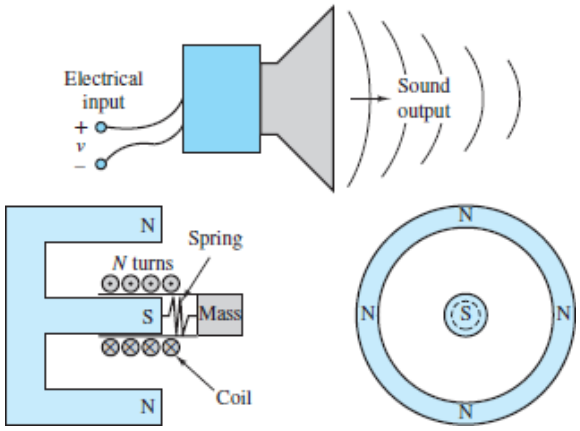


Figure 14.51 Loudspeaker

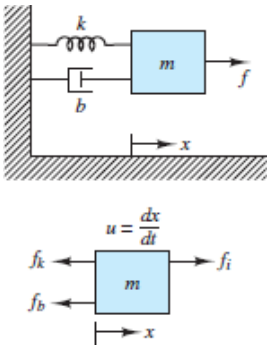


Figure 14.52 Forces acting on loudspeaker diaphragm

Page 852

The force exerted on the mass f_i is the magnetic force due to current flow in the coil. The electric circuit that describes the coil is shown in Figure 14.53, where L represents the inductance of the coil, R represents the resistance of the windings, and e is the emf induced by the coil moving through the magnetic field.

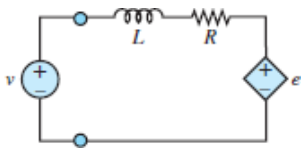


Figure 14.53 Model of transducer electrical side

Determine the frequency response $U(j\omega)/V(j\omega)$ of the speaker.

Solution

Known Quantities: Circuit and mechanical parameters; magnetic flux density; number of coil turns; coil radius.

Find: Frequency response of loudspeaker $U(j\omega)/V(j\omega)$.

Schematics, Diagrams, Circuits, and Given Data: Coil radius = 0.05 m; L = 10 mH; R = 8 Ω; m = 0.01 kg; b = 22.75 N-s^2/m; $k = 5 \times 10^4$ N/m; N = 47; B = 1 T.

Analysis: To determine the frequency response of the loudspeaker, we write the differential equations that describe the electrical and mechanical subsystems. We apply KVL to the electric circuit, using the circuit model of Figure 14.53, in which we have represented the Blu term (motional voltage) in the form of a *back electromotive force* e:

$$v - L\frac{di}{dt} - Ri - e = 0$$

or

$$L\frac{di}{dt} + Ri + Blu = v$$

Next, we apply Newton's second law to the mechanical system, consisting of a lumped mass representing the mass of the moving diaphragm m; an elastic (spring) term, which represents the elasticity of the diaphragm k; and a damping coefficient b, representing the frictional losses and aerodynamic damping affecting the moving diaphragm.

$$m\frac{du}{dt} = f_i - f_d - f_k = f_i - bu - kx$$

where $f_i = Bli$ and therefore

$$-Bli + m\frac{du}{dt} + bu + k\int_{-\infty}^{t} u(t')\,dt' = 0$$

Note that the two equations are *coupled;* that is, a mechanical variable appears in the electrical equation (velocity u in the Blu term), and an electrical variable appears in the mechanical equation (current i in the Bli term).

To derive the frequency response, we use the Laplace transform on the two equations to obtain

$$(sL + R)I(s) + BlU(s) = V(s)$$
$$-BlI(s) + \left(sm + b + \frac{k}{s}\right)U(s) = 0$$

We can write the above equations in matrix form and resort to Cramer's rule to solve for $U(s)$ as a function of $V(s)$:

$$\begin{bmatrix} sL + R & Bl \\ -Bl & sm + b + \frac{k}{s} \end{bmatrix} \begin{bmatrix} I(s) \\ U(s) \end{bmatrix} = \begin{bmatrix} V(s) \\ 0 \end{bmatrix}$$

with solution

$$U(s) = \frac{\det \begin{bmatrix} sL + R & V(s) \\ -Bl & 0 \end{bmatrix}}{\det \begin{bmatrix} sL + R & Bl \\ -Bl & sm + b + \frac{k}{s} \end{bmatrix}}$$

Page 853

or

$$\frac{U(s)}{V(s)} = \frac{Bl}{(sL + R)(sm + b + k/s) + (Bl)^2}$$
$$= \frac{Bls}{(Lm)s^3 + (Rm + Lb)s^2 + [Rb + kL + (Bl)^2]s + kR}$$

To determine the frequency response of the loudspeaker, we let $s \rightarrow j\omega$ in the above expression:

$$\frac{U(j\omega)}{V(j\omega)} = \frac{jBl\omega}{kR - (Rm + Lb)\omega^2 + j\{[Rb + kL + (Bl)^2]\omega - (Lm)\omega^3\}}$$

where $l = 2\pi Nr$, and substitute the appropriate numerical parameters:

$$\frac{U(j\omega)}{V(j\omega)} = \frac{j14.8\omega}{4 \times 10^5 - (0.08 + 0.2275)\omega^2 + j[(182 + 500 + 218)\omega - (10^{-4})\omega^3]}$$
$$= \frac{j14.8\omega}{4 \times 10^5 - 0.3075\omega^2 + j[(900)\omega - (10^{-4})\omega^3]}$$

The resulting frequency response is plotted in Figure 14.54.

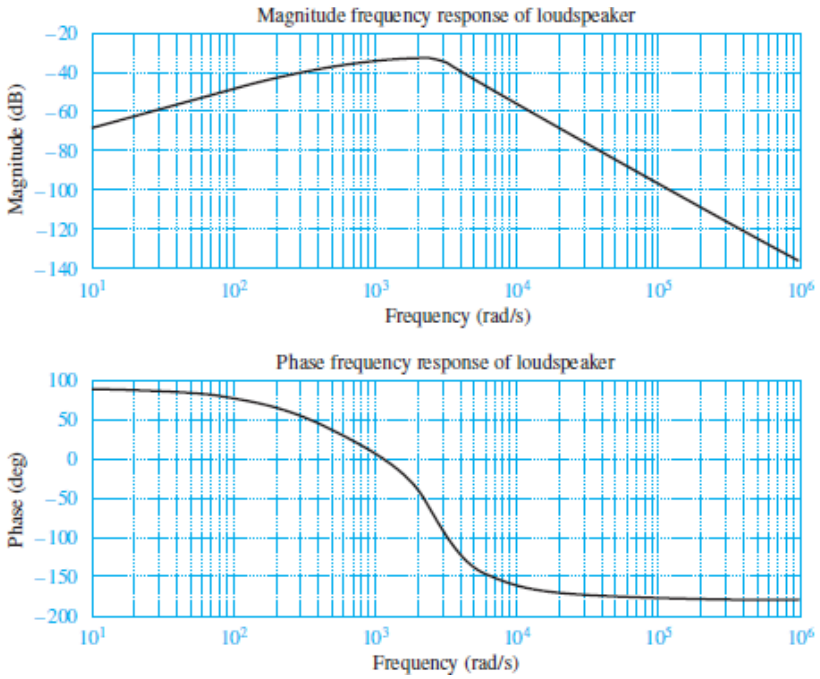


Figure 14.54 Frequency response of loudspeaker

CHECK YOUR UNDERSTANDING

In Example 14.12, we examined the frequency response of a loudspeaker. However, over time, permanent magnets may become demagnetized. Find the frequency response of the same loudspeaker if the permanent magnet has lost its strength to a point where $B = 0.95$ T.

Answer: $\dfrac{U(j\omega)}{V(j\omega)} = \dfrac{j(14.03\omega)}{(4 \times 10^{5} - 0.3075\omega^{2}) + j(889 \times 10^{-4}\omega^{3})}$

Page 854

FOCUS ON MEASUREMENTS

Seismic Transducer

Problem:

The device shown in Figure 14.55 is called a seismic transducer and can be used to measure the displacement, velocity, or acceleration of a body. The permanent magnet of mass m is supported on the case by a spring k, and there is some viscous damping b between the magnet and the case; the coil is fixed to the case. You may assume that the coil has length l and resistance and inductance R_{coil} and L_{coil}, respectively; the magnet exerts a magnetic field B. Find the transfer function between the output voltage v_{out} and the velocity of the body dx_c/dt. Note that $x(t)$ is not equal to zero when the system is at rest. We ignore this offset displacement.

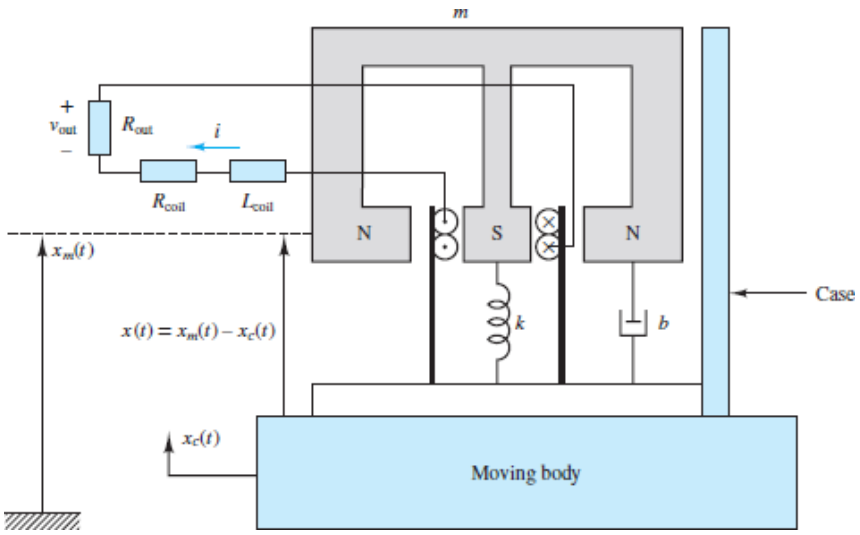


Figure 14.55 An electromagnetomechanical seismic transducer

Solution:

First we apply KVL around the electric circuit to write the differential equation describing the electrical systems:

$$L\frac{di}{dt} + (R_{coil} + R_{out})i + Bl\frac{dx}{dt} = 0$$

Also note that $v_{out} = -R_{out}i$. Next, we observe that the displacement of the magnet, x_m, is equal to the sum of the case displacement, x_c, and the relative displacement between the magnet and the case, $x(t)$: $x_m = x + x_c$. Apply Newton's second law to the mass of the magnet, m, we obtain

$$m\frac{d^2x_m}{dt^2} = -k(x_m - x_c) - b\left(\frac{dx_m}{dt} - \frac{dx_c}{dt}\right) - Bli$$

Substituting the relation $x_m = x + x_c$, we obtain

$$m\left(\frac{d^2x}{dt^2} + \frac{d^2x_c}{dt^2}\right) + kx + b\frac{dx}{dt} = -Bli$$

Page 855

From this expression we can now derive the transfer function between the displacement of the case, $X_c(s)$, and the output voltage, $V_{out}(s)$. Let $R = R_{coil} + R_{out}$. Then

$$(Ls + R)I(s) + Bls\,X(s) = 0$$

$$-BlI(s) + (ms^2 + bs + k)X(s) = -ms^2X_c(s)$$

$$I(s) = \frac{Blms^3X_c(s)}{mLs^3 + (bL + mR)s^2 + (kL + Rb + B^2l^2)s + kR}$$

Now, let the velocity of the case be $U_c(s) = sX_c(s)$; since $V_{out}(s) = -R_{out}I(s)$, the transfer function from case velocity to output voltage becomes

$$\frac{V_{out}(s)}{U_c(s)} = -\frac{Blm\,R_{out}s^2}{mLs^3 + (bL + mR)s^2 + (kL + Rb + B^2l^2)s + kR}$$

Conclusion

This chapter introduces electromechanical systems. Electromechanical devices include a variety of sensors and transducers that find common engineering

application in many fields. All electromechanical devices use the coupling between mechanical and electrical systems provided by a magnetic field. This magnetic coupling makes it possible to convert energy from electric to mechanical form, and back. Devices that convert electric to mechanical energy include all forms of electromagnetomechanical actuators, such as electromagnets, solenoids, relays, electrodynamic shakers, linear motors, and loudspeakers. Conversion from mechanical to electric energy results in generators, and various sensors that can detect mechanical displacement, velocity, or acceleration. Upon completing this chapter, you should have mastered the following learning objectives:

1. *Review the basic principles of electricity and magnetism.* The basic laws that govern electromagnetomechanical energy conversion are Faraday's law, stating that a changing magnetic field can induce a voltage, and Ampère's law, stating that a current flowing through a conductor generates a magnetic field.
2. *Use the concepts of reluctance and magnetic circuit equivalents to compute magnetic flux and currents in simple magnetic structures.* The two fundamental variables in the analysis of magnetic structures are the magnetomotive force and the magnetic flux; if some simplifying approximations are made, these quantities are linearly related through the reluctance parameter, in much the same way as voltage and current are related through resistance according to Ohm's law. This simplified analysis permits approximate calculation of forces and currents in electromagnetomechanical structures.
3. *Understand the properties of magnetic materials and their effects on magnetic circuit models.* Magnetic materials are characterized by a number of nonideal properties, which must be considered in a detailed analysis of any electromechanical transducer. The most important phenomena are saturation, eddy currents, and hysteresis.
4. *Use magnetic circuit models to analyze transformers.* One of the most common magnetic structures in use in electric power systems is the transformer. The methods developed in the earlier sections provide all the tools needed to perform an analysis of these important devices.

Page 856

5. *Model and analyze force generation in electromagnetomechanical systems. Analyze moving-iron transducers (electromagnets, solenoids, relays) and moving-coil transducers (electrodynamic shakers, loudspeakers, and seismic transducers).* Electromagnetomechanical transducers can be broadly divided into two categories: moving-iron transducers, which include all electromagnets, solenoids, and relays; and moving-coil transducers, which include loudspeakers, electrodynamic shakers, and all electric motors. Section 14.5 develops analysis and design methods for these devices.

HOMEWORK PROBLEMS

Section 14.1: Electricity and Magnetism

14.1 For the electromagnet of Figure P14.1:

a. Find the flux density in the core.

b. Sketch the magnetic flux lines and indicate their direction.

c. Indicate the north and south poles of the magnet.

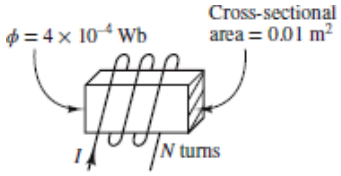


Figure P14.1

14.2 A single loop of wire carrying current I_2 is placed near the end of a solenoid having N turns and carrying current I_1, as shown in Figure P14.2. The solenoid is fastened to a horizontal surface, but the single coil is free to move. With the currents directed as shown, is there a resultant force on the single coil? If so, in what direction? Why?

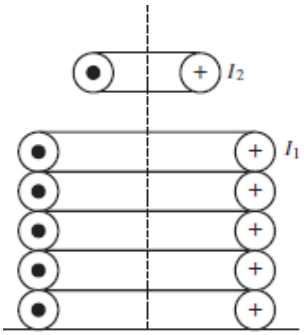


Figure P14.2

14.3 A practical LVDT is typically connected to a resistive load. Derive the LVDT equations in the presence of a resistive load R_L connected across the output terminals, using the results of the Focus on Measurements box, "Linear Variable Differential Transformer." Let R_S, L_S be the secondary coil parameters.

14.4 On the basis of the equations of the Focus on Measurements box, "Linear Variable Differential Transformer," and of the results of Problem 14.3, derive

the frequency response of the LVDT, and determine the range of frequencies for which the device will have maximum sensitivity for a given excitation. (*Hint:* Compute dv_{out}/dv_{ex}, and set the derivative equal to zero to determine the maximum sensitivity.)

14.5 An iron-core inductor has the following characteristic:

$$i = \frac{\lambda}{0.5 + \lambda}$$

a. Determine the energy, co-energy, and incremental inductance for λ = 1 V-s.

b. Given that the coil resistance is 1 Ω and that

$$i(t) = 0.625 + 0.01 \sin 400t \quad \text{A}$$

determine the voltage across the terminals on the inductor.

14.6 Repeat Problem 14.5 if

$$i = \frac{\lambda^2}{0.5 + \lambda^2}$$

14.7 An iron-core inductor has the characteristic shown in Figure P14.7:

a. Determine the energy and the incremental inductance for i = 1.0 A.

b. Given that the coil resistance is 2 Ω and that $i(t)$ = 0.5 sin 2π t, determine the voltage across the terminals of the inductor.

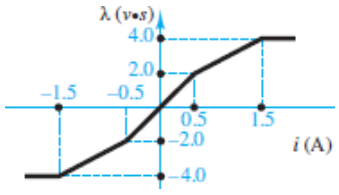


Figure P14.7

Page 857

14.8 Determine the reluctance of the structure of Figure 14.12 in the text if the cross-sectional area is A = 0.1 m^2 and μ_r = 2,000. Assume that each leg of the mean magnetic path is 0.1 m in length and that it runs through the exact center of the structure.

Section 14.2: Magnetic Circuits

14.9 a. Find the reluctance of a magnetic circuit if a magnetic flux $\phi = 4.2 \times 10^{-4}$ Wb is established by an impressed mmf of 400 A-turns.

b. Find the magnetizing force H in SI units if the magnetic circuit is 6 in long.

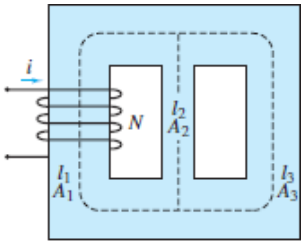
14.10 For the circuit shown in Figure P14.10:

a. Determine the reluctance values and show the magnetic circuit, assuming that $\mu = 3{,}000\mu_0$.

b. Determine the inductance of the device.

c. The inductance of the device can be modified by cutting an air gap in the magnetic structure. If a gap of 0.1 mm is cut in the arm of length l_3, what is the new value of inductance?

d. As the gap is increased in size (length), what is the limiting value of inductance? Neglect leakage flux and fringing effects.



$N = 100$ turns $A_2 = 25$ cm^2

$l_1 = 30$ cm $l_3 = 30$ cm

$A_1 = 100$ cm^2 $A_3 = 100$ cm^2

$l_2 = 10$ cm

Figure P14.10

14.11 The magnetic circuit shown in Figure P14.11 has two parallel paths. Find the flux and flux density in each leg of the magnetic circuit. Neglect fringing at the air gaps and any leakage fields. $N = 1{,}000$ turns, $i = 0.2$ A, $l_{g1} = 0.02$ cm, and $l_{g2} = 0.04$ cm. Assume the reluctance of the magnetic core to be negligible.

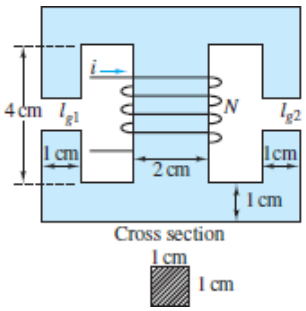


Figure P14.11

14.12 Find the current necessary to establish a flux of $\phi = 3 \times 10^{-4}$ Wb in the series magnetic circuit of Figure P14.12. Here $l_{iron} = l_{steel} = 0.3$ m, area (throughout) $= 5 \times 10^{-4}$ m^2, and $N = 100$ turns. Assume $\mu_r = 5{,}195$ for cast iron and $\mu_r = 1{,}000$ for cast steel.

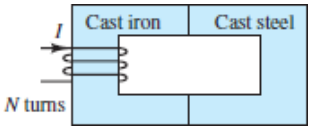


Figure P14.12

14.13 Find the magnetic flux ϕ established in the series magnetic circuit of Figure P14.13.

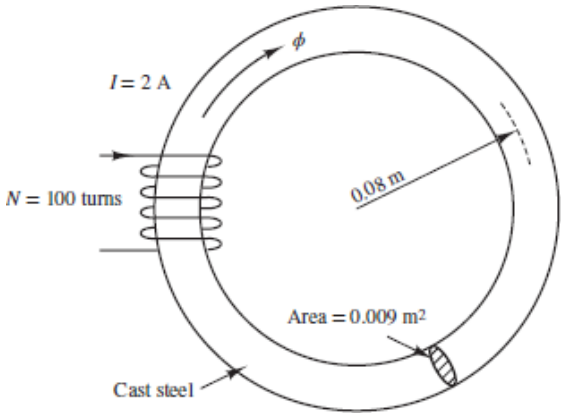


Figure P14.13

Page 858

14.14 a. Find the current I required to establish a flux $\phi = 2.4 \times 10^{-4}$ Wb in the magnetic circuit of Figure P14.14. Here area(throughout) $= 2 \times 10^{-4}$ m^2,

$l_{ab} = l_{ef} = 0.05$ m, $l_{af} = l_{be} = 0.02$ m, $l_{bc} = l_{dc}$, and the material is sheet steel.

b. Compare the mmf drop across the air gap to that across the rest of the magnetic circuit. Discuss your results, using the value of μ for each material.

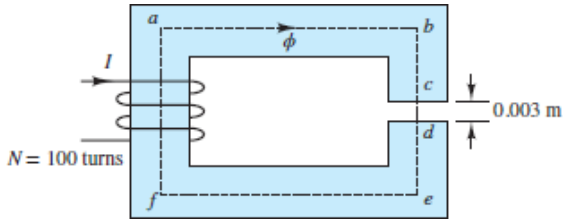


Figure P14.14

14.15 For the series-parallel magnetic circuit of Figure P14.15, find the value of I required to establish a flux in the gap of $\phi = 2 \times 10^{-4}$ Wb. Here, $l_{ab} = l_{bg} = l_{gh} = l_{ha} = 0.2$ m, $l_{bc} = l_{fg} = 0.1$ m, $l_{cd} = l_{ef} = 0.099$ m, and the material is sheet steel.

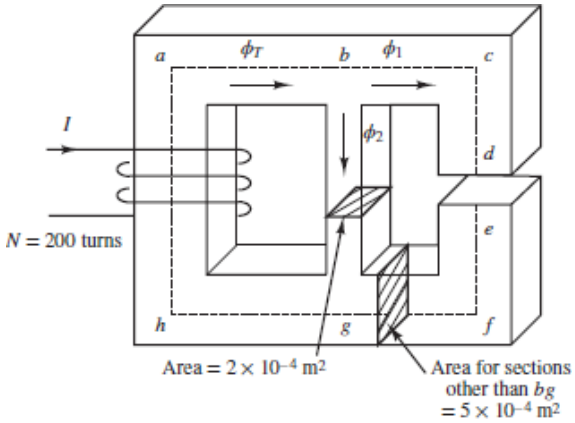


Figure P14.15

14.16 Refer to the actuator of Figure P14.16. The entire device is made of sheet steel. The coil has 2,000 turns. The armature is stationary so that the length of the air gaps, g = 10 mm, is fixed. A direct current passing through the coil produces a flux density of 1.2 T in the gaps. Assume μ_r = 4,000 for sheet steel. Determine:

a. The coil current.

b. The energy stored in the air gaps.

c. The energy stored in the steel.

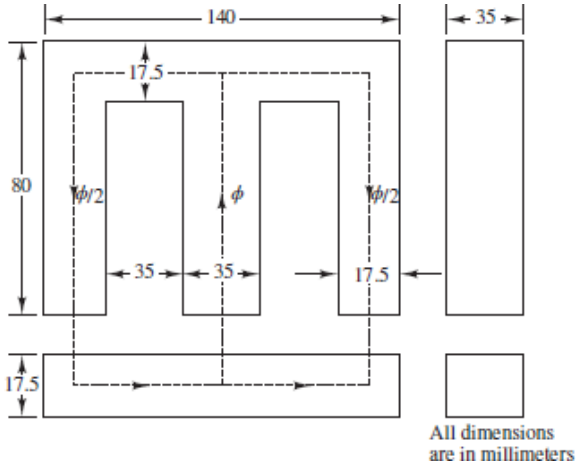


Figure P14.16

14.17 A core is shown in Figure P14.17, with $\mu_r = 2{,}000$ and $N = 100$. Find:

a. The current needed to produce a flux density of 0.4 Wb/m^2 in the center leg.

b. The current needed to produce a flux density of 0.8 Wb/m^2 in the center leg.

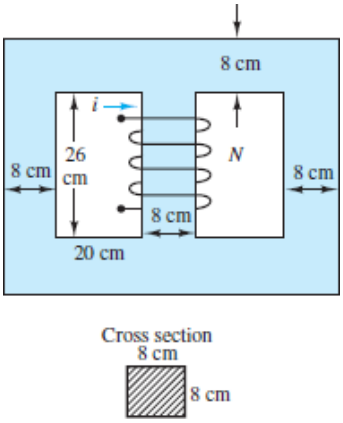


Figure P14.17

Page 859

Section 14.4: Transformers

14.18 For the transformer shown in Figure P14.18, $N = 1{,}000$ turns, $l_1 = 16$ cm, $A_1 = 4$ cm^2, $l_2 = 22$ cm, $A_2 = 4$ cm^2, $l_3 = 5$ cm, and $A_3 = 2$ cm^2. The relative permeability of the material is $\mu_r = 1{,}500$.

a. Construct the equivalent magnetic circuit, and find the reluctance associated with each part of the circuit.

b. Determine the self-inductance and mutual inductance for the pair of coils (that is, L_{11}, L_{22}, and $M = L_{12} = L_{21}$).

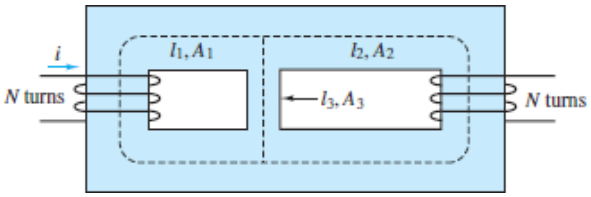


Figure P14.18

14.19 A transformer is delivering power to a 300-Ω resistive load. To achieve the desired power transfer, the turns ratio is chosen so that the resistive load referred to the primary is 7,500 Ω. The parameter values, *referred to the secondary winding*, are:

$$r_1 = 20\,\Omega \qquad L_1 = 1.0\text{ mH} \qquad L_m = 25\text{ mH}$$
$$r_2 = 20\,\Omega \qquad L_2 = 1.0\text{ mH}$$

Core losses are negligible.

a. Determine the turns ratio.

b. Determine the input voltage, current, and power and the efficiency when this transformer is delivering 12 W to the 300-Ω load at a frequency $f = 10{,}000/2\pi$ Hz.

14.20 A 220/20-V transformer has 50 turns on its low-voltage side. Calculate:

a. The number of turns on its high side.

b. The turns ratio α when it is used as a step-down transformer.

c. The turns ratio α when it is used as a step-up transformer.

14.21 The high-voltage side of a transformer has 750 turns, and the low-voltage side has 50 turns. When the high side is connected to a rated voltage of 120 V, 60 Hz, a rated load of 40 A is connected to the low side. Calculate:

a. The turns ratio.

b. The secondary voltage (assuming no internal transformer impedance voltage drops).

c. The resistance of the load.

14.22 A transformer is to be used to match an 8-Ω loudspeaker to a 500-Ω audio line. What is the turns ratio of the transformer, and what are the voltages at the primary and secondary terminals when 10 W of audio power is delivered to the speaker? Assume that the speaker is a resistive load and that the transformer is ideal.

14.23 The high-voltage side of a step-down transformer has 800 turns, and the low-voltage side has 100 turns. A voltage of 240 VAC is applied to the high side, and the load impedance is 3 Ω (low side). Find:

a. The secondary voltage and current.

b. The primary current.

c. The primary input impedance from the ratio of primary voltage to current.

d. The primary input impedance.

14.24 Calculate the transformer ratio of the transformer in Problem 14.23 when it is used as a step-up transformer.

14.25 A 2,300/240-V, 60-Hz, 4.6-kVA transformer is designed to have an induced emf of 2.5 V/turn. Assuming an ideal transformer, find:

a. The numbers of high-side turns N_h and low-side turns N_l.

b. The rated current of the high-voltage side I_h.

c. The transformer ratio when the device is used as a step-up transformer.

Section 14.5: Electromechanical Energy Conversion

14.26 Calculate the current required to lift the load for the electromagnet of Example 14.9. Calculate the holding current required to keep the load in place once it has been lifted and is attached to the magnet. Page 860Assume: N = 700; $\mu_0 = 4\pi \times 10^{-7}$; $\mu_r = 10^4$ (equal for electromagnet and load); initial distance (air gap) = 0.5 m; magnetic path length of electromagnet = l_1 = 0.80 m; magnetic path length of movable load = l_2 = 0.40 m; gap cross-sectional area = 5×10^{-4} m^2; m = mass of load = 10 kg; g = 9.8 m/s^2.

14.27 For the electromagnet of Example 14.9:

a. Calculate the current required to keep the bar in place. (*Hint:* The air gap becomes zero, and the iron reluctance cannot be neglected.) Assume μ_r =

1,000, $L = 1$ m.

b. If the bar is initially 0.1 m away from the electromagnet, what initial current would be required to lift the magnet?

14.28 The electromagnet of Figure P14.28 has reluctance given by $\mathcal{R}(x) = 7 \times 10^5(0.002 + x)\ \text{H}^{-1}$, where x is the length of the variable gap in meters. The coil has 980 turns and 30-Ω resistance. For an applied voltage of 120 VDC, find:

a. The energy stored in the magnetic field for $x = 0.005$ m.

b. The magnetic force for $x = 0.005$ m.

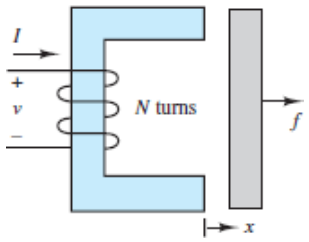


Figure P14.28

14.29 With reference to Example 14.10, determine the best combination of current magnitude and wire diameter to reduce the volume of the solenoid coil to a minimum. Will this minimum volume result in the lowest possible resistance? How does the power dissipation of the coil change with the wire gauge and current value? To solve this problem, you will need to find a table of wire gauge diameter, resistance, and current ratings. Table 1.1 in this book contains some information. The solution can only be found numerically.

14.30 Derive the same result obtained in Example 14.10, using equation 14.46 and the definition of inductance given in equation 14.30. You will first compute the inductance of the magnetic circuit as a function of the reluctance, then compute the stored magnetic energy, and finally write the expression for the magnetic force given in equation 14.46.

14.31 With reference to Example 14.11, calculate the required holding current to keep the relay closed. The mass of the moving element is $m = 0.05$ kg. Neglect damping. The initial position is $x = \epsilon = 0.001$ m.

14.32 The relay circuit shown in Figure P14.32 has the following parameters: $A_{\text{gap}} = 0.001\ \text{m}^2$; $N = 500$ turns; $L = 0.02$ m; $\mu = \mu_0 = 4\pi \times 10^{-7}$ (neglect the iron reluctance); $k = 1{,}000$ N/m; $R = 18\ \Omega$. What is the minimum DC supply

voltage v for which the relay will make contact when the electrical switch is closed?

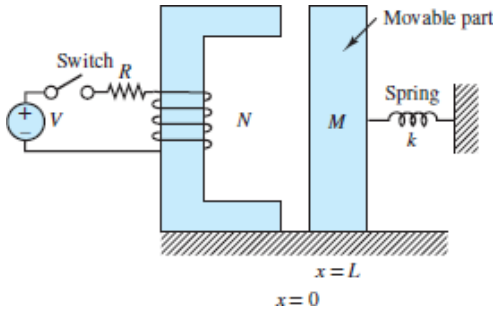


Figure P14.32

14.33 The magnetic circuit shown in Figure P14.33 is a very simplified representation of devices used as *surface roughness sensors*. The stylus is in contact with the surface and causes the plunger to move along with the surface. Assume that the flux ϕ in the gap is given by the expression $\phi = \beta/\mathcal{R}(x)$, , where β is a known constant and $\mathcal{R}(x)$ is the reluctance of the gap. The emf e is measured to determine the surface profile. Derive an expression for the displacement x as a function of the various parameters of the magnetic circuit and of the measured emf. (Assume a frictionless contact between the moving plunger and the magnetic structure and that the plunger is restrained to vertical motion only. The cross-sectional area of the plunger is A.)

Page 861

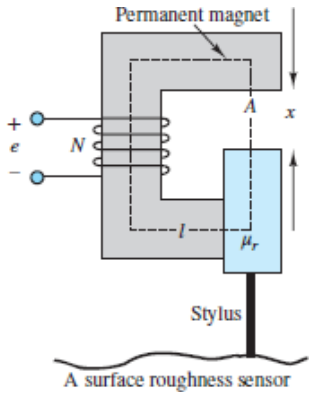


Figure P14.33

14.34 A cylindrical solenoid is shown in Figure P14.34. The plunger may move freely along its axis. The air gap between the shell and the plunger is uniform and equal to 1 mm, and the diameter d is 25 mm. If the exciting coil carries a

current of 7.5 A, find the force acting on the plunger when $x = 2$ mm. Assume $N = 200$ turns, and neglect the reluctance of the steel shell. Assume l_g is negligible.

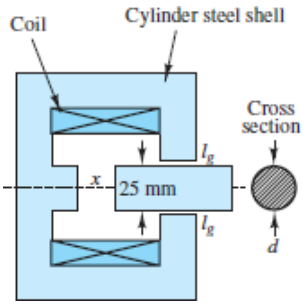


Figure P14.34

14.35 The double-excited electromechanical system shown in Figure P14.35 moves horizontally. Assume that resistance, magnetic leakage, and fringing are negligible; the permeability of the core is very large; and the cross section of the structure is $w \times w$. Find:

a. The reluctance of the magnetic circuit.

b. The magnetic energy stored in the air gap.

c. The force on the movable part as a function of its position.

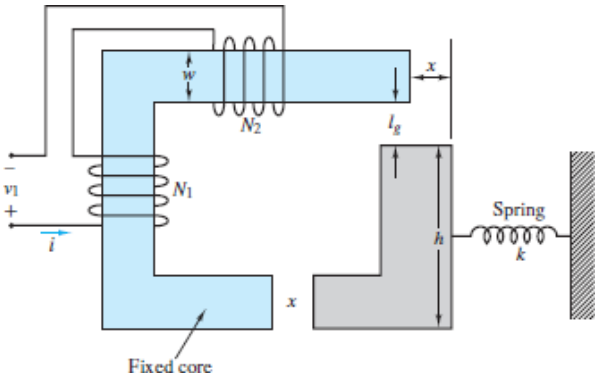


Figure P14.35

14.36 Determine the force F between the faces of the poles (stationary coil and plunger) of the solenoid pictured in Figure P14.36 when it is energized. When energized, the plunger is drawn into the coil and comes to rest with only a negligible air gap separating the two. The flux density in the cast steel pathway is 1.1 T. The diameter of the plunger is 10 mm. Assume that the reluctance of the steel is negligible.

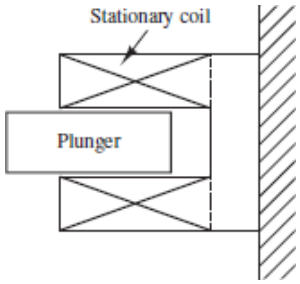


Figure P14.36

14.37 An electromagnet is used to support a solid piece of steel, as shown in Example 14.9. A force of 10,000 N is required to support the weight. The cross-sectional area of the magnetic core (the fixed part) is 0.01 m^2. The coil has 1,000 turns. Determine the minimum current that can keep the weight from falling for x = 1.0 mm. Assume negligible reluctance in steel and negligible fringing in the air gaps.

14.38 The armature, frame, and core of a 12-VDC control relay are made of sheet steel. The average length of the magnetic circuit is 12 cm when the relay is energized, and the average cross section of the magnetic circuit is 0.60 cm^2. The coil is wound with 250 turns and carries 50 mA. Determine:

a. The flux density $\mathcal{B}$ in the magnetic circuit of the relay when the coil is energized.

b. The force $\mathcal{F}$ exerted on the armature to close it when the coil is energized.

Page 862

14.39 A relay is shown in Figure P14.39. Find the differential equations describing the system.

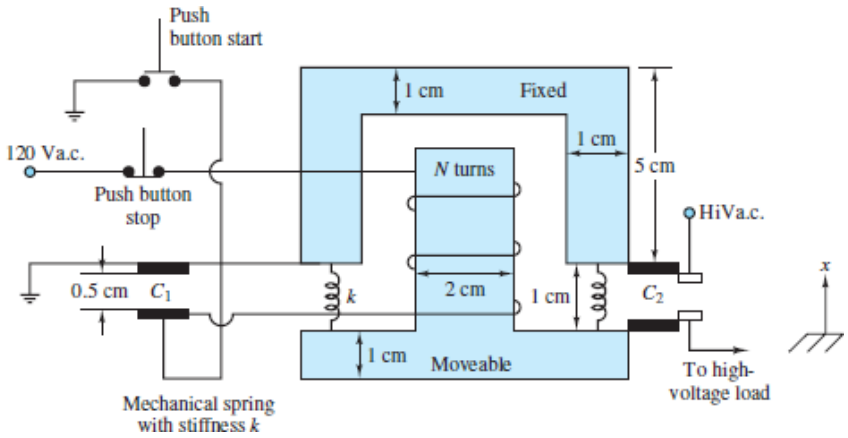


Figure P14.39

14.40 A solenoid having a cross section of 10 cm^2 is shown in Figure P14.40.

a. Calculate the force exerted on the plunger when the distance x is 2 cm and the current in the coil (where N = 100 turns) is 5 A. Assume that the fringing and leakage effects are negligible. The relative permeabilities of the magnetic material and the nonmagnetic sleeve are 2,000 and 1.

b. Develop a set of differential equations governing the behavior of the solenoid.

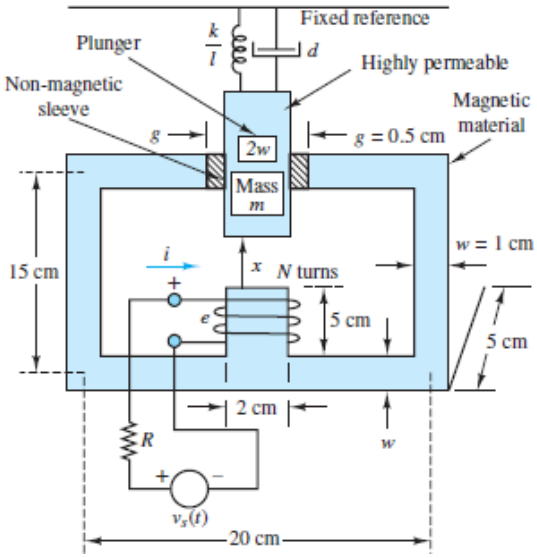


Figure P14.40

14.41 Derive the differential equations (electrical and mechanical) for the relay shown in Figure P14.41. Do not assume that the inductance is fixed; it is a function of x. You may assume that the iron reluctance is negligible.

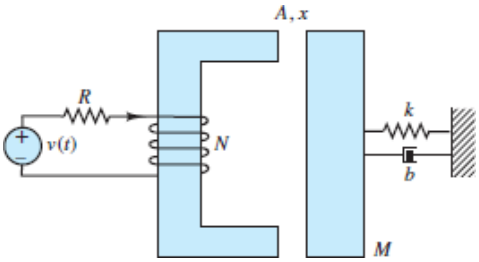


Figure P14.41

14.42 Derive the complete set of differential equations describing the relay shown in Figure P14.42.

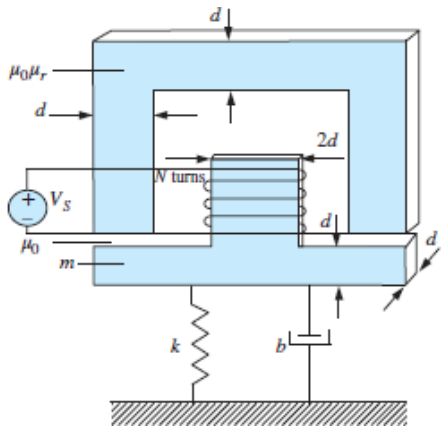


Figure P14.42

Page 863

14.43 A wire of length 20 cm vibrates in one direction in a constant magnetic field with a flux density of 0.1 T; see Figure P14.43. The position of the wire as a function of time is given by $x(t) = 0.1 \sin 10t$ m. Find the induced emf across the length of the wire as a function of time.

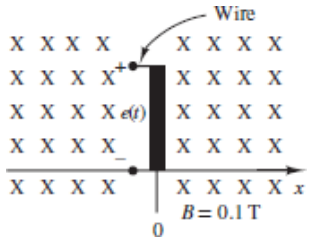


Figure P14.43

14.44 The wire of Problem 14.43 induces a time-varying emf of

$$e_1(t) = 0.02 \cos 10t$$

A second wire is placed in the same magnetic field but has a length of 0.1 m, as shown in Figure P14.44. The position of this wire is given by $x(t) = 1 - 0.1 \sin 10t$. Find the induced emf $e(t)$ defined by the difference between $e_1(t)$ and $e_2(t)$.

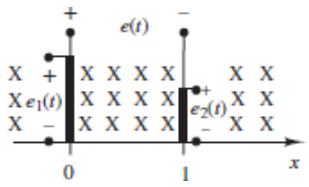


Figure P14.44

14.45 A conducting bar shown in Figure 14.47 in the text is carrying 4 A of current in the presence of a magnetic field $B = 0.3$ Wb/m^2. Find the magnitude and direction of the force induced on the bar.

14.46 A wire, shown in Figure P14.46, is moving in the presence of a magnetic field $B = 0.4$ Wb/m^2. Find the magnitude and direction of the induced voltage in the wire.

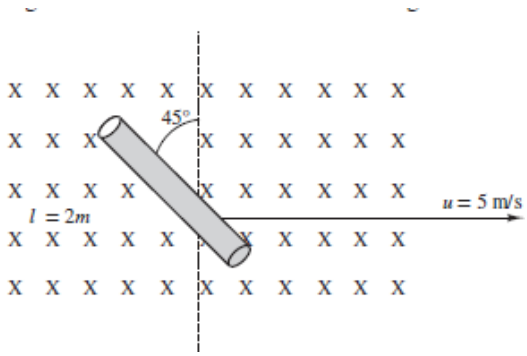


Figure P14.46

14.47 The electrodynamic shaker shown in Figure P14.47 is commonly used as a vibration tester. A constant current is used to generate a magnetic field in which the armature coil of length l is immersed. The shaker platform with mass m is mounted in the fixed structure by way of a spring with stiffness k. The platform is rigidly attached to the armature coil, which slides on the fixed structure thanks to frictionless bearings.

a. Neglecting iron reluctance, determine the reluctance of the fixed structure, and hence compute the strength of the magnetic flux density B in which the armature coil is immersed.

b. Knowing B, determine the dynamic equations of motion of the shaker, assuming that the moving coil has resistance R and inductance L.

c. Derive the transfer function and frequency response function of the shaker mass *velocity* in response to the input voltage V_S.

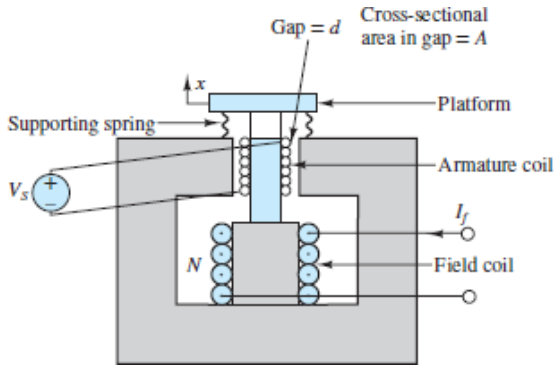


Figure P14.47

14.48 The electrodynamic shaker of Figure P14.47 is used to perform vibration testing of an electrical connector. The connector is placed on the test platform (with mass m), and it may be assumed to have negligible mass when compared to the platform. The test consists of shaking the connector at the frequency $\omega = 2\pi \times 100$ rad/s.

Given the parameter values $B = 1{,}000$ Wb/m^2, $l = 5$ m, $k = 1{,}000$ N/m, $m = 1$ kg, $b = 5$ N-s/m, $L = 0.8$ H, and $R = 0.5\ \Omega$, determine the peak amplitude of the sinusoidal voltage V_S required to generate an acceleration of $5g$ (49 m/s^2).

Page 864

14.49 Derive and sketch the frequency response of the loudspeaker of Example 14.12 for (1) $k = 50{,}000$ N/m and (2) $k = 5 \times 10^6$ N/m. Describe qualitatively how the loudspeaker frequency response changes as the spring stiffness k increases and decreases. What will the frequency response be in the limit as k approaches zero? What kind of speaker would this condition correspond to?

14.50 The loudspeaker of Example 14.12 has a midrange frequency response. Modify the mechanical parameters of the loudspeaker (mass, damping, and spring rate) so as to obtain a loudspeaker with a bass response centered on 400 Hz. Demonstrate that your design accomplishes the intended task, using frequency response plots. *Note: This is an open-ended design problem.*

14.51 The electrodynamic shaker shown in Figure P14.51 is used to perform vibration testing of an electronic circuit. The circuit is placed on a test table with mass m, and is assumed to have negligible mass when compared to the table. The test consists of shaking the circuit at the frequency $\omega = 2\pi(100)$ rad/s.

a. Write the dynamic equations for the shaker. Clearly indicate system input(s) and output(s).

b. Find the frequency response function of the table acceleration in response to the applied voltage.

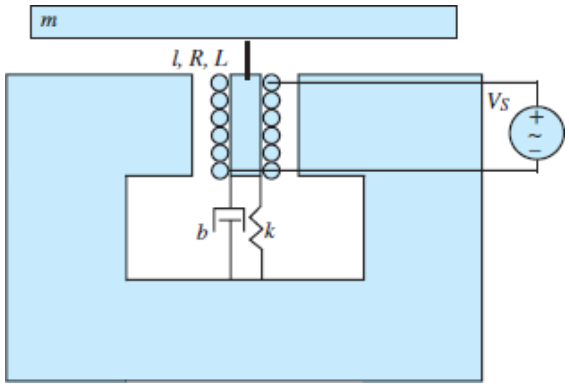


Figure P14.51

c. Given the following parameter values:

$B = 200\ \text{Wb/m}^2 \quad l = 5\ \text{m} \quad k = 100\ \text{N/m}$
$m = 0.2\ \text{kg} \quad b = 5\ \text{N-s/m}$
$L = 8\ \text{mH} \quad R = 0.5\ \Omega$

Determine the peak amplitude of the sinusoidal voltage V_S required to generate an acceleration of $5g$ (49 m/s^2) under the stated test conditions.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]We will use the boldface symbols **B** and **H** to denote the vector forms of B and H; the standard typeface will represent the scalar flux density or field intensity in a given direction.

[2]Note that although they are dimensionally equal to amperes, the units of magnetomotive force are ampere-turns.

[3]See the Focus on Measurements box, "Charge Amplifiers," in Chapter 6.

C H A P T E R

15

ELECTRIC MACHINES

The objective of this chapter is to introduce the basic operation of rotating electric machines. The operation of the three major classes of electric machines—DC, synchronous, and induction—is described as intuitively as possible, building on the material presented in Chapter 14.

The emphasis of this chapter is on explaining the properties of each type of machine, with its advantages and disadvantages with regard to other types; and on classifying these machines in terms of their performance characteristics and preferred field of application.

Learning Objectives

Students will learn to...

1. Understand the basic principles of operation of rotating electric machines, their classification, and basic efficiency and performance characteristics *Section 15.1*.

2. Understand the operation and basic configurations of separately excited, permanent-magnet, shunt and series DC machines. *Section 15.2*.
3. Analyze DC motors under steady-state and dynamic operation. *Section 15.3*.
4. Analyze DC generators at steady state. *Section 15.4*.
5. Understand the operation and basic configuration of AC machines, including the synchronous motor and generator, and the induction machine. *Sections 15.6 to 15.9*.

15.1 ROTATING ELECTRIC MACHINES

This introductory section is aimed at explaining the common properties of all rotating electric machines. We begin our discussion with reference to Figure 15.1, in which a hypothetical rotating machine is depicted in a cross-sectional view. In the figure, a box with a cross inscribed in it indicates current flowing into the page, while a dot represents current out of the plane of the page.

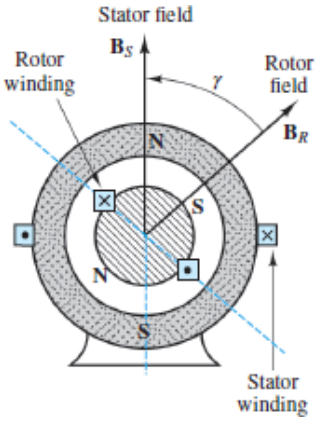


Figure 15.1 A rotating electric machine

In Figure 15.1, we identify a **stator**, of cylindrical shape, and a **rotor**, which, as the name indicates, rotates inside the stator, separated from the latter by means of an air gap. The rotor and stator each consist of a magnetic core, some electrical insulation, and the windings necessary to establish a magnetic flux (unless this is created by a permanent magnet). The rotor is mounted on a bearing-supported shaft, which can be connected to *mechanical loads* (if the machine is a motor) or to a *prime mover* (if the machine is a generator) by means of belts, pulleys, chains, or other mechanical couplings. The windings carry the electric currents that generate the magnetic fields and flow to the electrical loads, and also provide

the closed loops in which voltages will be induced (by virtue of Faraday's law, as discussed in Chapter 14).

Basic Classification of Electric Machines

An immediate distinction can be made between different types of windings characterized by the nature of the current they carry. If the current serves the sole purpose of providing a magnetic field and is independent of the load, it is called a *magnetizing*, or excitation, current, and the winding is termed a **field winding**. Field currents are nearly always direct current (DC) and are of relatively low power, since their only purpose is to magnetize the core (recall the important role of high-permeability cores in generating large magnetic fluxes from relatively small currents). On the other hand, if the winding carries only the load current, it is called an **armature**. In DC and alternating-current (AC) synchronous machines, separate windings exist to carry field and armature currents. In the induction motor, the magnetizing and load currents flow in the same winding, called the *input winding*, or *primary;* the output winding is then called the *secondary*. As we shall see, this terminology, which is reminiscent of transformers, is particularly appropriate for induction motors, which bear a significant analogy to the operation of the transformers studied in Chapters 13 and 14. Table 15.1 characterizes the principal machines in terms of their field and armature configuration.

Page 867

Table 15.1 Configurations of the three types of electric machines

Machine type	Winding	Winding type	Location	Current
DC	Input and output	Armature	Rotor	AC (winding) DC (at brushes)
	Magnetizing	Field	Stator	DC
Synchronous	Input and output	Armature	Stator	AC
	Magnetizing	Field	Rotor	DC
Induction	Input	Primary	Stator	AC
	Output	Secondary	Rotor	AC

It is also useful to classify electric machines in terms of their energy conversion characteristics. A machine acts as a **generator** if it converts mechanical energy from a prime mover, say, an internal combustion engine, to

electric energy. Examples of generators are the large machines used in power generating plants, or the common automotive alternator. A machine is classified as a **motor** if it converts electric energy to mechanical form. The latter class of machines is probably of more direct interest to you, because of its widespread application in engineering design. Electric motors are used to provide forces and torques to generate motion in countless industrial applications. Machine tools, robots, punches, presses, mills, and propulsion systems for electric vehicles are but a few examples of the application of electric machines in engineering.

Note that in Figure 15.1 we have explicitly shown the direction of two magnetic fields: that of the rotor $\mathbf{B}_R$ and that of the stator $\mathbf{B}_S$. Although these fields are generated by different means in different machines (e.g., permanent magnets, alternating currents, direct currents), the presence of these fields is what causes a rotating machine to turn and enables the generation of electric power. In particular, we see that in Figure 15.1 the north pole of the rotor field will seek to align itself with the south pole of the stator field. It is this magnetic attraction force that permits the generation of torque in an electric motor; conversely, a generator exploits the laws of electromagnetic induction to convert a changing magnetic field to an electric current.

To simplify the discussion in later sections, we now introduce some basic concepts that apply to all rotating electric machines. Referring to Figure 15.2, which depicts a permanent-magnet DC machine, note that the force on a wire is given by the expression:

$$\mathbf{f} = i_w \mathbf{l} \times \mathbf{B} \tag{15.1}$$

where i_w is the current in the wire, $\mathbf{l}$ is a vector along the direction of the wire, and $\times$ denotes the cross product of two vectors. Then the torque for a multiturn coil is:

$$T = KBi_w \sin\alpha \tag{15.2}$$

where:

B = magnetic flux density caused by stator field

K = constant depending on coil geometry

α = angle between $\mathbf{B}$ and normal to plane of coil

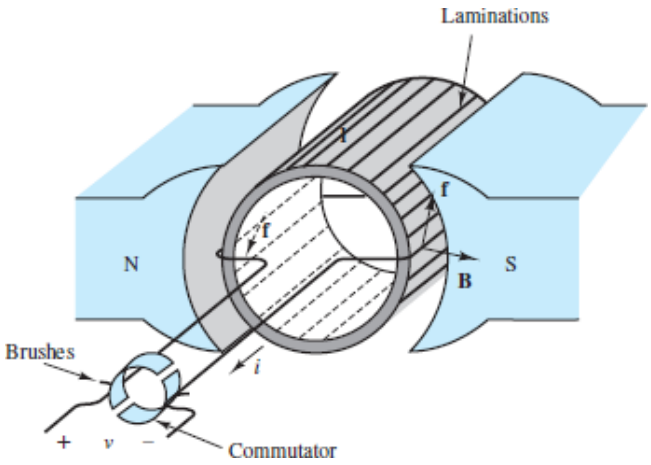


Figure 15.2 Stator and rotor fields and the force acting on a rotating permanent-magnet DC machine

In the machine of Figure 15.2, there are two magnetic fields: one generated within the stator, the other within the rotor windings. Either (but not both) of these fields Page 868could be generated by a current or by a permanent magnet. Thus, we could replace the permanent-magnet stator of Figure 15.2 with a suitably arranged winding to generate a stator field in the same direction. If the stator were made of a toroidal coil of radius R (see Chapter 14), then the magnetic field of the stator would generate a flux density B, where:

$$B = \mu H = \mu \frac{Ni}{2\pi R} \tag{15.3}$$

and where N is the number of turns and i is the coil current. The direction of the torque is always the direction determined by the rotor and stator fields as they seek to align to each other (i.e., counterclockwise in the diagram of Figure 15.1).

It is important to note that Figure 15.2 is only one example of the major features and characteristics of rotating machines. A variety of configurations exist, depending on whether each of the fields is generated by a current in a coil or by a permanent magnet and whether the load and magnetizing currents are direct or alternating. The type of excitation (AC or DC) provided to the windings permits a first classification of electric machines (see Table 15.1). According to this classification, one can define the following types of machines:

- *DC machines:* Direct current in both stator and rotor (the stator could also be realized by a permanent magnet, as in Figure 15.2)
- *Synchronous machines:* Alternating current in one stator, direct current in the rotor (the rotor could alternatively consist of a permanent magnet)
- *Induction machines:* Alternating current in both

In most industrial applications, the induction machine is the preferred choice, because of the simplicity of its construction. However, the analysis of the performance of an induction machine is rather complex. On the other hand, DC machines are quite complex in their construction but can be analyzed relatively simply with the analytical tools we have already acquired. Therefore, the progression of this Page 869chapter is as follows: We start with a section that discusses the physical construction of DC machines, both motors and generators. Then we continue with a discussion of synchronous machines, in which one of the currents is now alternating, since these can easily be understood as an extension of DC machines. Finally, we consider the case where both rotor and stator currents are alternating, and we analyze the induction machine.

Performance Characteristics of Electric Machines

As already stated earlier in this chapter, electric machines are **energy conversion devices**, and we are therefore interested in their energy conversion **efficiency**. Typical applications of electric machines as motors or generators must take into consideration the energy losses associated with these devices. Figure 15.3(a) and (b) represents the various loss mechanisms you must consider in analyzing the efficiency of an electric machine for the case of DC machines. It is important for you to keep in mind this conceptual flow of energy when analyzing electric machines. The sources of loss in a rotating machine can be separated into three fundamental groups: electrical (I^2R) losses, core losses, and mechanical losses.

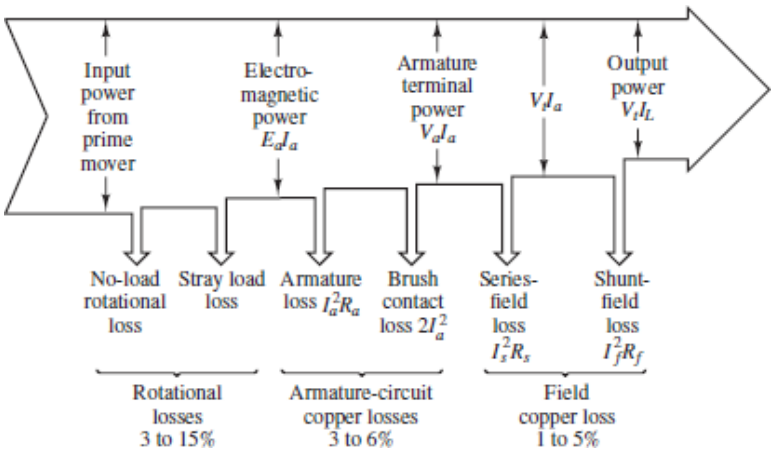


Figure 15.3a Generator losses, direct current

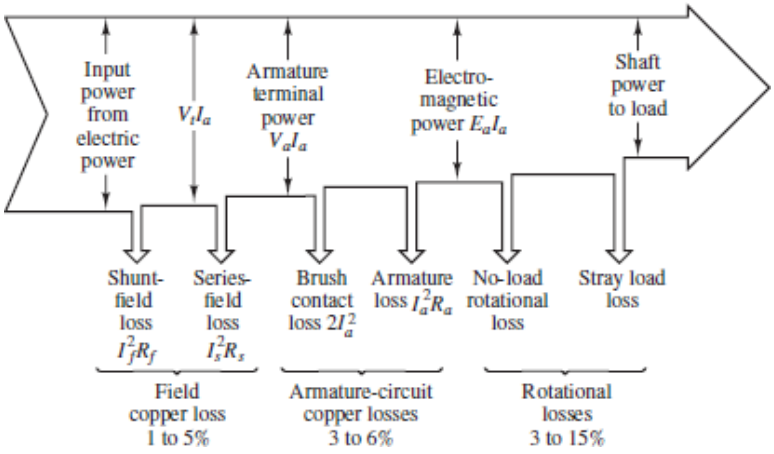


Figure 15.3b Motor losses, direct current

Page 870

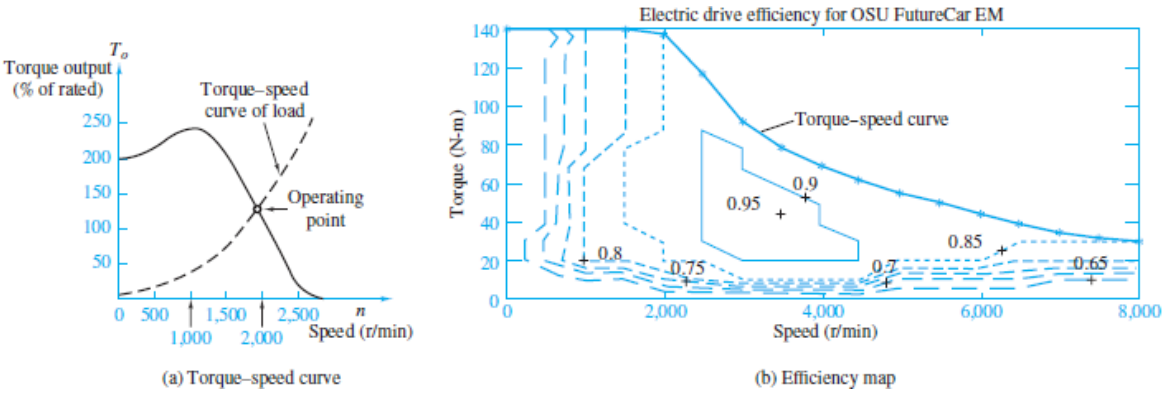
Usually I^2R losses are computed on the basis of the DC resistance of the windings at 75°C; in practice, these losses vary with operating conditions. The difference between the nominal and actual I^2R loss is usually lumped under the category of *stray-load loss.* In DC machines, it is also necessary to account for the *brush contact loss* associated with slip rings and commutators.

Mechanical losses are due to *friction* (mostly in the bearings) and *windage*, that is, the air drag force that opposes the motion of the rotor. In addition, if external devices (e.g., blowers) are required to circulate air through the machine

for cooling purposes, the energy expended by these devices is included in the mechanical losses.

Open-circuit core losses consist of *hysteresis* and *eddy current* losses, with only the excitation winding energized (see Chapter 14 for a discussion of hysteresis and eddy currents). Often these losses are summed with friction and windage losses to give rise to the *no-load rotational loss*. The latter quantity is useful if one simply wishes to compute efficiency. Since open-circuit core losses do not account for the changes in flux density caused by the presence of load currents, an additional magnetic loss is incurred that is not accounted for in this term. *Stray-load losses* are used to lump the effects of nonideal current distribution in the windings and of the additional core losses just mentioned. Stray-load losses are difficult to determine exactly and are often assumed to be equal to 1.0 percent of the output power for DC machines; these losses can be determined by experiment in synchronous and induction machines.

The performance of an electric machine can be quantified in a number of ways. In the case of an electric motor, it is usually portrayed in the form of a graphical **torque–speed characteristic** and **efficiency map**. The torque–speed characteristic of a motor describes how the torque supplied by the machine varies as a function of the speed of rotation of the motor for steady speeds. As we shall see in later sections, the torque–speed curves vary in shape with the type of motor (DC, induction, synchronous) and are very useful in determining the performance of the motor when connected to a mechanical load. Figure 15.4(a) depicts the torque–speed curve of induction motor. Figure 15.4(b) depicts a typical efficiency map for a permanent-magnet synchronous motor. In most engineering Page 871applications, it is quite likely that the engineer is required to make a decision regarding the performance characteristics of the motor best suited to a specified task. In this context, the torque–speed curve of a machine is a very useful piece of information.



(a) Torque–speed curve

(b) Efficiency map

Figure 15.4 Torque–speed and efficiency curves for an electric motor: (a) an induction machine; (b) an electric drive system for a hybrid-electric vehicle

The first feature we note of the torque–speed characteristic is that it bears a strong resemblance to the *i*-*υ* characteristics used in earlier chapters to represent the behavior of electrical sources. It should be clear that, according to this torque–speed curve, the motor is not an ideal source of torque (if it were, the curve would appear as a horizontal line across the speed range). One can readily see, for example, that the induction motor represented by the curves of Figure 15.4(a) would produce maximum torque in the range of speeds between approximately 800 and 1,400 r/min. What determines the actual speed of the motor (and therefore its output torque and power) is the torque–speed characteristic of the load connected to it, much as a resistive load determines the current drawn from a voltage source. In the figure, we display the torque–speed curve of a load, represented by the dashed line; the operating point of the motor-load pair is determined by the intersection of the two curves.

Another important observation pertains to the fact that the motor of Figure 15.4(a) produces a nonzero torque at zero speed. This fact implies that as soon as electric power is connected to the motor, the latter is capable of supplying a certain amount of torque; this zero-speed torque is called the **starting torque**. If the load requires less than the starting torque the motor can provide, then the motor can accelerate the load until the motor speed and torque settle to a stable value, at the operating point. As we discuss each type of machine in greater detail, we shall devote some time to the discussion of its torque–speed curve.

The efficiency of an electric machine is also an important design and performance characteristic. The 2005 Department of Energy's Energy Policy Act, also known as EPACT, has required electric motor manufacturers to guarantee a minimum efficiency. The efficiency of an electric motor is usually described using a contour plot of the efficiency value (a number between 0 and 1) in the torque–speed plane. This representation permits a determination of the motor efficiency as a function of its performance and operating conditions. Figure 15.4(b) depicts the efficiency map of an electric drive used in a hybrid-electric vehicle—a 20-kW permanent-magnet AC synchronous machine. We discuss this type of machine in Chapter 16. Note that the peak efficiency can be as high as 0.95 (95 percent), but that the efficiency decreases significantly away from the optimum point (around 3,500 r/min and 45 N-m), to values as low as 0.65.

The most common means of conveying information regarding electric machines is the *nameplate.* Typical information conveyed by the nameplate

includes

1. Type of device (e.g., DC motor, alternator)
2. Manufacturer
3. Rated voltage and frequency
4. Rated current and voltamperes
5. Rated speed and horsepower

The **rated voltage** is the terminal voltage for which the machine was designed, and which will provide the desired magnetic flux. Operation at higher voltages Page 872will increase magnetic core losses, because of excessive core saturation. The **rated current** and **rated voltamperes** are an indication of the typical current and power levels at the terminal that will not cause undue overheating due to copper losses (I^2R losses) in the windings. These ratings are not absolutely precise, but they give an indication of the range of excitations for which the motor will perform without overheating. Other name plate characteristics are introduced in Example 15.2.

Peak power operation in a motor may exceed rated torque, power, or currents by a substantial factor (up to as much as 6 or 7 times the rated value); however, continuous operation of the motor above the rated performance will cause the machine to overheat and eventually to sustain damage. Thus, it is important to consider both peak and continuous power requirements when selecting a motor for a specific application. An analogous discussion is valid for the speed rating: While an electric machine may operate above rated speed for limited periods of time, the large centrifugal forces generated at high rotational speeds will eventually cause undesirable mechanical stresses, especially in the rotor windings.

Another important feature of electric machines is the **regulation** of the machine speed or voltage, depending on whether it is used as a motor or as a generator, respectively. Regulation is the ability to maintain speed or voltage constant in the face of load variations. The ability to closely regulate speed in a motor or voltage in a generator is an important feature of electric machines; regulation is often improved by means of feedback control mechanisms, some of which are briefly introduced in this chapter. We take the following definitions as being adequate for the intended purpose of this chapter:

$$\text{Speed regulation} = \text{SR} = \frac{\text{Speed at no load} - \text{Speed at rated load}}{\text{Speed at rated load}} \tag{15.4}$$

$$\text{Voltage regulation} = \text{VR} = \frac{\text{Voltage at no load} - \text{Voltage at rated load}}{\text{Voltage at rated load}} \tag{15.5}$$

Please note that the rated value is usually taken to be the nameplate value, and that the meaning of *load* changes depending on whether the machine is a motor, in which case the load is mechanical, or a generator, in which case the load is electrical.

EXAMPLE 15.1 Regulation

Problem

Find the percentage of speed regulation of a shunt DC motor.

Solution

Known Quantities: No-load speed; speed at rated load.

Find: Percentage speed regulation, denoted by SR%.

Page 873

Schematics, Diagrams, Circuits, and Given Data:

n_{nl} = no-load speed = 1,800 r/min

n_{rl} = rated load speed = 1,760 r/min

Analysis:

$$\text{SR\%} = \frac{n_{nl} - n_{rl}}{n_{rl}} \times 100 = \frac{1{,}800 - 1{,}760}{1{,}760} \times 100 = 2.27\%$$

Comments: Speed regulation is an intrinsic property of a motor; however, external speed controls can be used to regulate the speed of a motor to any (physically achievable) desired value. Some motor control concepts are discussed later in this chapter.

CHECK YOUR UNDERSTANDING

The percentage of speed regulation of a motor is 10 percent. If the full-load speed is 50π rad/s, find (a) the no-load speed in radians per second and (b) the no-load speed in revolutions per minute. (c) If the percentage of voltage regulation for a 250-V generator is 10 percent, find the no-load voltage of the generator.

Answer: (a) $\omega = 55\pi$ rad/s; (b) $n = 1{,}650$ r/min; (c) $V_{\text{no-load}} = 275$ V

Table 15.2 summarizes important unit conversions that relate SI to English units, as the latter are still used in nameplate data in the United States.

Table 15.2 Unit conversions for electric machines

Quantity	SI unit	English unit
Length	1 m	3.281 ft
Mass	1 kg	2.205 lb (mass)
Force	1 N	0.224 lb (force)
Torque	1 N-m	0.738 lb-ft
		8.85 lb-in
Power	1 kW	1.341 hp
Moment of inertia	1 kg-m^2	23.73 lb-ft^2

EXAMPLE 15.2 Nameplate Data

Problem

Discuss the nameplate data, shown below, of a typical induction motor.

Solution

Known Quantities: Nameplate data.

Find: Motor characteristics.

Schematics, Diagrams, Circuits, and Given Data: The nameplate appears below.

MODEL	19308 J-X		
TYPE	CJ4B	FRAME	324TS
VOLTS	230/460	°C AMB.	40
		INS. CL.	B
FRT. BRG	210SF	EXT. BRG	312SF
SERV FACT	1.0	OPER INSTR	C-517
PHASE | 3	Hz | 60	CODE | G	WDGS | 1
H.P.	40		
R.P.M.	3,565		
AMPS	106/53		
NEMA NOM.	EFF		
NOM. P.F.			
DUTY	CONT.	NEMA DESIGN	B

Analysis: The nameplate of a typical induction motor is shown in the preceding table. The model number (sometimes abbreviated as MOD) uniquely identifies the motor to the manufacturer. It may be a style number, a model number, an identification number, or an instruction sheet reference number.

The term *frame* (sometimes abbreviated as FR) refers principally to the physical size of the machine, as well as to certain construction features.

Ambient temperature (abbreviated as AMB, or MAX. AMB) refers to the maximum ambient temperature in which the motor is capable of operating. Operation of the motor in a higher ambient temperature may result in shortened motor life and reduced torque.

Insulation class (abbreviated as INS. CL.) refers to the type of insulation used in the motor. The classes most often used are class A (105°C) and class B (130°C).

The duty (DUTY), or time rating, denotes the length of time the motor is expected to be able to carry the rated load under usual service conditions. “CONT.” means that the machine can be operated continuously.

The “CODE” letter sets the limits of starting kilovoltamperes per horsepower for the machine. There are 19 levels, denoted by the letters A through V, excluding I, O, and Q.

Service factor (abbreviated as SERV FACT) is a term defined by the National Electrical Manufacturers Association (NEMA) as follows: “The service factor of a general-purpose alternating-current motor is a multiplier which, when applied to the rated horsepower, indicates a permissible horsepower loading which may be carried under the conditions specified for the service factor.”

The voltage figure given on the nameplate refers to the voltage of the supply circuit to which the motor should be connected. Sometimes two voltages are given, for example, 230/460. In this case, the machine is intended for use on either a 230-V or a 460-V circuit. Special instructions will be provided for connecting the motor for each of the voltages.

The term “BRG” indicates the nature of the bearings supporting the motor shaft.

Page 875

CHECK YOUR UNDERSTANDING

The nameplate of a three-phase induction motor indicates the following values:

H.P. = 10	Volt = 220 V
R.P.M. = 1,750	Service factor = 1.15
Temperature rise = 60°C	Amp = 30 A

Find the rated torque, rated voltamperes, and maximum continuous output power.

Answer: $I_{rated} = 40.7$ N-m; rated VA = 11,431 VA; $P_{max} = 11.5$ hp.

EXAMPLE 15.3 Torque–Speed Curves

Problem

Discuss the significance of the torque–speed curve of an electric motor.

Solution

An induction motor has a torque output that varies directly with speed; hence, the power output varies directly with the speed. Motors with this characteristic are commonly used with fans, blowers, and centrifugal pumps. Figure 15.5 shows typical torque–speed curves for this type of motor. Superimposed on the motor torque–speed curve is the torque–speed curve for a typical fan where the input power to the fan varies as the cube of the fan speed. Point *A* is the actual operating point, which could be determined graphically by plotting the load line and the motor torque–speed curve on the same graph, as illustrated in Figure 15.5. The fan will operate at the speed corresponding to the intersection of the two curves.

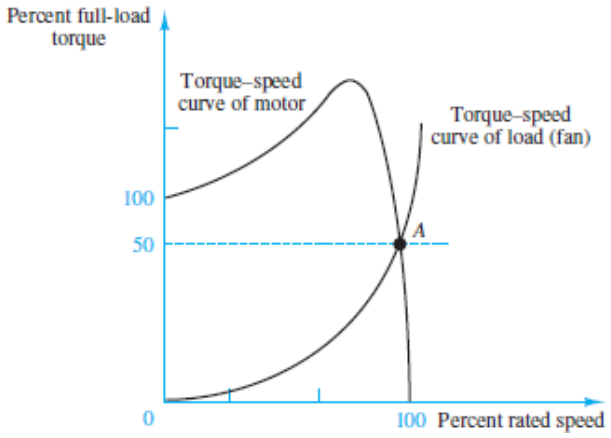


Figure 15.5 Torque–speed curves of electric motor and load

Page 876

CHECK YOUR UNDERSTANDING

A motor having the characteristics shown in Figure 15.4(a) is to drive a load; the load has a linear torque–speed curve and requires 150 percent of rated torque at 1,500 r/min. Find the operating point for this motor-load pair.

Answer: 170 percent of rated torque; 1,700 r/min.

Basic Operation of All Rotating Machines

We have already seen in Chapter 14 how the magnetic field in electromechanical devices provides a form of coupling between electrical and mechanical systems. Intuitively, one can identify two aspects of this coupling, both of which play a role in the operation of electric machines:

1. Magnetic attraction and repulsion forces generate mechanical torque.
2. The magnetic field can induce a voltage in the machine windings (coils) by virtue of Faraday's law.

Thus, an electric machine can serve either as a motor or a generator, depending on whether the input power is electric and mechanical power is produced (motor action), or the input power is mechanical and the output power is electric (generator action). Figure 15.6 illustrates the two cases graphically.

Figure 15.6 Generator (a) and motor (b) action in an electric machine

The coupling magnetic field performs a dual role, which may be explained as follows. When a current i flows through conductors placed in a magnetic field, a force is produced on each conductor, according to equation 15.1. If these conductors are attached to a cylindrical structure, a torque is generated; and if the structure is free to rotate, then it will rotate at an angular velocity ω_m. As the conductors rotate, however, they move through a magnetic field and cut through flux lines, thus generating an electromotive force in opposition to the excitation. This emf is also called *counter-emf*, as it opposes the source of the current i. If, on the other hand, the rotating element of the machine is driven by a prime mover (e.g., an internal combustion engine), then an emf is generated across the coil that is rotating in the magnetic field (the armature). If a load is connected to the

armature, a current i will flow to the load, and this current flow will in turn cause a reaction torque on the armature that opposes the torque imposed by the prime mover.

Page 877

You see, then, that for energy conversion to take place, two elements are required:

1. A coupling field **B**; generated in the field winding or by a permanent magnet.
2. An armature winding that supports the load current i and the emf e.

Magnetic Poles in Electric Machines

Before discussing the actual construction of a rotating machine, we should spend a few paragraphs to illustrate the significance of **magnetic poles** in an electric machine. In an electric machine, torque is developed as a consequence of magnetic forces of attraction and repulsion between magnetic poles on the stator and on the rotor; these poles produce a torque that accelerates the rotor and a reaction torque on the stator. It is also important to observe that the number of poles must be even, since there have to be equal numbers of north and south poles.

The motion and associated electromagnetic torque of an electric machine are the result of two magnetic fields that are trying to align with each other so that the south pole of one field attracts the north pole of the other. Figure 15.7 illustrates this action by analogy with two permanent magnets, one of which is allowed to rotate about its center of mass.

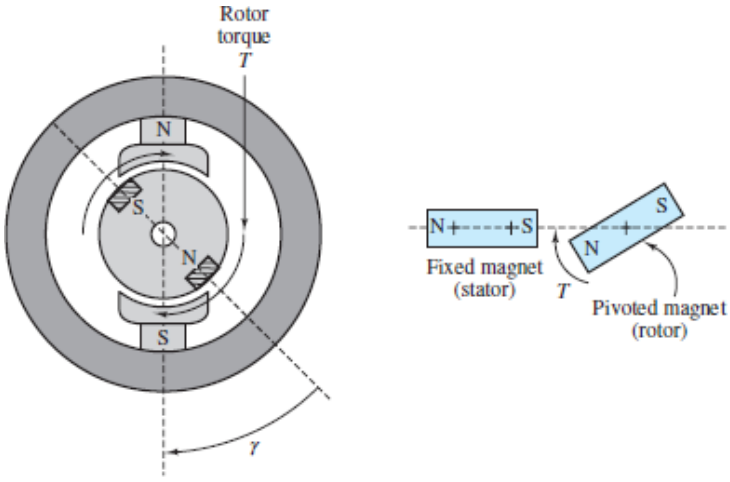


Figure 15.7 Alignment action of poles

Figure 15.8 depicts a two-pole machine in which the stator poles are constructed in such a way as to project closer to the rotor than to the stator structure. This type of construction is rather common, and poles constructed in this fashion are called **salient poles**. Note that the rotor could also be constructed to have salient poles.

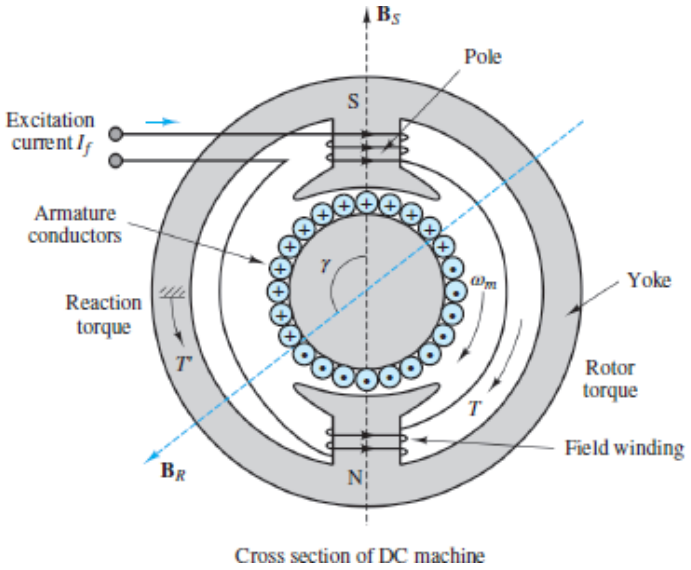


Figure 15.8 A two-pole machine with salient stator poles

To understand magnetic polarity, we need to consider the direction of the magnetic field in a coil carrying current. Figure 15.9 shows how the *right-hand rule* can be employed to determine the direction of the magnetic flux. If one were to grasp the coil with the right hand, with the fingers curling in the direction of current flow, then the thumb would be pointing in the direction of the magnetic flux. Magnetic flux by convention is viewed as entering the south pole and exiting from the north pole. Thus, to determine whether a magnetic pole is Page 878north or south, we must consider the direction of the flux. Figure 15.10 shows a cross section of a coil wound around a pair of salient rotor poles. In this case, one can readily identify the direction of the magnetic flux in the rotor and therefore the magnetic polarity of the poles by applying the right-hand rule, as illustrated in the figure.

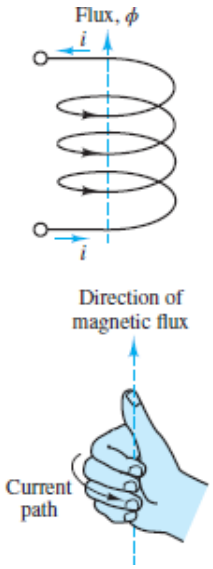


Figure 15.9 Right-hand rule

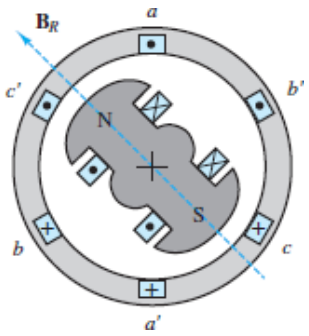


Figure 15.10 Magnetic field in a salient rotor winding

Often, however, the coil windings are not arranged as simply as in the case of salient poles. In many machines, the windings are embedded in slots cut into the stator or rotor, so that the situation is similar to that of the stator depicted in Figure 15.11. This figure is a cross section in which the wire connections between "crosses" and "dots" have been cut away. In Figure 15.11, the dashed line indicates the axis of the stator flux according to the right-hand rule, showing that the slotted stator in effect behaves as a pole pair. The north and south poles indicated in the figure are a consequence of the fact that the flux exits the top part of the structure (thus, the north pole indicated in the figure) and enters the bottom half of the structure (thus, the south pole). In particular, if you consider Page 879that the windings are arranged so that the current entering the right-hand side of the stator (to the right of the dashed line) flows through the back end of the stator and then flows outward from the left-hand side of the stator slots (left of the

dashed line), you can visualize the windings in the slots as behaving in a manner similar to the coils of Figure 15.10, where the flux axis of Figure 15.11 corresponds to the flux axis of each of the coils of Figure 15.10. The actual circuit that permits current flow is completed by the front and back ends of the stator, where the wires are connected according to the pattern *a*-*a*′, *b*-*b*′, *c*-*c*′, as depicted in the figure.

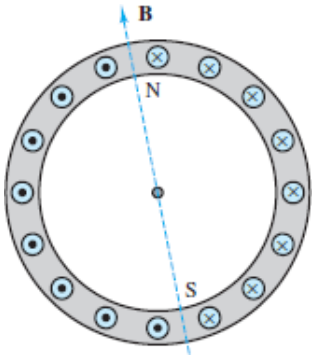


Figure 15.11 Magnetic field of stator

Another important consideration that facilitates understanding of the operation of electric machines pertains to the use of alternating currents. It should be apparent by now that if the current flowing into the slotted stator is alternating, the direction of the flux will also alternate, so that in effect the two poles will reverse polarity every time the current reverses direction, that is, every half-cycle of the sinusoidal current. Further—since the magnetic flux is approximately proportional to the current in the coil—as the amplitude of the current oscillates in a sinusoidal fashion, so will the flux density in the structure. Thus, *the magnetic field developed in the stator changes both spatially and in time.*

This property is typical of AC machines, where a *rotating magnetic field* is established by energizing the coil with an alternating current. As explained in Section 15.2, the principles underlying the operation of DC and AC machines are quite different: In a direct-current machine, there is no rotating field, but a mechanical switching arrangement (the *commutator*) makes it possible for the rotor and stator magnetic fields to always align at right angles to each other.

The book website includes two-dimensional “animations” of the most common types of electric machines. You might wish to explore these animations to better understand the basic concepts described in this section.

15.2 DIRECT-CURRENT MACHINES

As explained in the introductory section, DC machines are easier to analyze than their AC counterparts although their actual construction is made rather complex by the need to have a commutator, which switches the load winding connection to the source so as to always maintain an angle close to 90° between the stator and the rotor magnetic fields. The objective of this section is to describe the major construction features and the operation of DC machines, as well as to develop simple circuit models that are useful in analyzing the performance of this class of machines.

Physical Structure of DC Machines

A representative DC machine was depicted in Figure 15.8, with the magnetic poles clearly identified, for both the stator and the rotor. Figure 15.12 is a photograph of the same type of machine. Note the salient pole construction of the stator and the slotted rotor. As previously stated, the torque developed by the machine is a consequence of the magnetic forces between stator and rotor poles. This torque is maximum when the angle γ between the rotor and stator poles is 90°. Also, as you can see from the figure, in a DC machine the armature circuit is on the rotor, and the field winding is on the stator.

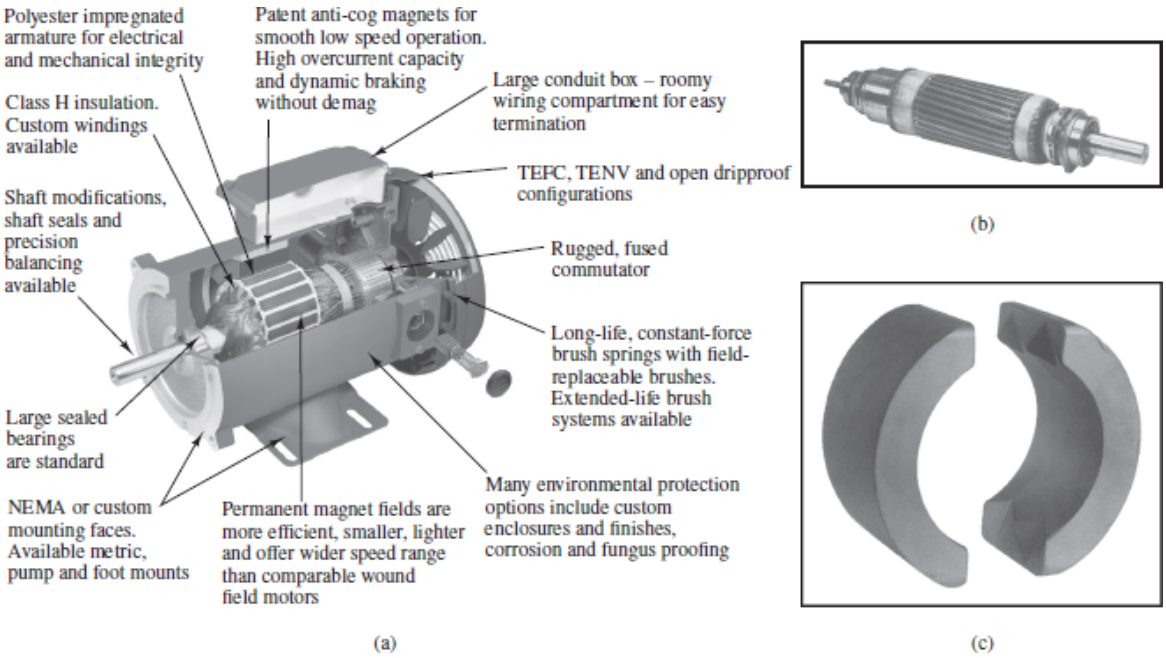


Figure 15.12 (a) DC machine; (b) rotor; (c) permanent-magnet stator (*Photos copyright © 2005, Rockwell Automation. All rights reserved. Used with permission.*)

To keep this torque angle close to 90° as the rotor spins on its shaft, a mechanical switch, called a **commutator**, is configured so the rotor poles are consistently close to 90° with respect to the fixed stator poles. In a DC machine, the magnetizing current is DC so that there is no spatial alternation of the stator poles due to time-varying currents. To understand the operation of the commutator, consider the simplified diagram of Figure 15.13. In the figure, the brushes are fixed, and the rotor revolves at an angular velocity ω_m; the instantaneous position of the rotor is given by the expression $\theta = \omega_m t - \gamma$.

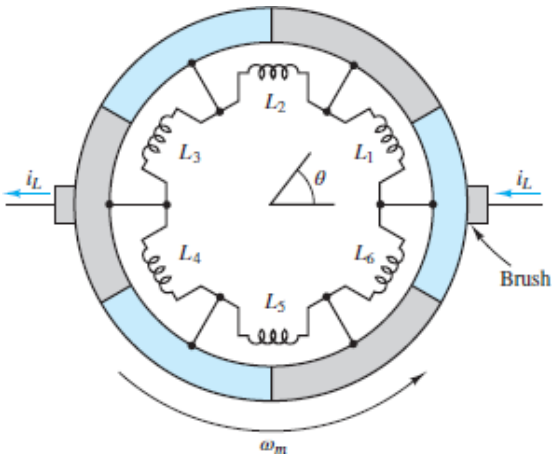


Figure 15.13 Rotor winding and commutator

The commutator is fixed to the rotor and is made up in this example of six segments that are made of electrically conducting material but are insulated from one another. Further, the rotor windings are configured so that they form six coils, connected to the commutator segments as shown in Figure 15.13.

As the commutator rotates counterclockwise, the rotor magnetic field rotates with it up to $\theta = 30°$. At that point, the direction of the current changes in coils L_3 and L_6 as the brushes make contact with the next segment. Now the direction of the magnetic field is –30°. As the commutator continues to rotate, the direction of the rotor field will again change from –30° to +30°, and it will switch again when the brushes switch to the next pair of segments. In this machine, then, the torque angle γ is not always 90°, but can vary by as much as ±30°; the actual torque produced by the machine would fluctuate by as much as ±14 percent, since the torque is proportional to sin γ. As the number of segments increases, the torque fluctuation produced by the commutation is greatly reduced. In a practical machine, for example, one might have as many as 60 segments, and the variation of γ from 90° would be only ±3°, with a torque fluctuation of less than 1 percent. Thus, the DC machine can produce a nearly constant torque (as a motor) or voltage (as a generator).

Configuration of DC Machines

The DC machine of Figure 15.12 employs a permanent magnet to generate a constant magnetic field in the stator. However, in DC machines, the field excitation that provides the magnetizing current may be provided by an external source, in which case the machine is said to be **separately excited** [Figure 15.14(a)]. More often, the field excitation is derived from the armature voltage, and the machine is said to be **self-excited**. The latter configuration does not require the use of a separate source for the field excitation and is therefore frequently preferred. If a machine is in the separately excited configuration, an additional source V_f is required. In the self-excited case, one method used to provide the field excitation is to connect the field in parallel with the armature; since the field winding typically has significantly higher resistance than the armature circuit (remember that it is the armature that carries the load current), this will not draw excessive current from the armature. Further, a series resistor can be added to the field circuit to provide the means for adjusting the field current independent of the armature voltage. This configuration is called a **shunt-**

connected machine and is depicted in Figure 15.14(b). Another method for self-exciting a DC machine consists of connecting the field in series with the armature, leading to the **series-connected** machine, depicted in Figure 15.14(c); in this case, the field winding will support the entire armature current, and thus the field coil must have low resistance (and therefore relatively few turns). This configuration is rarely used for generators, since the generated voltage and the load voltage must always differ by the voltage drop across the field coil, which varies with the load current. Thus, a series generator would have poor (large) regulation. However, series-connected motors are commonly used in traction applications.

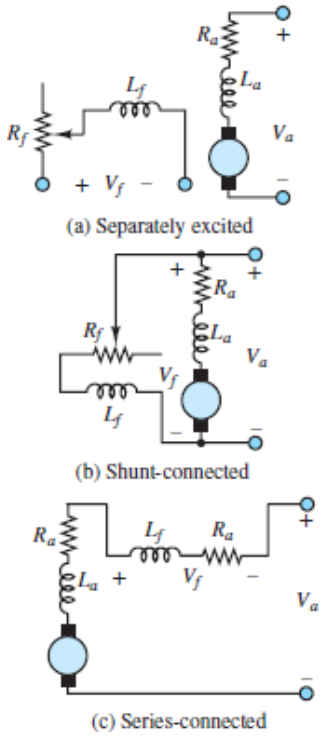


Figure 15.14 Basic configurations of DC machines with field excitation.

DC Machine Models

As stated earlier, it is relatively easy to develop a simple model of a DC machine, which is well suited to performance analysis, without the need to resort to the Page 882details of the construction of the machine itself. This section illustrates

the development of such models in two steps. First, algebraic equations relating field and armature currents and voltages to speed and torque are introduced; second, the differential equations describing the dynamic behavior of DC machines are derived.

When a field excitation is established, a magnetic flux ϕ is generated by the field current I_f. From equation 15.2, we know that the torque acting on the rotor is proportional to the product of the magnetic field and the current in the load-carrying wire; the latter current is the armature current I_a (i_w in equation 14.2). Assuming that, by virtue of the commutator, the torque angle γ is kept very close to 90°, and therefore sin $\gamma = 1$, we obtain the following expression for the torque (in units of newton-meters) in a DC machine:

$$T = k_T \phi I_a \quad \text{for } \gamma = 90° \quad \text{DC machine torque} \tag{15.6}$$

You may recall that this is simply a consequence of the *Bli* law of Chapter 14. The mechanical power generated (or absorbed) is equal to the product of the machine torque and the mechanical speed of rotation ω_m rad/s, and is therefore given by

$$P_m = \omega_m T = \omega_m k_T \phi I_a \tag{15.7}$$

Recall now that the rotation of the armature conductors in the field generated by the field excitation causes a **back emf** E_b in a direction that opposes the rotation of the armature. According to the *Blu* law (see Chapter 14), then, this back emf is given by

$$E_b = k_a \phi \omega_m \quad \text{DC machine back emf} \tag{15.8}$$

where k_a is called the **armature constant** and is related to the geometry and magnetic properties of the structure. The voltage E_b represents a countervoltage (opposing the DC excitation) in the case of a motor and the generated voltage in

the case of a generator. Thus, the electric power dissipated (or generated) by the machine is given by the product of the back emf and the armature current:

$$P_e = E_b I_a \tag{15.9}$$

The constants k_T and k_a in equations 15.6 and 15.8 are related to geometry factors, such as the dimension of the rotor and the number of turns in the armature winding, and to properties of materials, such as the permeability of the magnetic materials. Note that in the ideal energy conversion case $P_m = P_e$, and therefore $k_a = k_T$. We shall in general assume such ideal conversion of electric to mechanical energy (or vice versa) and will therefore treat the two constants as being identical: $k_a = k_T$. The constant k_a is given by

$$k_a = \frac{pN}{2\pi M} \tag{15.10}$$

where:

p = number of magnetic poles

N = number of conductors per coil

M = number of parallel paths in armature winding

Page 883

An important observation concerning the units of angular speed must be made at this point. The equality (under the no-loss assumption) between the constants k_a and k_T in equations 15.6 and 15.8 results from the choice of consistent units, namely, volts and amperes for the electrical quantities and newton-meters and radians per second for the mechanical quantities. You should be aware that it is fairly common practice to refer to the speed of rotation of an electric machine in units of revolutions per minute (r/min).[1] In this book, we shall uniformly use the symbol n to denote angular speed in revolutions per minute; the following relationship should be committed to memory:

$$n\,(\text{r/min}) = \frac{60}{2\pi}\omega_m \qquad \text{rad/s} \tag{15.11}$$

If the speed is expressed in revolutions per minute, the armature constant changes as follows:

$$E_b = k_a' \phi n \tag{15.12}$$

where

$$k_a' = \frac{pN}{60M} \tag{15.13}$$

Having introduced the basic equations relating torque, speed, voltages, and currents in electric machines, we may now consider the interaction of these quantities in a DC machine at steady state, that is, operating at constant speed and field excitation. Figure 15.15 depicts the electric circuit model of a separately excited DC machine, illustrating both motor and generator action. It is very important to note the reference direction of armature current flow, and of the developed torque, to make a distinction between the two modes of operation. The field excitation is shown as a voltage V_f generating the field current I_f that flows through a variable resistor R_f and through the field coil L_f. The variable resistor permits adjustment of the field excitation. The armature circuit, on the other hand, consists of a voltage source representing the back emf E_b, the armature resistance R_a, and the armature voltage V_a. This model is appropriate both for motor and for generator action. When $V_a < E_b$, the machine acts as a generator (I_a flows out of the machine). When $V_a > E_b$, the machine acts as a motor (I_a flows into the machine). Thus, according to the circuit model of Figure 15.15, the operation of a DC machine at steady state (i.e., with the inductors in the circuit replaced by short-circuits) is described by the following equations:

$$\begin{aligned} -I_f + \frac{V_f}{R_f} = 0 \quad &\text{and} \quad V_a - R_a I_a - E_b = 0 \quad \text{(motor action)} \\ -I_f + \frac{V_f}{R_f} = 0 \quad &\text{and} \quad V_a + R_a I_a - E_b = 0 \quad \text{(generator action)} \end{aligned} \tag{15.14}$$

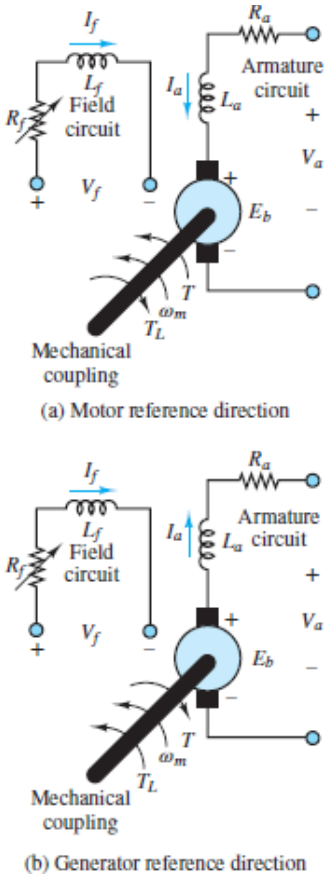


Figure 15.15 Electric circuit model of a separately excited DC machine

Equation 15.14 together with equations 15.6 and 15.8 may be used to determine the steady-state operating condition of a DC machine.

Page 884

The circuit model of Figure 15.15 permits the derivation of a simple set of differential equations that describe the *dynamic* analysis of a DC machine. The dynamic equations describing the behavior of a separately excited DC machine are as follows:

$$V_a(t) - I_a(t)R_a - L_a\frac{dI_a(t)}{dt} - E_b(t) = 0 \qquad \text{(armature circuit)} \tag{15.15a}$$

$$V_f(t) - I_f(t)R_f - L_f\frac{dI_f(t)}{dt} = 0 \qquad \text{(field circuit)} \tag{15.15b}$$

These equations can be related to the operation of the machine in the presence of a load. If we assume that the motor is rigidly connected to an inertial load with

moment of inertia J and that the friction losses in the load are represented by a viscous friction coefficient b, then the torque developed by the machine (in the motor mode of operation) can be written as

$$T(t) = T_L + b\omega_m(t) + J\frac{d\omega_m(t)}{dt} \tag{15.16}$$

where T_L is the load torque. Typically T_L is either constant or some function of speed ω_m in a motor. In the case of a generator, the load torque is replaced by the torque supplied by a prime mover, and the machine torque $T(t)$ opposes the motion of the prime mover, as shown in Figure 15.15. Since the machine torque is related to the armature and field currents by equation 15.6, equations 15.16 and 15.17 are coupled to each other; this coupling may be expressed as follows:

$$T(t) = k_a \phi I_a(t) \tag{15.17}$$

or

$$k_a \phi I_a(t) = T_L + b\omega_m(t) + J\frac{d\omega_m(t)}{dt} \tag{15.18}$$

The dynamic equations described in this section apply to any DC machine. In the case of a *separately excited* machine, a further simplification is possible, since the flux is established by virtue of a separate field excitation, and therefore

$$\phi = \frac{N_f}{\mathcal{R}} I_f = k_f I_f \tag{15.19}$$

where N_f is the number of turns in the field coil, $\mathcal{R}$ is the reluctance of the structure, and I_f is the field current.

DC Machine Steady-State Equations

The equations that describe the steady-state behavior of DC motors and generators are summarized below. The key to interpreting these equations is in correctly evaluating the expression for the flux ϕ for each of the four cases of interest in this chapter: field generated by a separate excitation, field generated by a shunt connection, field generated by a series connection, and field generated by a permanent magnet (constant field). See Figure 15.14 for a reference to the first three configurations.

DC Motor Steady-State Equations

$$E_b = k_a \phi \omega_m \quad \text{V}$$
$$T = k_a \phi I_a \quad \text{N-m}$$

In a *separately excited* machine [Figure 15.14(a)]:

$$V_s = E_b - I_a R_a \quad \text{V}$$
$$\phi = k_f I_f = k_f \frac{V_f}{R_f}$$

where V_s is the external source voltage.

In a *shunt-connected* machine [Figure 15.14(b)]:

$$V_s = E_b - I_a R_a \quad \text{V}$$
$$\phi = \phi_{\text{shunt}} = k_f I_f = k_f \frac{V_a}{R_f}$$

In a *series-connected* machine [Figure 15.14(c)]:

$$V_s = E_b - I_a R_a - I_a R_f \quad \text{V}$$
$$\phi = \phi_{\text{series}} = k_f I_f = k_f I_a$$

Finally, in a *permanent-magnet* machine, where the field excitation is provided by a permanent magnet:

$$V_s = E_b - I_a R_a \quad \text{V}$$
$$\phi = \phi_{\text{PM}} = \text{constant}$$

DC Generator Steady-State Equations

$$E_b = k_a \phi \omega_m \quad \text{V}$$
$$T = \frac{P}{\omega_m} = \frac{E_b I_a}{\omega_m} = k_a \phi I_a \quad \text{N-m}$$

where V_g is the generator open-circuit output voltage, with no load connected. In a *separately excited* machine [Figure 15.14(a)]:

$$V_g = E_b - I_a R_a \quad \text{V}$$
$$\phi = k_f I_f = k_f \frac{V_f}{R_f} \quad \text{Wb}$$

In a *shunt-connected* machine [Figure 15.14(b)]:

$$V_g = E_b - I_a R_a \quad \text{V}$$
$$\phi = \phi_{\text{shunt}} = k_f I_f = k_f \frac{V_g}{R_f} \quad \text{Wb}$$

In a *series-connected* machine [Figure 15.14(b)]:

$$V_g = E_b - I_a R_a - I_a R_f \quad \text{V}$$
$$\phi = \phi_{\text{series}} = k_f I_f = k_f I_a \quad \text{Wb}$$

Finally, in a *permanent-magnet* machine:

$$V_g = E_b - I_a R_a \quad \text{V}$$
$$\phi = \phi_{\text{PM}} = \text{constant} \quad \text{Wb}$$

Page 886

15.3 DIRECT-CURRENT MOTORS

DC motors are widely used in applications requiring accurate speed or torque control, for example in servo systems. In the preceding section, we had introduced the analysis of a separately excited DC machine; in this section we extend that analysis to include a review of the other three commonly used configurations (*shunt, series*, and *permanent-magnet motors*) to study their torque–speed characteristics and dynamic behavior.

The Shunt Motor

In a shunt motor [see Figure 15.14(b)], the armature current is found by dividing the net voltage across the armature circuit (source voltage minus back emf) by the armature resistance:

$$I_a = \frac{V_s - k_a \phi \omega_m}{R_a} \tag{15.20}$$

An expression for the armature current may also be obtained from equation 15.17, as follows:

$$I_a = \frac{T}{k_a \phi} \tag{15.21}$$

It is then possible to relate the torque requirements to the speed of the motor by substituting equation 15.20 in equation 15.21:

$$\frac{T}{k_a \phi} = \frac{V_s - k_a \phi \omega_m}{R_a} \tag{15.22}$$

Equation 15.22 describes the steady-state torque–speed characteristic of the shunt motor. To understand this performance equation, we observe that if V_s, k_a, ϕ, and R_a are fixed in equation 15.22 (the flux is essentially constant in the shunt motor for a fixed V_s), then the speed of the motor is directly related to the armature current. Now consider the case where the load applied to the motor is suddenly increased, causing the speed of the motor to drop. As the speed decreases, the armature current increases, according to equation 15.20. The excess armature current causes the motor to develop additional torque, according to equation 15.21 until a new equilibrium is reached between the higher armature current and developed torque and the lower speed of rotation. The equilibrium point is dictated by the balance of mechanical and electric power, in accordance with the relation:

$$E_b I_a = T \omega_m \tag{15.23}$$

Thus, the shunt DC motor will adjust to variations in load by changing its speed to preserve this power balance. The torque–speed curves for the shunt motor may be obtained by rewriting the equation relating the speed to the armature current:

$$\omega_m = \frac{V_s - I_a R_a}{k_a \phi} = \frac{V_s}{k_a \phi} - \frac{R_a T}{(k_a \phi)^2} \qquad T\text{-}\omega \text{ curve for shunt motor} \tag{15.24}$$

Page 887

To interpret equation 15.24, one can start by considering the motor operating at rated speed and torque. As the load torque is reduced, the armature current will also decrease, causing the speed to increase in accordance with equation 15.24. The increase in speed depends on the extent of the voltage drop across the armature resistance $I_a R_a$. The change in speed will be on the same order of magnitude as this drop; it typically takes values around 10 percent. This

corresponds to a relatively good speed regulation, which is an attractive feature of the shunt DC motor (recall the discussion of regulation in Section 15.1). The dynamic behavior of the shunt motor is described by equations 15.15 through 15.18, with the additional relation:

$$I_a(t) = I_s(t) - I_f(t) \tag{15.25}$$

Series Motors

The series motor [see Figure 15.14(c)] behaves somewhat differently from the shunt and separately excited motors because the flux is established solely by virtue of the series current flowing through the armature. It is relatively simple to derive an expression for the emf and torque equations for the series motor if we approximate the relationship between flux and armature current by assuming that the motor operates in the linear region of its magnetization curve. Then we can write

$$\phi = k_S I_a \tag{15.26}$$

and the emf and torque equations become, respectively,

$$E_b = k_a \omega_m \phi = k_a \omega_m k_S I_a \tag{15.27}$$

$$T = k_a \phi I_a = k_a k_S I_a^2 \tag{15.28}$$

The circuit equation for the series motor becomes

$$V_s = E_b + I_a(R_a + R_S) = (k_a \omega_m k_S + R_T) I_a \tag{15.29}$$

where R_a is the armature resistance, R_S is the series field winding resistance, and R_T is the total series resistance. From equation 15.29, we can solve for I_a and substitute in the torque expression (equation 15.28) to obtain the following torque–speed relationship:

$$T = k_a k_S \frac{V^2}{(k_a \omega_m k_S + R_T)^2} \qquad T\text{-}\omega \text{ curve for series DC motor} \tag{15.30}$$

which indicates the inverse squared relationship between torque and speed in the series motor. This expression describes a behavior that can, under certain conditions, become unstable. Since the speed increases when the load torque is reduced, one can readily see that if one were to disconnect the load altogether, the speed would tend to increase to dangerous values. To prevent excessive speeds, series motors are always mechanically coupled to the load. This feature is not necessarily a drawback, though, because series motors can develop very high torque at low speeds and therefore can serve very well for traction-type loads (e.g., conveyor belts or vehicle propulsion systems).

Page 888

The differential equation for the armature circuit of the motor can be given as

$$\begin{aligned} V_s &= I_a(t)(R_a + R_S) + L_a\frac{dI_a(t)}{dt} + L_S\frac{dI_a(t)}{dt} + E_b \\ &= I_a(t)(R_a + R_S) + L_a\frac{dI_a(t)}{dt} + L_S\frac{dI_a(t)}{dt} + k_a k_S I_a \omega_m \end{aligned} \tag{15.31}$$

Permanent-Magnet DC Motors

Permanent-magnet (PM) DC motors have become increasingly common in applications requiring relatively low torques and efficient use of space. The construction of PM DC motors differs from that of the motors considered thus far in that the magnetic field of the stator is produced by suitably located poles made of magnetic materials. Thus, the basic principle of operation, including the idea of commutation, is unchanged with respect to the wound-stator DC motor. What changes is that there is no need to provide a field excitation, whether separately or by means of the self-excitation techniques discussed in the preceding sections. Therefore, the PM motor is intrinsically simpler than its wound-stator counterpart.

The equations that describe the operation of the PM motor follow. The torque produced is related to the armature current by a torque constant k_{PM}, which is determined by the geometry of the motor:

$$T = k_{T,\text{PM}} I_a \tag{15.32}$$

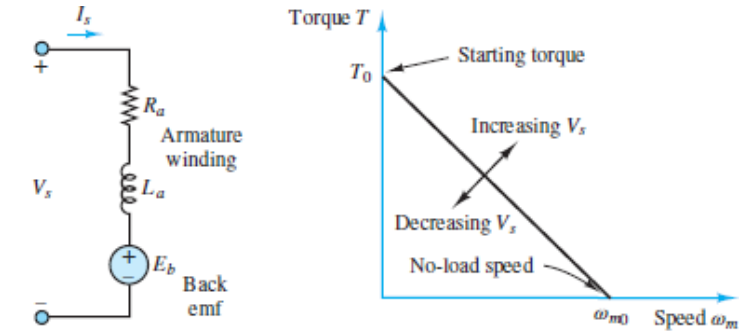
As in the conventional DC motor, the rotation of the rotor produces the usual count or back emf E_b, which is linearly related to speed by a voltage constant $k_{a,\text{PM}}$:

$$E_b = k_{a,\mathrm{PM}}\omega_m \tag{15.33}$$

The equivalent circuit of the PM motor is particularly simple, since we need not model the effects of a field winding. Figure 15.16 shows the circuit model and the torque–speed curve of a PM motor.

We can use the circuit model of Figure 15.16 to derive the torque–speed curve shown in the same figure as follows. From the circuit model, for a constant speed (and therefore constant current), we may consider the inductor a short-circuit and write the equation:

$$\begin{aligned} V_s &= I_a R_a + E_b = I_a R_a + k_{a,\mathrm{PM}}\omega_m \\ &= \frac{T}{k_{T,\mathrm{PM}}} R_a + k_{a,\mathrm{PM}}\omega_m \end{aligned} \tag{15.34}$$



Circuit model for PM motor

Torque-speed curve for PM motor

Figure 15.16 Circuit model and torque–speed curve of PM motor

Page 889

thus obtaining the equations relating speed and torque:

LO

$$\omega_m = \frac{V_s}{k_{a,\mathrm{PM}}} - \frac{TR_a}{k_{a,\mathrm{PM}}} k_{T,\mathrm{PM}} \qquad T\text{-}\omega \text{ curve for PM DC motor} \tag{15.35}$$

and

$$T = \frac{V_s}{R_a} k_{T,\mathrm{PM}} - \frac{\omega_m}{R_a} k_{a,\mathrm{PM}}\, k_{T,\mathrm{PM}} \tag{15.36}$$

From these equations, one can extract the stall torque T_0, that is, the zero-speed torque:

$$T_0 = \frac{V_s}{R_a} k_{T,\text{PM}} \tag{15.37}$$

and the no-load speed ω_{m0}:

$$\omega_{m0} = \frac{V_s}{k_{a,\text{PM}}} \tag{15.38}$$

Under dynamic conditions, assuming an inertia plus viscous friction load, the torque produced by the motor can be expressed as

$$T = k_{T,\text{PM}} I_a(t) = T_{\text{load}}(t) + b\,\omega_m(t) + J\frac{d\omega_m(t)}{dt} \tag{15.39}$$

The differential equation for the armature circuit of the motor is therefore given by

$$\begin{aligned} V_s &= I_a(t)R_a + L_a\frac{dI_a(t)}{dt} + E_b \\ &= I_a(t)R_a + L_a\frac{dI_a(t)}{dt} + k_{a,\text{PM}}\omega_m(t) \end{aligned} \tag{15.40}$$

The fact that the airgap flux is constant in a PM DC motor makes its characteristics somewhat different from those of the wound DC motor. A direct comparison of PM and wound-field DC motors reveals the following advantages and disadvantages of each configuration.

Comparison of Wound-Field and PM DC Motors

1. PM motors are smaller and lighter than wound motors for a given power rating. Further, their efficiency is greater because there are no field winding losses.
2. An additional advantage of PM motors is their essentially linear speed–torque characteristic, which makes analysis (and control) much easier. Reversal of rotation is also accomplished easily, by reversing the polarity of the source.

3. A major disadvantage of PM motors is that they can become demagnetized by exposure to excessive magnetic fields, application of excessive voltage, or operation at excessively high or low temperatures.
4. A less obvious drawback of PM motors is that their performance is subject to greater variability from motor to motor than is the case for wound motors, because of variations in the magnetic materials.

EXAMPLE 15.4 DC Shunt Motor Analysis

Problem

Find the speed and torque generated by a four-pole DC shunt motor.

Solution

Known Quantities: Motor ratings; circuit and magnetic parameters.

Find: ω_m, T.

Schematics, Diagrams, Circuits, and Given Data:

Motor ratings: 3 hp, 240 V, 120 r/min.

Circuit and magnetic parameters: $I_S = 30$ A; $I_f = 1.4$ A; $R_a = 0.6\ \Omega$; $\phi = 20$ mWb; $N = 1{,}000$; $M = 4$ (see equation 15.10).

Analysis: We convert the power to SI units:

$$P_{\text{RATED}} = 3\ \text{hp} \times 746 \frac{\text{W}}{\text{hp}} = 2{,}238\ \text{W}$$

Next we compute the armature current as the difference between source and field current (equation 15.25):

$$I_a = I_s - I_f = 30 - 1.4 = 28.6\ \text{A}$$

The no-load armature voltage E_b is given by:

$$E_b = V_s - I_a R_a = 240 - 28.6 \times 0.6 = 222.84 \text{ V}$$

and equation 15.10 can be used to determine the armature constant:

$$k_a = \frac{pN}{2\pi M} = \frac{4 \times 1{,}000}{2\pi \times 4} = 159.15 \frac{\text{V-s}}{\text{Wb-rad}}$$

Knowing the motor constant, we can calculate the speed, after equation 15.25:

$$\omega_m = \frac{E_a}{k_a \phi} = \frac{222.84 \text{ V}}{(159.15 \text{ V-s/Wb-rad})(0.02 \text{ Wb})} = 70 \frac{\text{rad}}{\text{s}}$$

Finally, the torque developed by the motor can be found as the ratio of the power to the angular velocity:

$$T = \frac{P}{\omega_m} = \frac{2{,}238 \text{ W}}{70 \text{ rad/s}} = 32 \text{ N-m}$$

CHECK YOUR UNDERSTANDING

A 200-V DC shunt motor draws 10 A at 1,800 r/min. The armature circuit resistance is 0.15 Ω, and the field winding resistance is 350 Ω. What is the torque developed by the motor?

Answer: $T = \frac{P}{\omega_m} = 9.93$ N-m

Page 891

LO

EXAMPLE 15.5 DC Shunt Motor Analysis

Problem

Determine the following quantities for the DC shunt motor, connected as shown in the circuit of Figure 15.17:

1. Field current required for full-load operation.
2. No-load speed.
3. Plot of the speed torque curve of the machine in the range from no-load torque to rated torque.
4. Power output at rated torque.

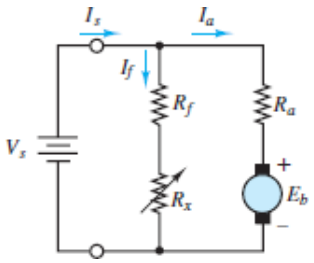


Figure 15.17 Shunt motor configuration

Solution

Known Quantities: Magnetization curve, rated current, rated speed, circuit parameters.

Find: I_f; $n_{\text{no-load}}$; *T-n* curve, P_{rated}.

Schematics, Diagrams, Circuits, and Given Data:

Figure 15.18 (magnetization curve)

Motor ratings: 8 A, 120 r/min

Circuit parameters: $R_a = 0.2\ \Omega$; $V_s = 7.2$ V; N = number of coil turns in winding = 200

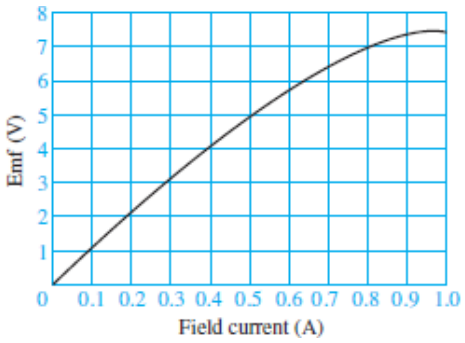


Figure 15.18 Magnetization curve for a small DC motor

Analysis:

1. To find the field current, we must find the generated emf since R_f is not known. Writing KVL around the armature circuit, we obtain

$$V_s = E_b + I_a R_a$$
$$E_b = V_s - I_a R_a = 7.2 - 8(0.2) = 5.6 \text{ V}$$

Having found the back emf, we can find the field current from the magnetization curve. At E_b = 5.6 V, we find that the field current and field resistance are

$$I_f = 0.6 \text{ A} \quad \text{and} \quad R_f = \frac{7.2}{0.6} = 12 \ \Omega$$

2. To obtain the no-load speed, we use the equations:

$$E_b = k_a \phi \frac{2\pi n}{60} \qquad T = k_a \phi I_a$$

Page 892

leading to

$$V_s = I_a R_a + E_b = I_a R_a + k_a \phi \frac{2\pi}{60} n$$

or

$$n = \frac{V_s - I_a R_a}{k_a \phi (2\pi/60)}$$

At no load, and assuming no mechanical losses, the torque is zero, and we see that the current I_a must also be zero in the torque equation ($T = k_a \phi I_a$). Thus, the motor speed at no load is given by

$$n_{\text{no-load}} = \frac{V_s}{k_a \phi (2\pi/60)}$$

We can obtain an expression for $k_a \phi$ knowing that, at full load:

$$E_b = 5.6 \text{ V} = k_a \phi \frac{2\pi n}{60}$$

so that, for constant field excitation:

$$k_a\phi = E_b\left(\frac{60}{2\pi n}\right) = 5.6\left[\frac{60}{2\pi(120)}\right] = 0.44563\,\frac{\text{V-s}}{\text{rad}}$$

Finally, we may solve for the no-load speed.

$$n_{\text{no-load}} = \frac{V_s}{k_a\phi(2\pi/60)} = \frac{7.2}{(0.44563)(2\pi/60)}$$
$$= 154.3 \text{ r/min}$$

3. The torque at rated speed and load may be found as follows:

$$T_{\text{rated load}} = k_a\phi I_a = (0.44563)(8) = 3.565 \text{ N-m}$$

Now we have the two points necessary to construct the torque–speed curve for this motor, which is shown in Figure 15.19.

4. The power is related to the torque by the frequency of the shaft:

$$P_{\text{rated}} = T\omega_m = (3.565)\left(\frac{120}{60}\right)(2\pi) = 44.8 \text{ W}$$

or, equivalently:

$$P = 44.8\,\text{W} \times \frac{1}{746}\frac{\text{hp}}{\text{W}} = 0.06 \text{ hp}$$

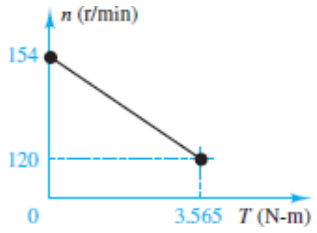


Figure 15.19 Torque-speed curve for motor of Example 15.5.

EXAMPLE 15.6 DC Series Motor Analysis

Problem

Determine the torque developed by a DC series motor when the current supplied to the motor is 60 A.

Solution

Known Quantities: Motor ratings; operating conditions.

Find: T_{60}, torque delivered at 60-A series current.

Page 893

Schematics, Diagrams, Circuits, and Given Data:

Motor ratings: 10 hp, 115 V, full-load speed = 1,800 r/min

Operating conditions: motor draws 40 A

Assumptions: The motor operates in the linear region of the magnetization curve.

Analysis: Within the linear region of operation, the flux per pole is directly proportional to the current in the field winding. That is,

$$\phi = k_S I_a$$

The full-load speed is

$$n = 1{,}800 \text{ r/min}$$

or

$$\omega_m = \frac{2\pi n}{60} = 60\pi \qquad \text{rad/s}$$

Rated output power is

$$P_{\text{rated}} = 10 \text{ hp} \times 746 \text{ W/hp} = 7{,}460 \text{ W}$$

and full-load torque is

$$T_{40\text{ A}} = \frac{P_{\text{rated}}}{\omega_m} = \frac{7{,}460}{60\pi} = 39.58 \text{ N-m}$$

Thus, the machine constant may be computed from the torque equation for the series motor:

$$T = k_a k_s I_a^2 = K I_a^2$$

At full load:

$$K = k_a k_s = \frac{39.58 \text{ N-m}}{40^2 \text{ A}^2} = 0.0247 \frac{\text{N-m}}{\text{A}^2}$$

and we can compute the torque developed for a 60-A supply current to be

$$T_{60\text{ A}} = K I_a^2 = 0.0247 \times 60^2 = 88.92 \text{ N-m}$$

CHECK YOUR UNDERSTANDING

A series motor draws a current of 25 A and develops a torque of 100 N-m. Find (a) the torque when the current rises to 30 A if the field is unsaturated and (b) the torque when the current rises to 30 A and the increase in current produces a 10 percent increase in flux.

Answer: (a) 144 N-m; (b) 132 N-m

LO

EXAMPLE 15.7 Dynamic Response of PM DC Motor

Problem

Develop a set of differential equations and a transfer function describing the dynamic response of the motor angular velocity of a PM DC motor connected to a mechanical load.

Page 894

Solution

Known Quantities: PM DC motor circuit model; mechanical load model.

Find: Differential equations and transfer functions of electromechanical system.

Analysis: The dynamic response of the electromechanical system can be determined by applying KVL to the electric circuit (Figure 15.16) and Newton's second law to the mechanical system. These equations will be coupled to one another, as you shall see, because of the nature of the motor back emf and torque equations.

Applying KVL and equation 15.33 to the electric circuit, we obtain

$$V_L(t) - R_a I_a(t) - L_a \frac{dI_a(t)}{dt} - E_b(t) = 0$$

or

$$L_a \frac{dI_a(t)}{dt} + R_a I_a(t) + K_{a\text{-PM}} \omega_m(t) = V_L(t)$$

Applying Newton's second law and equation 15.32 to the load inertia, we obtain

$$J \frac{d\omega(t)}{dt} = T(t) - T_{\text{load}}(t) - b\omega$$

or

$$-K_{T\text{-PM}} I_a(t) + J \frac{d\omega(t)}{dt} + b\omega(t) = -T_{\text{load}}(t)$$

These two differential equations are coupled because the first depends on ω_m and the second on I_a. Thus, they need to be solved simultaneously.

To derive the transfer function, we use the Laplace transform on the two equations to obtain

$$\begin{aligned}(sL_a + R_a) I_a(s) + K_{a\text{-PM}} \Omega(s) &= V_L(s) \\ -K_{T\text{-PM}} I_a(s) + (sJ + b)\Omega(s) &= -T_{\text{load}}(s)\end{aligned}$$

We can write the above equations in matrix form and resort to Cramer's rule to solve for $\Omega_m(s)$ as a function of $V_L(s)$ and $T_{\text{load}}(s)$.

$$\begin{bmatrix} sL_a + R_a & K_{a\text{-PM}} \\ -K_{T\text{-PM}} & sJ + b \end{bmatrix} \begin{bmatrix} I_a(s) \\ \Omega_m(s) \end{bmatrix} = \begin{bmatrix} V_L(s) \\ -T_{\text{load}}(s) \end{bmatrix}$$

with solution:

$$\Omega_m(s) = \frac{\det\begin{bmatrix} sL_a + R_a & V_L(s) \\ K_{T\text{-PM}} & -T_{\text{load}}(s) \end{bmatrix}}{\det\begin{bmatrix} sL_a + R_a & K_{a\text{-PM}} \\ -K_{T\text{-PM}} & sJ + b \end{bmatrix}}$$

or

$$\Omega_m(s) = -\frac{sL_a + R_a}{(sL_a + R_a)(sJ + b) + K_{a\text{-PM}}K_{T\text{-PM}}} T_{\text{load}}(s) + \frac{K_{T\text{-PM}}}{(sL_a + R_a)(sJ + b) + K_{a\text{-PM}}K_{T\text{-PM}}} V_L(s)$$

Comments: Note that the dynamic response of the motor angular velocity depends on both the input voltage and the load torque. This problem is explored further in the homework problems.

Page 895

DC Drives and DC Motor Speed Control

The advances made in power semiconductors have made it possible to realize low-cost speed control systems for DC motors. In this section we describe some of the considerations that are behind the choice of a specific drive type, and some of the loads that are likely to be encountered.

Constant-torque loads are quite common and are characterized by a need for constant torque over the entire speed range. This need is usually due to friction; the load will demand increasing horsepower at higher speeds, since power is the product of speed and torque. Thus, the power required will increase linearly with speed. This type of loading is characteristic of conveyors, extruders, and surface winders.

Another type of load is one that requires *constant horsepower* over the speed range of the motor. Since torque is inversely proportional to speed with constant horsepower, this type of load will require higher torque at low speeds. Examples of constant-horsepower loads are machine tool spindles (e.g., lathes). This type of application requires very high starting torques.

Variable-torque loads are also common. In this case, the load torque is related to the speed in some fashion, either linearly or geometrically. For some loads, for example, torque is proportional to the speed (and thus horsepower is proportional to speed squared); examples of loads of this type are positive displacement pumps. More common than the linear relationship is the squared-speed

dependence of inertial loads such as centrifugal pumps, some fans, and all loads in which a flywheel is used for energy storage.

To select the appropriate motor and adjustable-speed drive for a given application, we need to examine how each method for speed adjustment operates on a DC motor. Armature voltage control serves to smoothly adjust speed from 0 to 100 percent of the nameplate rated value (i.e., base speed), provided that the field excitation is also equal to the rated value. Within this range, it is possible to fully control motor speed for a constant-torque load, thus providing a linear increase in horsepower, as shown in Figure 15.20. Field weakening allows for increases in speed of up to several times the base speed; however, field control changes the characteristics of the DC motor from constant torque to constant horsepower, and therefore the torque output drops with speed, as shown in Figure 15.20. Operation above base speed requires special provision for field control, in addition to the circuitry required for armature voltage control, and is therefore more complex and costly.

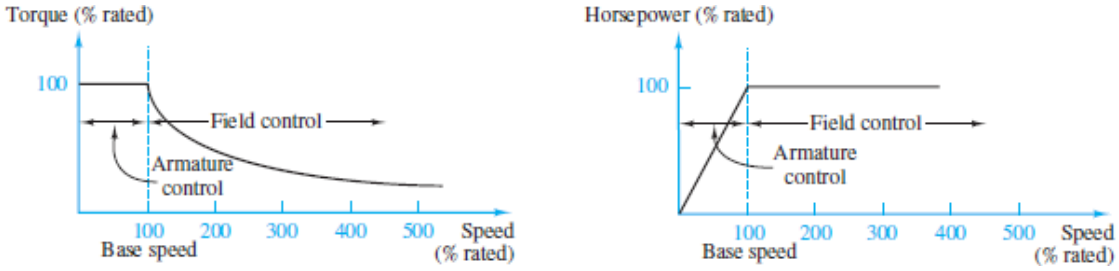


Figure 15.20 Speed control in DC motors

CHECK YOUR UNDERSTANDING

Describe the cause-and-effect behavior of the speed control method of changing armature voltage for a shunt DC motor.

current to drop and the motor torque to decrease until a balance condition is reached between motor and load torque and the motor runs at constant speed.

Answer: Increasing the armature voltage leads to an increase in armature current. Consequently, the motor torque increases until it exceeds the load torque, causing the speed to increase as well. The corresponding increase in back emf, however, causes the armature

15.4 DIRECT-CURRENT GENERATORS

The same analysis and equations used in the preceding section can be applied to DC generators, with the understanding that in a motor, the external voltage V_s is a DC supply that enables the motor to generate a torque, while in a generator the torque provided by a prime mover results in the motor rotating at a speed Ω, which in turn generates an open-circuit voltage V_g. When the generator is connected to a load, armature current flows, and a load voltage V_L is generated. Figure 15.21 depicts the configuration of a separately excited DC generator, and Figure 15.22 depicts a magnetization curve for a generator that can be used to calculate the back emf (generator open-circuit voltage) as a function of field current. Two examples follow, to illustrate methods of analysis for DC generators.

Figure 15.21 Separately excited DC generator

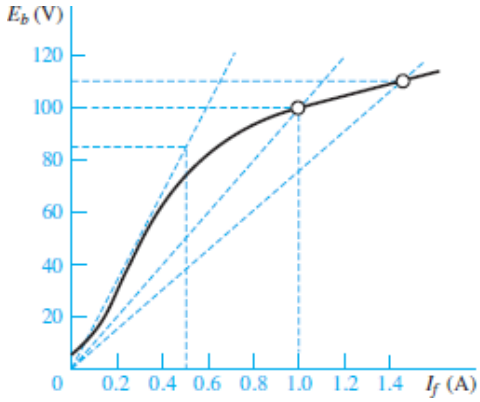


Figure 15.22 Separately excited DC generator magnetization curve

EXAMPLE 15.8 Separately Excited DC Generator

Problem

A separately excited DC generator is characterized by the magnetization of Figure 15.22.

1. If the prime mover is driving the generator at 800 r/min, what is the no-load terminal voltage V_a?
2. If a 1-Ω load is connected to the generator, what is the generator voltage?
3. Assume steady-state operation.

Page 897

Solution

Known Quantities: Generator magnetization curve and ratings.

Find: Terminal voltage with no load and 1-Ω load.

Schematics, Diagrams Circuits and Green Data: Generator ratings: 100 V, 100 A, 1,000 r/min. Circuit parameters: $R_a = 0.14\ \Omega$; $V_f = 100$ V; $R_f = 100\ \Omega$.

Analysis:

1. The field current in the machine at steady state is

$$I_f = \frac{V_f}{R_f} = \frac{100\,\text{V}}{100\,\Omega} = 1\,\text{A}$$

From the magnetization curve, it can be seen that this field current will produce 100 V at a speed of 1,000 r/min. Since this generator is actually running at 800 r/min, the induced emf may be found by assuming a linear relationship between speed and emf. This approximation is reasonable, provided that the departure from the nominal operating condition is small. Let n_0 and E_{b0} be the nominal speed and emf, respectively (that is, 1,000 r/min and 100 V). Then:

$$\frac{E_b}{E_{b0}} = \frac{n}{n_0}$$

and therefore:

$$E_b = \frac{n}{n_0} E_{b0} = \frac{800\ \text{r/min}}{1{,}000\ \text{r/min}} \times 100\ \text{V} = 80\ \text{V}$$

The open-circuit (output) terminal voltage of the generator is equal to the emf from the circuit model of Figure 15.15: therefore:

$$V_a = E_b = 80\ \text{V}$$

2. When a load resistance is connected to the circuit (the practical situation), the terminal (or load) voltage is no longer equal to E_b, since there will be a voltage drop across the armature winding resistance. The armature (or load) current may be determined from

$$I_a = I_L = \frac{E_b}{R_a + R_L} = \frac{80\ \text{V}}{(0.14 + 1)\ \Omega} = 70.2\ \text{A}$$

where $R_2 = 1\ \Omega$ is the load resistance. The terminal (load) voltage is therefore given by

$$V_L = I_L R_L = 70.2 \times 1 = 70.2\ \text{V}$$

CHECK YOUR UNDERSTANDING

A 24-coil, two-pole DC generator has 16 turns per coil in its armature winding. The field excitation is 0.05 Wb per pole, and the armature angular velocity is 180

rad/s. Find the machine constant and the total induced voltage.

Answer: $k_a = 5.1$; $E_b = 45.9$ V

Page 898

EXAMPLE 15.9 Separately Excited DC Generator

Problem

Determine the following quantities for a separately excited DC:

1. Induced voltage
2. Machine constant
3. Torque developed at rated conditions
4. Assume steady-state operation

Solution

Known Quantities: Generator ratings and machine parameters.

Find: E_b, k_a, T.

Schematics, Diagrams, Circuits, and Given Data: Generator ratings: 1,000 kW, 2,000 V, 3,600 r/min. Circuit parameters: $R_0 = 0.1\ \Omega$, flux per pole $\phi = 0.5$ Wb.

Analysis:

1. The armature current may be found by observing that the rated power is equal to the product of the terminal (load) voltage and the current. Then:

$$I_a = \frac{P_{\text{rated}}}{V_L} = \frac{1{,}000 \times 10^3}{2{,}000} = 500\ \text{A}$$

The generated voltage is equal to the sum of the terminal voltage and the voltage drop across the armature resistance (see Figure 15.14):

$$E_b = V_a + I_a R_a = 2{,}000 + 500 \times 0.1 = 2{,}050 \text{ V}$$

2. The speed of rotation of the machine in units of radians per second is

$$\omega_a = \frac{2\pi n}{60} = \frac{2\pi \times 3{,}600 \text{ r/min}}{60 \text{ r/min}} = 377 \text{ rad/s}$$

Thus, the machine constant is found to be

$$L_a = \frac{E_b}{\phi \omega_m} = \frac{2.050 \text{ V}}{0.5 \text{ Wb} \times 377 \text{ rad/s}} = 10.876 \frac{\text{V-s}}{\text{Wb-rad}}$$

3. The torque developed is found from equation 15.6:

$$T = k_a \phi I_a = 10.876 \text{ V-s/Wb-rad} \times 0.5 \text{ Wb} \times 500 \text{ A} = 2.718.9 \text{ N-m}$$

Comments: In many practical cases, it is not actually necessary to know the armature constant and the flux separately, but it is sufficient to know the value of the product $k_a \phi$. For example, suppose that the armature resistance of a DC machine is known and that, given a known field excitation, the armature current, load voltage, and speed of the machine can be measured. Then the product $k_a \phi$ may be determined from equation 15.8, as follows:

$$k_a \phi = \frac{E_b}{\omega_m} = \frac{V_L + (R_a + R_x)}{\omega_m}$$

where V_L, I_a and ω_m are measured quantities for given operating conditions.

Page 899

CHECK YOUR UNDERSTANDING

A 1,000-kW, 1,000-V, 2,400 r/min separately excited DC generator has an armature circuit resistance of 0.04 Ω. The flux per pole is 0.4 Wb. Find (a) the induced voltage, (b) the machine constant, and (c) the torque developed at the rated conditions.

Answer: (a) $E_b = 1{,}040$ V; (b) $k_a = 10.34 \frac{\text{V-s}}{\text{Wb-rad}}$; (c) $T = 4{,}138$ N-m

CHECK YOUR UNDERSTANDING

A 100-kW, 250-V shunt generator has a field circuit resistance of 50 Ω and an armature circuit resistance of 0.05 Ω. Find (a) the full-load line current flowing to the load, (b) the field current, (c) the armature current, and (d) the full-load generator voltage.

Answer: (a) 400 A; (b) 5 A; (c) 405 A; (d) 270.25 V

15.5 ALTERNATING-CURRENT MACHINES

AC machines represent the vast majority of industrial applications. The objective of this section is to explain the basic operation of both synchronous and induction machines and to outline their performance characteristics. In doing so, we also point out the relative advantages and disadvantages of these machines in comparison with DC machines.

Rotating Magnetic Fields

As mentioned in Section 15.1, the fundamental principle of operation of AC machines is the generation of a rotating magnetic field, which causes the rotor to turn at a speed that depends on the speed of rotation of the magnetic field. We now explain how a rotating magnetic field can be generated in the stator and air gap of an AC machine by means of alternating currents.

Consider the stator shown in Figure 15.23, which supports windings *a*-*a′*, *b*-*b′*, and *c*-*c′*. The coils are geometrically spaced 120° apart, and a three-phase voltage is applied to the coils. As you may recall from the discussion of AC power in Chapter 13, the currents generated by a three-phase source are also spaced by 120°, as illustrated in Figure 15.24. The phase voltages referenced to the neutral terminal would then be given by the expressions

$$v_a = A\cos(\omega_e t)$$
$$v_b = A\cos\left(\omega_e t - \frac{2\pi}{3}\right)$$
$$v_c = A\cos\left(\omega_e t + \frac{2\pi}{3}\right)$$

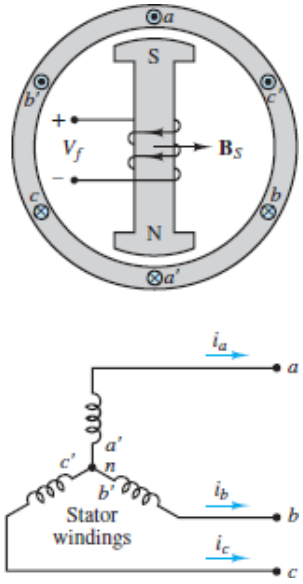


Figure 15.23 Two-pole three-phase stator

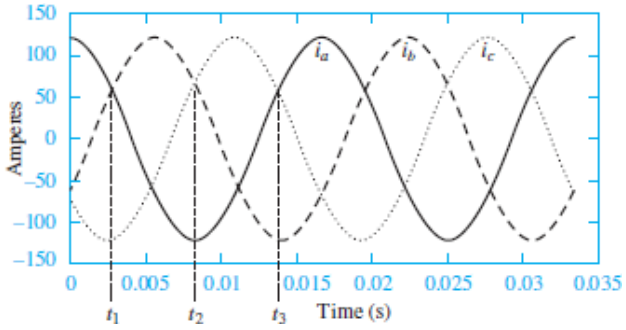


Figure 15.24 Three-phase stator winding currents

Page 900

where ω_e is the frequency of the AC supply, or line frequency. The coils in each winding are arranged in such a way that the flux distribution generated by any one winding is approximately sinusoidal. Such a flux distribution may be obtained by appropriately arranging groups of coils for each winding over the stator surface. Since the coils are spaced 120° apart, the flux distribution resulting from the sum of the contributions of the three windings is the sum of the fluxes

due to the separate windings, as shown in Figure 15.25. Thus, the flux in a three-phase machine rotates in space according to the vector diagram of Figure 15.26, and the flux is constant in amplitude. A stationary observer on the machine's stator would see a sinusoidally varying flux distribution, as shown in Figure 15.25.

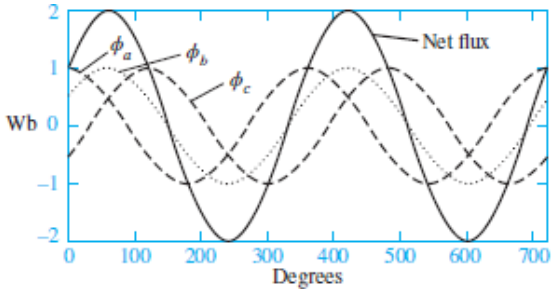


Figure 15.25 Flux distribution in a three-phase stator winding as a function of angle of rotation

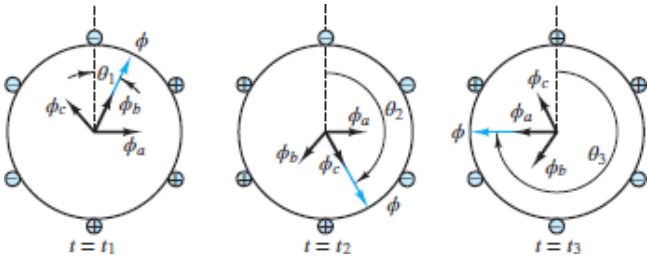


Figure 15.26 Rotating flux in a three-phase machine

Since the resultant flux of Figure 15.25 is generated by the currents of Figure 15.24, the speed of rotation of the flux must be related to the frequency of the sinusoidal phase currents. In the case of the stator of Figure 15.23, the number of magnetic poles resulting from the winding configuration is two; however, it is also possible to configure the windings so that they have more poles. For example, Figure 15.27 depicts a simplified view of a four-pole stator.

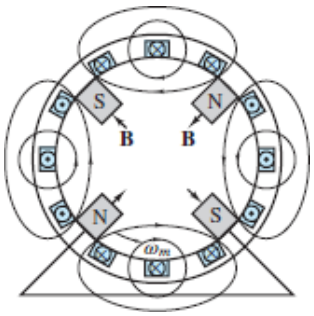


Figure 15.27 Four-pole stator

In general, the speed of the rotating magnetic field is determined by the frequency of the excitation current f and by the number of poles present in the stator p according to

$$n_s = \frac{120f}{p} \text{ r/min} \quad \text{Synchronous speed}$$

or

$$\omega_s = \frac{2\pi n_S}{60} = \frac{2\pi \times 2f}{p} \quad \text{Synchronous speed} \tag{15.41}$$

where n_s (or ω_s) is usually called the **synchronous speed**.

Now, the structure of the windings in the preceding discussion is the same whether the AC machine is a motor or a generator; the distinction between the two depends on the direction of power flow. In a generator, the electromagnetic torque is a reaction torque that opposes rotation of the machine; this is the torque against which the prime mover does work. In a motor, on the other hand, the rotational (motional) voltage generated in the armature opposes the applied voltage; this voltage is the counter- (or back) emf. Thus, the description of the rotating magnetic field given thus far applies to both motor and generator action in AC machines.

Page 901

As described a few paragraphs earlier, the stator magnetic field rotates in an AC machine, and therefore the rotor cannot "catch up" with the stator field and is in constant pursuit of it. The speed of rotation of the rotor will therefore depend on the number of magnetic poles present in the stator and in the rotor. The magnitude of the torque produced in the machine is a function of the angle γ between the stator and rotor magnetic fields; precise expressions for this torque depend on how the magnetic fields are generated and will be given separately for the two cases of synchronous and induction machines. What is common to all rotating machines is that the number of stator and rotor poles must be identical if any torque is to be generated. Further, the number of poles must be even, since for each north pole there must be a corresponding south pole.

One important desired feature in an electric machine is an ability to generate a constant electromagnetic torque. With a constant-torque machine, one can avoid torque pulsations that could lead to undesired mechanical vibration in the motor itself and in other mechanical components attached to the motor (e.g., mechanical loads, such as spindles or belt drives). A constant torque may not always be achieved although it will be shown that it is possible to accomplish this goal when the excitation currents are multiphase. A general rule of thumb, in this respect, is that it is desirable, insofar as possible, to produce a constant flux per pole.

15.6 THE ALTERNATOR (SYNCHRONOUS GENERATOR)

One of the most common AC machines is the **synchronous generator**, or **alternator**. In this machine, the field winding is on the rotor, and the connection is made by means of brushes, in an arrangement similar to that of the DC machines studied earlier. The rotor field is obtained by means of a direct current provided to the rotor winding, or by permanent magnets. The rotor is then connected to a mechanical source of power and rotates at a speed that we will consider constant to simplify the analysis.

Figure 15.28 depicts a two-pole three-phase synchronous machine. Figure 15.29 depicts a four-pole three-phase alternator, in which the rotor poles are generated Page 902by means of a wound salient pole configuration and the stator poles are the result of windings embedded in the stator according to the simplified arrangement shown in the figure, where each of the pairs a/a', b/b', and so on contributes to the generation of the magnetic poles, as follows. The group a/a', b/b', c/c' produces a sinusoidally distributed flux (see Figure 15.25) corresponding to one of the pole pairs, while the group $-a/-a'$, $-b/-b'$, $-c/-c'$ contributes the other pole pair. The connections of the coils making up the windings are also shown in Figure 15.29. Note that the coils form a wye connection (see Chapter 13). The resulting flux distribution is such that the flux completes two sinusoidal cycles around the circumference of the air gap. Note also that each arm of the three-phase wye connection has been divided into two coils, wound in different locations, according to the schematic stator diagram of Figure 15.29. One could then envision analogous configurations with greater numbers of poles, obtained in the same fashion, that is, by dividing each arm of a wye connection into more windings.

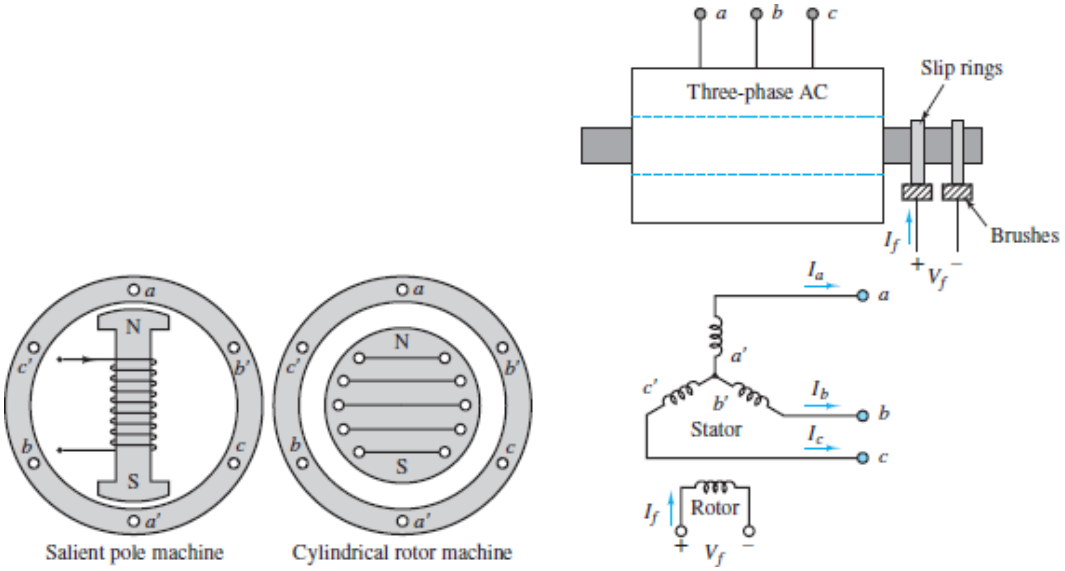


Figure 15.28 Two-pole synchronous machine

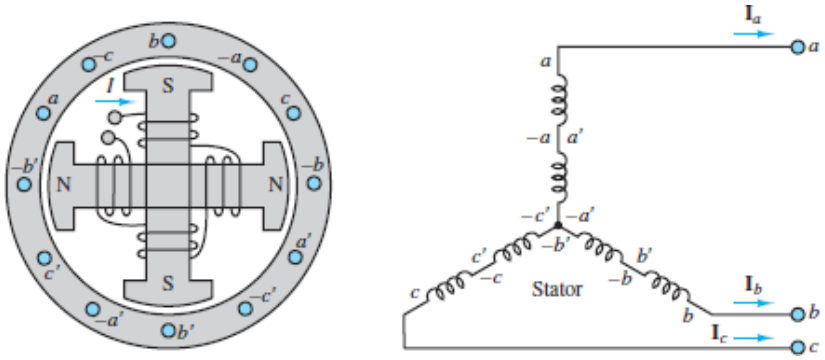


Figure 15.29 Four-pole three-phase alternator

The arrangement shown in Figure 15.29 requires that a further distinction be made between mechanical degrees θ_m and electrical degrees θ_e. In the four-pole alternator, the flux will see two complete cycles during one rotation of the rotor, and therefore the voltage that is generated in the coils will also oscillate at twice the frequency of rotation. In general, the electrical degrees (or radians) are related to the mechanical degrees by the expression:

$$\theta_e = \frac{p}{2}\theta_m \tag{15.42}$$

where p is the number of poles. In effect, the voltage across a coil of the machine goes through one cycle every time a pair of poles moves past the coil. Thus, the frequency of the voltage generated by a synchronous generator is

$$f = \frac{p}{2}\frac{n}{60} \quad \text{Hz} \tag{15.43}$$

where n is the mechanical speed in revolutions per minute. Alternatively, if the speed is expressed in radians per second, we have

$$\omega_e = \frac{p}{2}\omega_m \tag{15.44}$$

where ω_m is the mechanical speed of rotation in radians per second. The number of poles employed in a synchronous generator is then determined by two factors: the frequency desired of the generated voltage (e.g., 60 Hz, if the generator is used to produce AC power) and the speed of rotation of the prime mover. In the latter Page 903respect, there is a significant difference, for example, between the speed of rotation of a steam turbine generator and that of a hydroelectric generator, the former being much greater.

A common application of the alternator is seen in automotive battery-charging systems, in which, however, the generated AC voltage is rectified to provide the DC required for charging the battery. Figure 15.30 depicts an automotive alternator.

Figure 15.30 Automotive alternator (© *2012 Remy International, Inc. All rights reserved.*)

CHECK YOUR UNDERSTANDING

A synchronous generator has a multipolar construction that permits changing its synchronous speed. If only two poles are energized, at 50 Hz, the speed is 3,000 r/min. If the number of poles is progressively increased to 4, 6, 8, 10, and 12, find the synchronous speed for each configuration. Draw the complete equivalent circuit of a synchronous generator and its phasor diagram.

Answer: 1,500, 1,000, 750, 600, and 500 r/min

15.7 THE SYNCHRONOUS MOTOR

Synchronous motors are virtually identical to synchronous generators with regard to their construction, except for an additional winding for helping start the motor and minimizing motor speed over- and undershoots. The principle of operation is, of course, the opposite: An AC excitation provided to the armature generates a magnetic field in the air gap between stator and rotor, resulting in a mechanical torque. To generate the rotor magnetic field, some direct current must be provided to the field windings; this is often accomplished by means of an **exciter**, which consists of a small DC generator propelled by the motor itself, and therefore mechanically connected to it. It was mentioned earlier that to obtain a constant torque in an electric motor, it is necessary to keep the rotor and stator magnetic fields constant relative to each other. This means that the electromagnetically rotating field in the stator and the mechanically rotating rotor field should be aligned at all times. The only condition for which this is possible occurs if both fields are rotating at the synchronous speed $n_s = 120\ f/p$. Thus, synchronous motors are by their very nature constant-speed motors, if the excitation frequency is constant.

For a non–salient pole (cylindrical rotor) synchronous machine, the torque can be written in terms of the stator alternating current $i_S(t)$ and the rotor direct current, I_f:

$$T = k\, i_S(t) I_f \sin(\gamma) \qquad \text{Synchronous motor torque} \tag{15.45}$$

where γ is the angle between the stator and rotor fields (see Figure 15.7). Let the angular speed of rotation be

$$\omega_m = \frac{d\theta_m}{dt} \quad \text{rad/s} \tag{15.46}$$

where $\omega_m = 2\pi n/60$, and let ω_e be the electrical frequency of $i_S(t)$, where $i_S(t) = \sqrt{2}\, I_S \sin(\omega_e t)$. Then the torque may be expressed as

$$T = k\sqrt{2}\, I_S \sin(\omega_e t) I_f \sin(\gamma) \tag{15.47}$$

where k is a machine constant, I_S is the rms value of the stator current, and I_f is the rotor direct current. Now, the rotor angle γ can be expressed as a function of time by

$$\gamma = \gamma_0 + \omega_m t \tag{15.48}$$

where γ_0 is the angular position of the rotor at $t = 0$; the torque expression then becomes

$$\begin{aligned} T &= k\sqrt{2}\, I_S I_f \sin(\omega_e t) \sin(\omega_m t + \gamma_0) \\ &= k\frac{\sqrt{2}}{2} I_S I_f \cos[(\omega_m - \omega_e)t - \gamma_0] - \cos[(\omega_m + \omega_e)t + \gamma_0] \end{aligned} \tag{15.49}$$

It is a straightforward matter to show that the average value of this torque, denoted by $\langle T \rangle$, is different from zero only if $\omega_m = \pm\omega_e$, that is, only if the motor is turning at the synchronous speed. The resulting average torque is then given by

$$\langle T \rangle = k\sqrt{2}\, I_S I_f \cos(\gamma_0) \tag{15.50}$$

Note that equation 15.49 corresponds to the sum of an average torque plus a fluctuating component at twice the original electrical (or mechanical) frequency. The fluctuating component results because, in the foregoing derivation, a single-phase current was assumed. The use of multiphase currents reduces the torque fluctuation to zero and permits the generation of a constant torque.

A per-phase circuit model describing the synchronous motor is shown in Figure 15.31, where the rotor circuit is represented by a field winding equivalent resistance and inductance, R_f and L_f, respectively, and the stator circuit is represented by equivalent stator winding inductance and resistance, L_S and R_S, respectively, and by the induced emf E_b. From the exact equivalent circuit as given in Figure 15.31, we have

$$V_S = E_b + I_S(R_S + jX_S) \tag{15.51}$$

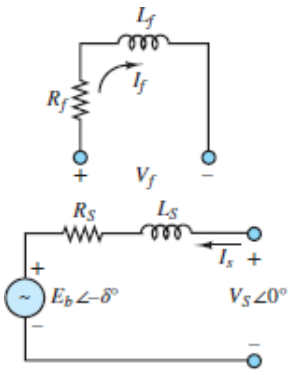


Figure 15.31 Per-phase circuit model

where X_S is known as the *synchronous reactance* and includes magnetizing reactance.

The motor power is

$$P_{\text{out}} = \omega_S T = |V_S||I_S|\cos(\theta) \tag{15.52}$$

for each phase, where T is the developed torque and θ is the angle between the stator voltage and current, V_S and I_S.

When the phase winding resistance R_S is neglected, the circuit model of a synchronous machine can be redrawn as shown in Figure 15.32. The input power (per phase) is equal to the output power in this circuit, since no power is dissipated in the circuit:

$$P_\phi = P_{\text{in}} = P_{\text{out}} = |\mathbf{V}_S||\mathbf{I}_S|\cos(\theta) \tag{15.53}$$

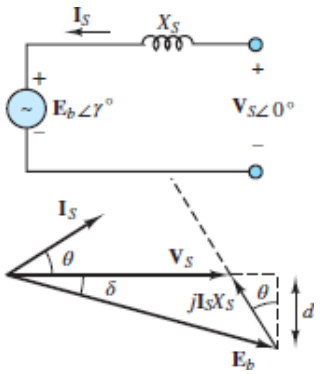


Figure 15.32 Per-phase circuit model of synchronous machines with winding resistance neglected.

Also by inspection of Figure 15.32, we have

$$d = |\mathbf{E}_b|\sin(\delta) = |\mathbf{I}_S|X_S\cos(\theta) \tag{15.54}$$

Then:

$$|\mathbf{E}_b||\mathbf{V}_S|\sin(\delta) = |\mathbf{V}_S||\mathbf{I}_S|X_s\cos(\theta) = X_S P_\phi \tag{15.55}$$

Page 905

The total power of a three-phase synchronous machine is then given by

$$P = 3\frac{|\mathbf{V}_S||\mathbf{E}_b|}{X_S}\sin(\delta) \tag{15.56}$$

Because of the dependence of the power upon the angle δ, this angle has come to be called the **power angle**. If δ is zero, the synchronous machine cannot develop useful power. The developed power has its maximum value at δ equal to 90°. If we assume that $|\mathbf{E}_b|$ and $|\mathbf{V}_S|$ are constant, we can draw the curve shown in Figure 15.33, relating the power and power angle in a synchronous machine.

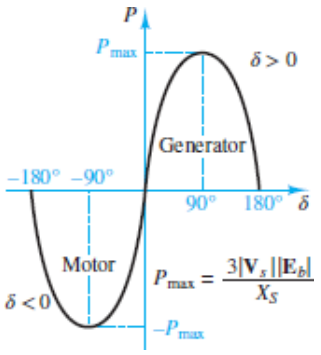


Figure 15.33 Power versus power angle for a synchronous machine

A synchronous generator is usually operated at a power angle varying from 15° to 25°. For synchronous motors and small loads, δ is close to 0°, and the motor torque is just sufficient to overcome its own windage and friction losses; as the load increases, the rotor field falls further out of phase with the stator field (although the two are still rotating at the same speed) until δ reaches a maximum at 90°. If the load torque exceeds the maximum torque, which is produced for δ =

90°, the motor is forced to slow down below synchronous speed. This condition is undesirable, and provisions are usually made to shut down the motor automatically whenever synchronism is lost. The maximum torque is called the **pull-out torque** and is an important measure of the performance of the synchronous motor.

Accounting for each of the phases, the total torque is given by

$$T = \frac{m}{\omega_s}|\mathbf{V}_S||\mathbf{I}_S|\cos(\theta) \tag{15.57}$$

where m is the number of phases. From Figure 15.32, we have $E_b \sin(\delta) = X_S I_S \cos(\theta)$. Therefore, for a three-phase machine, the developed torque is

$$T = \frac{P}{\omega_s} = \frac{3}{\omega_s}\frac{|\mathbf{V}_S||\mathbf{E}_b|}{X_S}\sin(\delta) \quad \text{N-m} \tag{15.58}$$

Typically, analysis of multiphase motors is performed on a per-phase basis, as illustrated in Examples 15.10 and 15.11.

EXAMPLE 15.10 Synchronous Motor Analysis

Problem

Find the kilovoltampere rating, the induced voltage, and the power angle of the rotor for a fully loaded synchronous motor.

Solution

Known Quantities: Motor ratings; motor synchronous impedance.

Find: S; $\mathbf{E}_b$; δ.

Schematics, Diagrams, Circuits, and Given Data: Motor ratings: 460 V; three-phase; power factor = 0.707 lagging; full-load stator current: 12.5 A. $Z_S = 1 + j12$ Ω.

Assumptions: Use per-phase analysis.

Analysis: The circuit model for the motor is shown in Figure 15.34. The per-phase current in the wye-connected stator winding is

$$I_S = |\mathbf{I}_S| = 12.5 \text{ A}$$

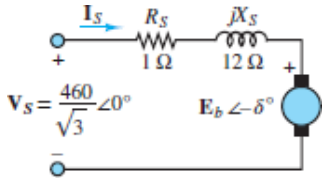


Figure 15.34 Circuit model of synchronous motor.

Page 906

The per-phase voltage is

$$V_S = |\mathbf{V}_S| = \frac{460 \text{ V}}{\sqrt{3}} = 265.58 \text{ V}$$

The kilovoltampere rating of the motor is expressed in terms of the apparent power S (see Chapter 13):

$$S = 3 V_S I_S = 3 \times 265.58 \text{ V} \times 12.5 \text{ A} = 9{,}959 \text{ W}$$

From the equivalent circuit, we have

$$\begin{aligned} \mathbf{E}_b &= \mathbf{V}_S - \mathbf{I}_S(R_S + jX_S) \\ &= 265.58 - (12.5 \angle -45° \text{ A}) \times (1 + j12 \ \Omega) = 179.31 \angle -32.83° \text{ V} \end{aligned}$$

The induced line voltage is defined to be

$$V_{\text{line}} = \sqrt{3} E_b = \sqrt{3} \times 179.31 \text{ V} = 310.57 \text{ V}$$

From the expression for $\mathbf{E}_b$, we can find the power angle:

$$\delta = -32.83°$$

Comments: The minus sign indicates that the machine is in the motor mode.

EXAMPLE 15.11 Synchronous Motor Analysis

Problem

Find the stator current, the line current, and the induced voltage for a synchronous motor, with reference to Figure 15.34, where $Z_S = R_S + jX_S$.

Solution

Known Quantities: Motor ratings; motor synchronous impedance.

Find: $\mathbf{I}_S$; $\mathbf{I}_{\text{line}}$; $\mathbf{E}_b$.

Schematics, Diagrams, Circuits, and Given Data: Motor ratings: 208 V; three-phase; 45 kVA; 60 Hz; power factor = 0.8 leading; $Z_S = 0 + j2.5\ \Omega$. Friction and windage losses: 1.5 kW; core losses: 1.0 kW; load power: 15 hp.

Assumptions: Use per-phase analysis.

Analysis: The output power of the motor is 15 hp; that is,

$$P_{\text{out}} = 15\ \text{hp} \times 0.746\ \text{kW/hp} = 11.19\ \text{kW}$$

The electric power supplied to the machine is

$$\begin{aligned} P_{\text{in}} &= P_{\text{out}} + P_{\text{mech}} + P_{\text{core loss}} + P_{\text{elec loss}} \\ &= 11.19\ \text{kW} + 1.5\ \text{kW} + 1.0\ \text{kW} + 0\ \text{kW} = 13.69\ \text{kW} \end{aligned}$$

As discussed in Chapter 13, the resulting line current is

$$I_{\text{line}} = \frac{P_{\text{in}}}{\sqrt{3}\,V\cos\theta} = \frac{13{,}690\ \text{W}}{\sqrt{3} \times 208\ \text{V} \times 0.8} = 47.5\ \text{A}$$

Page 907

Because of the delta connection, the armature current is

$$\mathbf{I}_S = \frac{1}{\sqrt{3}}\mathbf{I}_{\text{line}} = 27.4\angle 36.87^\circ\ \text{A}$$

The emf may be found from the equivalent circuit and KVL:

$$\begin{aligned} \mathbf{E}_b &= \mathbf{V}_S - jX_S\mathbf{I}_S \\ &= 208\angle 0^\circ - (j2.5\ \Omega)(27.4\angle 36.87^\circ\ \text{A}) = 255\angle -12.4^\circ\ \text{V} \end{aligned}$$

The power angle is

$$\delta = -12.4^\circ$$

CHECK YOUR UNDERSTANDING

Find an expression for the maximum pull-out torque of the synchronous motor.

Answer: $T_{max} = \frac{3V_S E_b}{\omega_m X_S}$

Synchronous motors are not very commonly used in practice, for various reasons, among which are that they are essentially required to operate at constant speed (unless a variable-frequency AC supply is available) and that they are not self-starting. Further, separate AC and DC supplies are required. It will be seen shortly that the induction motor overcomes most of these drawbacks.

15.8 THE INDUCTION MOTOR

The induction motor is the most widely used electric machine, because of its relative simplicity of construction. The stator winding of an induction machine is similar to that of a synchronous machine; thus, the description of the three-phase winding of Figure 15.23 also applies to induction machines. The primary advantage of the induction machine, which is almost exclusively used as a motor (its performance as a generator is not very good), is that no separate excitation is required for the rotor. The rotor typically consists of one of two arrangements: a **squirrel cage** or a **wound rotor**. The former contains conducting bars short-circuited at the end and embedded within it; the latter consists of a multiphase winding similar to that used for the stator, but electrically short-circuited.

In either case, the induction motor operates by virtue of currents induced from the stator field in the rotor. In this respect, its operation is similar to that of a transformer, in that currents in the stator (which acts as a primary coil) induce currents in the rotor (acting as a secondary coil). In most induction motors, no external electrical connection is required for the rotor, thus permitting a simple, rugged construction without the need for slip rings or brushes. Unlike the synchronous motor, the induction motor operates not at synchronous speed, but at a somewhat lower speed, which is dependent on the load. Figure 15.35 illustrates

the appearance of a squirrel cage induction motor. The following discussion focuses mainly on this very common configuration.

Page 908

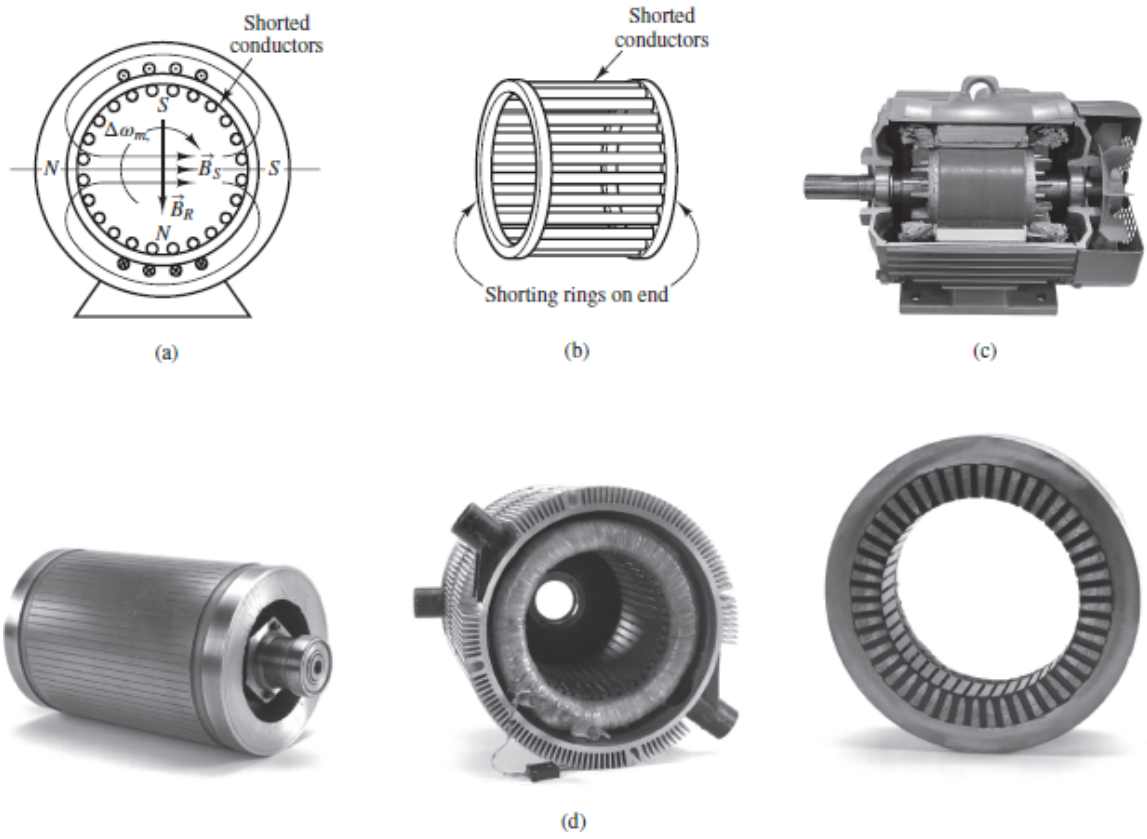


Figure 15.35 (a) Squirrel cage induction motor; (b) conductors in rotor; (c) photograph of squirrel cage induction motor; (d) views of Smokin' Buckeye motor: rotor, stator, and cross section of stator ((*c*) *Normal Life/Shutterstock;* (*d*) *Courtesy: David H. Koether Photography*)

By now you are acquainted with the notion of a rotating stator magnetic field. Imagine now that a squirrel cage rotor is inserted in a stator in which such a rotating magnetic field is present. The stator field will induce voltages in the cage conductors, and if the stator field is generated by a three-phase source, the resulting rotor currents—which circulate in the bars of the squirrel cage, with the conducting path completed by the shorting rings at the end of the cage—are also three-phase and are determined by the magnitude of the induced voltages and by

the impedance of the rotor. Since the rotor currents are induced by the stator field, the number of poles and the speed of rotation of the induced magnetic field are the same as those of the stator field, *if the rotor is at rest.* Thus, when a stator field is initially applied, the rotor field is synchronous with it, and the fields are stationary with respect to one another. Thus, according to the earlier discussion, a *starting torque* is generated.

If the starting torque is sufficient to cause the rotor to start spinning, the rotor will accelerate up to its operating speed. However, an induction motor can never reach synchronous speed; if it did, the rotor would appear to be stationary with respect to the rotating stator field, since it would be rotating at the same speed. But in the absence of relative motion between the stator and rotor fields, no voltage would be induced in the rotor. Thus, an induction motor is limited to speeds somewhere below the synchronous speed n_s. Let the speed of rotation of the rotor be n; then the rotor is losing ground with respect to the rotation of the stator field Page 909at a speed $n_s - n$. In effect, this is equivalent to backward motion of the rotor at the **slip speed**, defined by $n_s - n$. The **slip** s is usually defined as a fraction of n_s:

$$s = \frac{n_s - n}{n_s} \qquad \text{Slip in induction machine} \tag{15.59}$$

which leads to the following expression for the rotor speed:

$$n = n_s(1 - s) \tag{15.60}$$

The slip s is a function of the load, and the amount of slip in a given motor is dependent on its construction and rotor type (squirrel cage or wound rotor). Since there is a relative motion between the stator and rotor fields, voltages will be induced in the rotor at a frequency called the **slip frequency**, $f_R = sf$, where f is the frequency of the sinusoidal excitation related to the relative speed of the two fields. This gives rise to an interesting phenomenon: The rotor field travels relative to the rotor at the slip speed sn_s, but the rotor is mechanically traveling at the speed $(1 - s)n_s$, so that the net effect is that the rotor field travels at the speed:

$$sn_s + (1 - s)n_s = n_s \tag{15.61}$$

that is, at synchronous speed. The fact that the rotor field rotates at synchronous speed—although the rotor itself does not—is extremely important because it means that the stator and rotor fields will continue to be stationary with respect to each other, and therefore a net torque can be produced.

As in the case of DC and synchronous motors, important characteristics of induction motors are the starting torque, the maximum torque, and the torque–speed curve. These will be discussed shortly, after some analysis of the induction motor is performed.

EXAMPLE 15.12 Induction Motor Analysis

Problem

Find the full-load rotor slip and frequency of the induced voltage at rated speed in a four-pole induction motor.

Solution

Known Quantities: Motor ratings.

Find: s; f_R.

Schematics, Diagrams, Circuits, and Given Data: Motor ratings: 230 V; 60 Hz; full-load speed: 1,725 r/min.

Analysis: The synchronous speed of the motor is

$$n_s = \frac{120f}{p} = \frac{60f}{p/2} = \frac{60 \text{ s/min} \times 60 \text{ r/s}}{4/2} = 1{,}800 \text{ r/min}$$

The slip is

$$s = \frac{n_s - n}{n_s} = \frac{1{,}800 \text{ r/min} - 1{,}725 \text{ r/min}}{1{,}800 \text{ r/min}} = 0.0417$$

The rotor frequency f_R is

$$f_R = sf = 0.0417 \times 60 \text{ Hz} = 2.5 \text{ Hz}$$

CHECK YOUR UNDERSTANDING

A three-phase induction motor has six poles. (a) If the line frequency is 60 Hz, calculate the speed of the magnetic field in revolutions per minute. (b) Repeat the calculation if the frequency is changed to 50 Hz.

Answer: (a) $n = 1{,}200$ r/min; (b) $n = 1{,}000$ r/min

The induction motor can be described by means of an equivalent circuit, which is essentially that of a rotating transformer. (See Chapter 13 for a circuit model of the transformer.) Figure 15.36 depicts such a circuit model, where:

R_S = stator resistance per phase R_R = rotor resistance per phase
X_S = stator reactance per phase X_R = rotor reactance per phase
X_m = magnetizing (mutual) reactance
R_C = equivalent core-loss resistance
E_S = per-phase induced voltage in stator windings
E_R = per-phase induced voltage in rotor windings

The primary internal stator voltage $\mathbf{E}_S$ is coupled to the secondary rotor voltage $\mathbf{E}_R$ by an ideal transformer with an effective turns ratio of α. For the rotor circuit, the induced voltage at any slip will be

$$\mathbf{E}_R = s\mathbf{E}_{R0} \tag{15.62}$$

where $\mathbf{E}_{R0}$ is the induced rotor voltage at the condition in which the rotor is stationary. Also, $X_R = \omega_R L_R = 2\pi f_R L_R = 2\pi s f L_R = sX_{R0}$, where $X_{R0} = 2\pi f L_R$ is the reactance when the rotor is stationary. The rotor current is given by

$$\mathbf{I}_R = \frac{\mathbf{E}_R}{R_R + jX_R} = \frac{s\mathbf{E}_{R0}}{R_R + jsX_{R0}} = \frac{\mathbf{E}_{R0}}{R_R/s + jX_{R0}} \tag{15.63}$$

The resulting rotor equivalent circuit is shown in Figure 15.37.

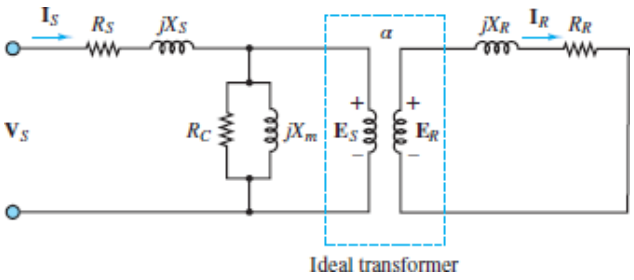


Figure 15.36 Circuit model for induction machine

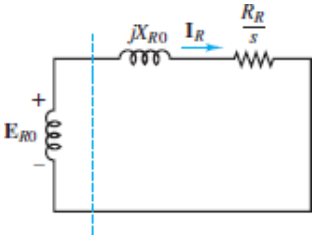


Figure 15.37 Rotor circuit

Page 911

The voltages, currents, and impedances on the secondary (rotor) side can be reflected to the primary (stator) by means of the effective turns ratio. When this transformation is effected, the transformed rotor voltage is given by

$$\mathbf{E}_2 = \mathbf{E}_R' = \alpha \mathbf{E}_{R0} \tag{15.64}$$

The transformed (reflected) rotor current is

$$\mathbf{I}_2 = \frac{\mathbf{I}_R}{\alpha} \tag{15.65}$$

The transformed rotor resistance can be defined as

$$R_2 = \alpha^2 R_R \tag{15.66}$$

and the transformed rotor reactance can be defined by

$$X_2 = \alpha^2 X_{R0} \tag{15.67}$$

The final per-phase equivalent circuit of the induction motor is shown in Figure 15.38.

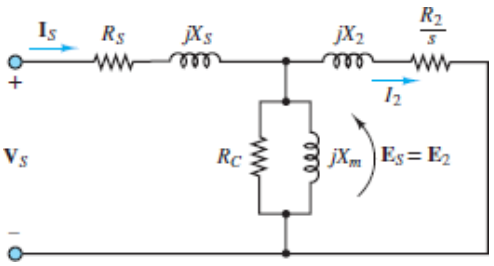


Figure 15.38 Equivalent circuit of an induction machine

Examples 15.13 and 15.14 illustrate the use of the circuit model in determining the performance of the induction motor.

EXAMPLE 15.13 Induction Motor Analysis

Problem

Determine the following quantities for an induction motor, using the circuit model of Figures 15.36 to 15.38.

1. Speed
2. Stator current
3. Power factor
4. Output torque

Solution

Known Quantities: Motor ratings; circuit parameters.

Find: n; ω_m; $\mathbf{I}_S$; power factor (pf); T.

Page 912

Schematics, Diagrams, Circuits, and Given Data: Motor ratings: 460 V; 60 Hz; four poles; $s = 0.022$; $P_{\text{out}} = 14$ hp; $R_S = 0.641\ \Omega$; $R_2 = 0.332\ \Omega$; $X_S = 1.106\ \Omega$; $X_2 = 0.464\ \Omega$; $X_m = 26.3\ \Omega$

Assumptions: Use per-phase analysis. Neglect core losses ($R_C = 0$).

Analysis:

1. The per-phase equivalent circuit is shown in Figure 15.38. The synchronous speed is found to be

$$n_s = \frac{120f}{p} = \frac{60 \text{ s/min} \times 60 \text{ r/s}}{4/2} = 1{,}800 \text{ r/min}$$

or

$$\omega_s = 1{,}800 \frac{\text{r}}{\text{min}} \times \frac{2\pi \text{ rad}}{60 \text{ s/min}} = 188.5 \text{ rad/s}$$

The rotor mechanical speed is

$$n = (1 - s)n_s = 1{,}760 \text{ r/min}$$

or

$$\omega_m = (1 - s)\omega_s = 184.4 \text{ rad/s}$$

2. The reflected rotor impedance is found from the parameters of the per-phase circuit to be

$$Z_2 = \frac{R_2}{s} + jX_2 = \frac{0.332}{0.022} + j0.464 \ \Omega$$
$$= 15.09 + j0.464 \ \Omega$$

The combined magnetization plus rotor impedance is therefore equal to

$$Z = \frac{1}{1/jX_m + 1/Z_2} = \frac{1}{-j0.038 + 0.0662\angle -1.76^\circ} = 12.93\angle 31.2^\circ \ \Omega$$

and the total impedance is

$$Z_{\text{total}} = Z_S + Z = 0.641 + j1.106 + 11.06 + j6.69$$
$$= 11.70 + j7.8 = 14.06\angle 33.7^\circ \ \Omega$$

Finally, the stator current is given by

$$\mathbf{I}_S = \frac{\mathbf{V}_S}{Z_{\text{total}}} = \frac{460/\sqrt{3}\angle 0^\circ \text{ V}}{14.07\angle 33.6^\circ \ \Omega} = 18.88\angle -33.7^\circ \text{ A}$$

3. The power factor is

$$\text{pf} = \cos 33.6^\circ = 0.832 \text{ lagging}$$

4. The output power P_{out} is

$$P_{\text{out}} = 14 \text{ hp} \times 746 \text{ W/hp} = 10.444 \text{ kW}$$

and the output torque is

$$T = \frac{P_{\text{out}}}{\omega_m} = \frac{10{,}444 \text{ W}}{184.4 \text{ rad/s}} = 56.64 \text{ N-m}$$

Page 913

CHECK YOUR UNDERSTANDING

A four-pole induction motor operating at a frequency of 60 Hz has a full-load slip of 4 percent. Find the frequency of the voltage induced in the rotor (a) at the instant of starting and (b) at full load.

Answer: (a) $f_R = 60$ Hz; (b) $f_R = 2.4$ Hz

LO

EXAMPLE 15.14 Induction Motor Analysis

Problem

Determine the following quantities for a three-phase induction motor, using the circuit model of Figure 15.38.

1. Stator current
2. Power factor
3. Full-load electromagnetic torque

Solution

Known Quantities: Motor ratings; circuit parameters.

Find: $\mathbf{I}_S$; pf; T.

Schematics, Diagrams, Circuits, and Given Data: Motor ratings: 500 V; three-phase; 50 Hz; $p = 8$; $s = 0.05$; $P = 14$ hp.

Circuit parameters: $R_S = 0.13\ \Omega$; $R'_R = 0.32\ \Omega$; $X_S = 0.6\ \Omega$; $X'_R = 1.48\ \Omega$; $Y_m = G_C + jB_m =$ magnetic branch admittance describing core loss and mutual inductance = $0.004 - j0.05\ \Omega^{-1}$; stator/rotor turns ratio = $1{:}\alpha = 1{:}1.57$.

Assumptions: Use per-phase analysis. Neglect mechanical losses.

Analysis: The approximate equivalent circuit of the three-phase induction motor on a perphase basis is shown in Figure 15.39. The parameters of the model are calculated as follows:

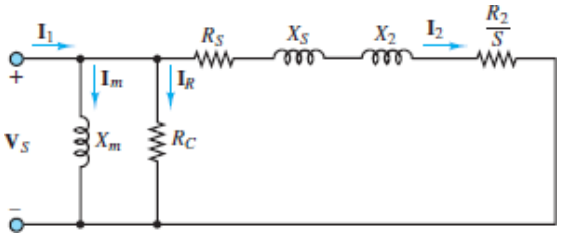


Figure 15.39 Per-phase equivalent circuit of induction machine.

Page 914

$$R_2 = R'_R \times \left(\frac{1}{\alpha}\right)^2 = 0.32 \times \left(\frac{1}{1.57}\right)^2 = 0.13\ \Omega$$

$$X_2 = X'_R \times \left(\frac{1}{\alpha}\right)^2 = 1.48 \times \left(\frac{1}{1.57}\right)^2 = 0.6\ \Omega$$

$$Z = R_S + \frac{R_2}{s} + j(X_S + X_2)$$

$$= 0.13 + \frac{0.13}{0.05} + j(0.6 + 0.6) = 2.73 + j1.2\ \Omega$$

Using the approximate circuit, we have

$$\mathbf{I}_2 = \frac{\mathbf{V}_S}{Z} = \frac{(500/\sqrt{3})\angle 0^\circ \text{ V}}{2.73 + j1.2\ \Omega} = 88.6 - 38.9 \text{ A}$$

$$\mathbf{I}_R = \mathbf{V}_S G_C = 288.7 \text{ V} \times 0.004\ \Omega^{-1} = 1.15 \text{ A}$$

$$\mathbf{I}_m = -j\mathbf{V}_S B_m = 288.7 \text{ V} \times (-j0.05)\ \Omega = -j14.4 \text{ A}$$

$$\mathbf{I}_1 = \mathbf{I}_2 + \mathbf{I}_R + \mathbf{I}_m = 89.75 - j53.3 \text{ A}$$

$$\text{Input power factor} = \frac{\text{Re}[\mathbf{I}_1]}{|\mathbf{I}_1|} = \frac{89.95}{104.6} = 0.86 \text{ lagging}$$

$$\text{Torque} = \frac{3P}{\omega_S} = \frac{3I_2^2 R_2/s}{4\pi f/p} = 931 \text{ N-m}$$

CHECK YOUR UNDERSTANDING

A four-pole, 1,746 r/min, 220-V, three-phase, 60-Hz, 10-hp, Y-connected induction machine has the following parameters: $R_S = 0.4\ \Omega$, $R_2 = 0.14\ \Omega$, $X_m = 16\ \Omega$, $X_S = 0.35\ \Omega$, $X_2 = 0.35\ \Omega$, $R_C = \infty$. Using Figure 15.38 find (a) the stator current, (b) the rotor current, (c) the motor power factor, and (d) the total stator power input.

Answer: (a) $25.92\angle -22.43^\circ$ A; (b) $24.35\angle -6.51^\circ$ A; (c) 0.9243; (d) 9,129 W

Performance of Induction Motors

The performance of induction motors can be described by torque–speed curves similar to those already used for DC motors. Figure 15.40 depicts an induction motor torque–speed curve, with five torque ratings marked *a* through *e*. Point *a* is the *starting torque*, also called **breakaway torque**, and is the torque available with the rotor "locked," that is, in a stationary position. At this condition, the frequency of the voltage induced in the rotor is highest, since it is equal to the frequency of rotation of the stator field; consequently, the inductive reactance of the rotor is greatest. As the rotor accelerates, the torque drops off, reaching a minimum value called the **pull-up torque** (point *b*); this typically occurs somewhere between 25 and 40 percent of synchronous speed. As the rotor speed continues to increase, the rotor reactance decreases further (since the frequency of the induced voltage is determined by the relative speed of rotation of the rotor

with respect to the stator field). The torque becomes a maximum when the rotor inductive reactance is equal Page 915to the rotor resistance; maximum torque is also called **breakdown torque** (point *c*). Beyond this point, the torque drops off until it is zero at synchronous speed, as discussed earlier. Also marked on the curve are the *150 percent torque* (point *d*) and the *rated torque* (point *e*).

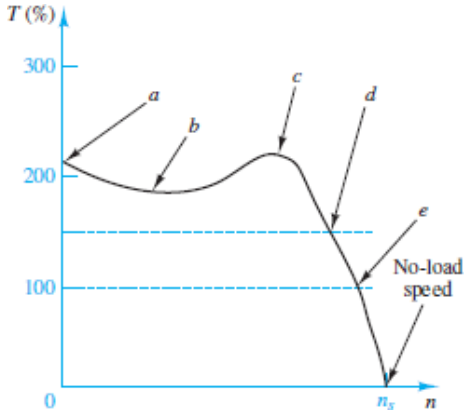


Figure 15.40 Performance curve for induction motor

A general formula for the computation of the induction motor steady-state torque–speed characteristic is

$$T = \frac{1}{\omega_e} \frac{mV_S^2 R_R/s}{(R_S + R_R/s)^2 + (X_S + X_R)^2} \qquad (15.68)$$

T-ω equation for induction machine

where *m* is the number of phases.

Different construction arrangements permit the design of induction motors with different torque–speed curves, thus permitting the user to select the motor that best suits a given application. Figure 15.41 depicts the four basic classifications—classes A, B, C, and D—as defined by NEMA. The determining features in the classification are the locked-rotor torque and current, the breakdown torque, the pull-up torque, and the percentage of slip. Class A motors have a higher breakdown torque than class B motors, and a slip of 5 percent or less. Motors in this class are often designed for a specific application. Class B motors are general-purpose motors; this is the most commonly used type of induction motor, with typical values of slip of 3 to 5 percent. Class C motors have a high starting torque for a given starting current, and a low slip. These motors

are typically used in applications demanding high starting torque but having relatively normal running loads, once the running speed has been reached. Class D motors are characterized by high starting torque, high slip, low starting current, and low full-load speed. A typical value of slip is around 13 percent.

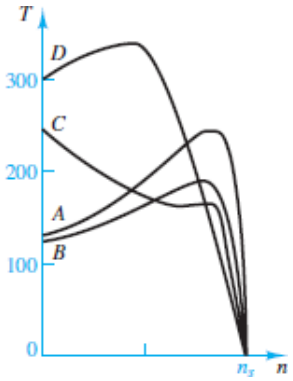


Figure 15.41 Induction motor classification

Factors that should be considered in the selection of an AC motor for a given application are the *speed range*, both minimum and maximum, and the speed variation. For example, it is important to determine whether constant speed is required; what variation might be allowed, either in speed or in torque; or whether variable-speed operation is required, in which case a variable-speed drive will be needed. Page 916The torque requirements are obviously important as well. The starting and running torque should be considered; they depend on the type of load. Starting torque can vary from a small percentage of full-load torque to several times full-load torque. Furthermore, the excess torque available at start-up determines the *acceleration characteristics* of the motor. Similarly, *deceleration characteristics* should be considered, to determine whether external braking might be required.

Another factor to be considered is the *duty cycle* of the motor. The duty cycle, which depends on the nature of the application, is an important consideration when the motor is used in repetitive, noncontinuous operation, such as is encountered in some types of machine tools. If the motor operates at zero or reduced load for periods of time, the duty cycle—that is, the percentage of the time the motor is loaded—is an important selection criterion. Last, but by no means least, are the *thermal properties* of a motor. Motor temperature is determined by internal losses and by ventilation; motors operating at a reduced speed may not generate sufficient cooling, and forced ventilation may be required.

Thus far, we have not considered the dynamic characteristics of induction motors. Among the integral-horsepower induction motors (i.e., motors with horsepower rating greater than 1), the most common dynamic problems are associated with starting and stopping and with the ability of the motor to continue operation during supply system transient disturbances. Dynamic analysis methods for induction motors depend to a considerable extent on the nature and complexity of the problem and the associated precision requirements. When the electric transients in the motor are to be included as well as the motion transients, and especially when the motor is an important element in a large network, the simple transient equivalent circuit of Figure 15.42 provides a good starting approximation. There, X'_S is called the *transient reactance.* The voltage E'_S is called the *voltage behind the transient reactance* and is assumed to be equal to the initial value of the induced voltage, at the start of the transient. The stator resistance is R_S. The dynamic analysis problem consists of selecting a sufficiently simple but reasonably realistic representation that will not unduly complicate the dynamic analysis, particularly through the introduction of nonlinearities.

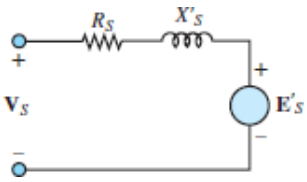


Figure 15.42 Simplified induction motor dynamic model

It should be remarked that the basic equations of the induction machine, as derived from first principles, are quite nonlinear. Thus, an accurate dynamic analysis of the induction motor, without any linearizing approximations, requires the use of computer simulation.

AC Motor Speed and Torque Control

As explained in an earlier section, AC machines are constrained to fixed-speed or near fixed-speed operation when supplied by a constant-frequency source. Several simple methods exist to provide limited speed control in AC induction machines; more complex methods, involving the use of advanced power electronics circuits, can be used if the intended application requires wide-bandwidth control of motor speed or torque. In this subsection we provide a general overview of available solutions.

Pole Number Control

The (conceptually) easiest method to implement speed control in an induction machine is by *varying the number of poles.* Equation 15.41 explains the dependence Page 917of synchronous speed in an AC machine on the supply frequency and on the number of poles. For machines operated at 60 Hz, the following speeds can be achieved by varying the number of magnetic poles in the stator winding:

Number of poles	2	4	6	8	12
n (r/min)	3,600	1,800	1,200	800	600

While for machines operating at 50 Hz, the speeds are

Number of poles	2	4	6	8	12
n (r/min)	3,000	1,500	1,000	667	500

Motor stators can be wound so that the number of pole pairs in the stators can be varied by switching between possible winding connections. Such switching requires that care be taken in timing it to avoid damage to the machine.

Slip Control

Since the rotor speed is inherently dependent on the slip, *slip control* is a valid means of achieving some speed variation in an induction machine. Since motor torque falls with the square of the voltage (see equation 15.68), it is possible to change the slip by changing the motor torque through a change in motor voltage. This procedure allows for speed control over the range of speeds that allow for stable motor operation. With reference to Figure 15.40, this is possible only above point *c*, that is, above the *breakdown torque.*

Rotor Control

For motors with wound rotors, it is possible to connect the rotor slip rings to resistors; adding resistance to the rotor increases the losses in the rotor and therefore causes the rotor speed to decrease. This method is also limited to operation above the *breakdown torque* although it should be noted that the shape of the motor torque–speed characteristic changes when the rotor resistance is changed.

Frequency Regulation

The last two methods cause additional losses to be introduced in the machine. If a variable-frequency supply is used, motor speed can be controlled without any additional losses. As seen in equation 15.41, the motor speed is directly

dependent on the supply frequency, as the supply frequency determines the speed of the rotating magnetic field. However, to maintain the same motor torque characteristics over a range of speeds, the motor voltage must change with frequency, to maintain a constant torque. Thus, generally, the volts/hertz ratio should be held constant. This condition is difficult to achieve at start-up and at very low frequencies, in which cases the voltage must be raised above the constant volts/hertz ratio that will be appropriate at higher frequency.

Page 918

15.9 ELECTRIC MOTOR DRIVES

High-power semiconductor devices make it possible to design effective and relatively low-cost electronic supplies that take full advantage of the device capabilities. Electronic power supplies for DC and AC motors is one of the major fields of application of power electronic devices. This section introduces two families of power supplies, or electric drives: **choppers**, or **DC-DC converters**; and **inverters**, or **DC-AC converters**. These families find widespread use in the control of AC and DC motors in a variety of applications and power ranges.

Depending on the relationship between the voltage across and the current through a load, an electronic drive can operate in one of four possible modes, as indicated in Figure 15.43. It is important to notice that in quadrants I and III power is *supplied by the drive* and *absorbed by the load.* By contrast, in quadrants II and IV power is *absorbed by the drive* and *supplied by the load.*

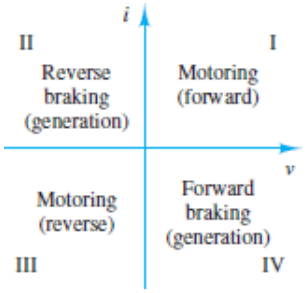


Figure 15.43 The four quadrants of an electric drive

DC-DC Converters

As the name suggests, a DC-DC converter is capable of converting a fixed DC supply to a variable DC supply. This feature is particularly useful for speed control of a DC motor, which is depicted in Figure 15.44. The torque T_m is

proportional to the current I_a supplied to the motor **armature**, while the speed of rotation of the motor ω_m is proportional to the voltage E_a (emf) across the armature. A DC motor is an example of an electromechanical energy conversion system, which converts electrical energy to mechanical energy (or vice versa when used as a generator). In the mechanical domain, the product of torque and speed is power; in the electrical domain, the product of current and voltage is power. Thus, in the ideal case of lossless energy conversion:

$$E_a \times I_a = T_m \times \omega_m \tag{15.69}$$

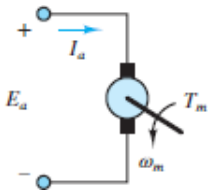


Figure 15.44 DC motor

Of course, no energy conversion process is lossless; however, there is a correspondence between the four electrical quadrants of Figure 15.43 and the mechanical power output of the motor. In particular, the mechanical power of a DC motor will have the same sign as the electrical power of its drive. That is, when the drive is supplying power, the DC motor will be doing work on its load, as in the **forward** and **reverse motoring** modes. When the drive is absorbing power, the DC motor will be having work done on it by its load, as in the forward and reverse **regenerative braking** modes.

A simple circuit that can accomplish the task of providing a variable DC supply from a fixed DC source is the **buck converter (step-down chopper)**, shown in Figure 15.45. The circuit consists of a switch, denoted by the symbol *S*, and a snubber diode. The switch can be any power switch, for example, a power BJT or MOSFET. The circuit to the right of the diode is a model of a DC motor, including the inductance and resistance of the armature windings, and the effect of the back-emf E_a. When the switch is turned on (say, at $t = 0$), the supply V_S is connected to the load and $v_o = V_S$. The load current i_o is determined by the motor parameters. When the switch is turned off, the load current continues to flow through the snubber diode, but the output voltage is now $v_o = 0$. At time T, the switch is turned on again, and the cycle repeats.

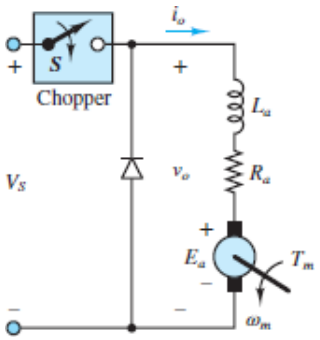


Figure 15.45 Buck converter (step-down chopper)

Page 919

Figure 15.46 depicts the v_o and i_o waveforms. The average value of the output voltage $\langle vo \rangle$ is given by:

$$\langle v_o \rangle = \frac{t_1}{T} V_S = \delta V_S \tag{15.70}$$

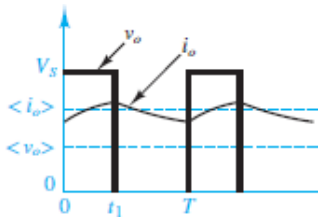


Figure 15.46 Step-down chopper waveforms

where δ is the **duty cycle** of the chopper. The step-down chopper has the useful range:

$$0 \le \langle v_o \rangle \le V_S \tag{15.71}$$

It is also possible to expand the range of a DC-DC converter to above the supply voltage by using the energy storage properties of an inductor; the resulting circuit is shown in Figure 15.47. When the chopper switch S is closed, the supply current is through the inductor L_o and the switch, so energy is stored in the inductor; the output voltage v_o is zero since the switch is a short-circuit. When the switch is open, the supply current is through the diode and the load. However, the inductor voltage is negative during the transient following the opening of the switch and therefore adds to the source voltage: the energy stored in the inductor while the switch was closed is now released and transferred to the load. This

stored energy makes it possible for the output voltage to be higher than the supply voltage for a finite period.

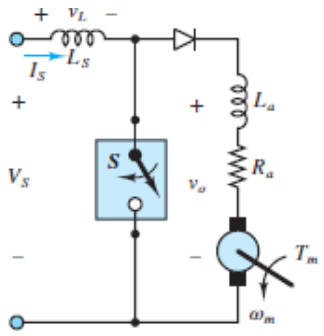


Figure 15.47 Boost converter (step-up chopper)

To maintain a constant average load current, the current increase between 0 and t_1 must equal the current decrease from t_1 to T. Therefore:

$$\frac{1}{L}\int_0^{t_1} V_S\,dt = \frac{1}{L}\int_{t_1}^{T} (\langle v_o\rangle - V_S)\,dt \tag{15.72}$$

from which the following relationship can be found:

$$V_S t_1 = (\langle v_o\rangle - V_S)(T - t_1) \tag{15.73}$$

This result can be rewritten to express the average output voltage as:

$$\langle v_o\rangle = \frac{T}{T - t_1}V_S = \frac{1}{1 - t_1/T}V_S = \frac{1}{1-\delta}V_s \geq V_S \tag{15.74}$$

Since the duty cycle δ is always less than 1, the theoretical range of the supply is:

$$V_S \leq \langle v_o\rangle < \infty \tag{15.75}$$

The waveforms for the boost converter are shown in Figure 15.48.

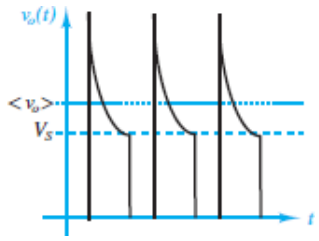


Figure 15.48 Boost converter output voltage waveform (ideal)

A **boost converter** (**step-up chopper**) can also be used for regenerative braking, where the motor armature voltage is used as the "supply" voltage and the output voltage is that across a DC battery; then power will flow from the motor and recharge the battery. This configuration is shown in Figure 15.49.

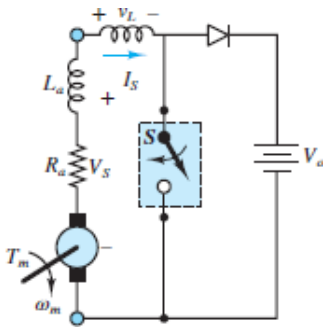


Figure 15.49 Boost converter used for regenerative braking in an electric vehicle.

The operations of the buck converter and boost converter can be combined to form a **buck-boost converter**, shown in Figure 15.50. This same circuit can act as a **two-quadrant chopper** to provide both regenerative braking and forward motoring operation in a DC motor. When switch S_2 is open, switch S_1 can serve as a chopper, and the circuit operates as a buck converter. In this mode, the drive and motor operate in quadrant I. The output voltage v_o will switch between V_S and zero, as shown in Figure 15.46, and the current i_0 is directed toward the motor, as indicated in Figure 15.50.

Page 920

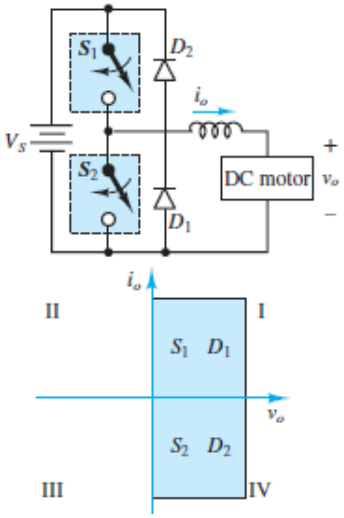


Figure 15.50 Two-quadrant DC-DC converter

When switch S_1 is open, switch S_2 can serve as a chopper, and the circuit operates as a boost converter. The source is the motor emf E_a and the load is the battery, as depicted in Figure 15.49. The current i_0 is directed away from the motor (the sum of the motor emf and the voltage across the inductor is greater than the battery voltage), and the drive operates in quadrant IV.

Inverters (DC-AC Converters)

Variable-speed drives for AC motors require a multiphase variable-frequency, variable-voltage supply. Such drives are called *DC-AC converters*, or *inverters*. Inverter circuits can be quite complex; only a brief introduction illustrating the basic principles is presented here.

A **voltage source inverter (VSI)** converts the output of a fixed DC supply (e.g., a battery) to a variable-frequency AC supply. Figure 15.51 depicts a **half-bridge VSI**; once again, the switches can be bipolar or MOS transistors, or thyristors. When switch S_1 is closed, the output voltage is in the positive half-cycle, and $v_o = V_S/2$. When switch S_2 is closed, $v_o = -V_S/2$. A switching sequence for S_1 and S_2 is shown in Figure 15.52. It is important that each switch be turned off before the other is turned on; otherwise, the DC supply will be short-circuited. Since a motor drive load is inductive, the load current i_o will lag the voltage, as shown in Figure 15.52. As a result, there will be periods when the voltage is positive but the current is negative. During these periods the diodes D_1 and D_2 conduct the load current, which would otherwise be forced to zero. Figure 15.52 indicates which diode is conducting in each of these periods.

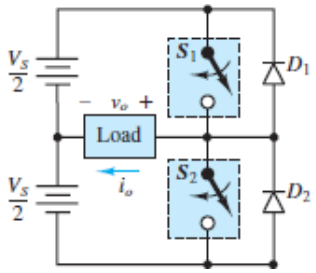


Figure 15.51 Half-bridge voltage source inverter

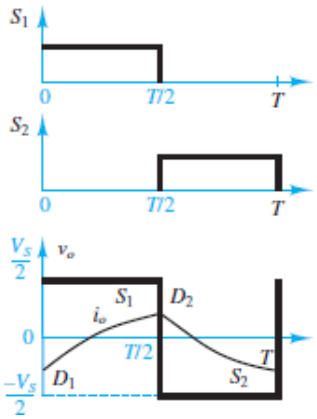


Figure 15.52 Half-bridge voltage source inverter waveforms

A full-bridge version of the VSI can also be designed as shown in Figure 15.53; the associated output voltage waveform is shown in Figure 15.54. The operation of this circuit is analogous to that of the half-bridge VSI; switches S_1 and S_2 are fired during the first half-cycle, and switches S_3 and S_4 during the second half. Note that the full-bridge configuration allows the output voltage to swing from V_S to $-V_S$. The diodes provide a path for the load current whenever the load voltage and current are of opposite polarity.

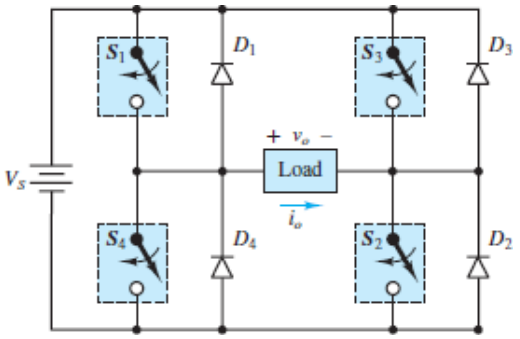


Figure 15.53 Full-bridge voltage source inverter

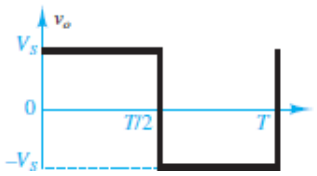


Figure 15.54 Half-bridge voltage source inverter output waveform

A three-phase version of the VSI and its related waveforms are shown in Figures 15.55 and 15.56. Once again, the operation is analogous to that of the previous VSI circuits. The top three waveforms depict the **pole voltages**, which are referenced to the DC supply neutral point *o*. The pole voltages are obtained by firing switches S_1 through S_6 at appropriate times. For example, if S_1 is fired at $\omega t = 0$, then pole *a* is connected to the positive side of the DC supply and $v_{ao} = V_S/2$; Page 921if S_4 is subsequently turned on at $\omega t = \pi$, then pole *a* is connected to the negative side of the DC supply and $v_{ao} = -V_S/2$. The other pairs of switches are then fired in an analogous sequence, shifted by 120 electrical degrees with respect to each other, to obtain the waveforms shown in the top three graphs of Figure 15.56. The **line voltages** are obtained from the pole voltages by using the relations

$$\begin{aligned} v_{ab} &= v_{ao} - v_{bo} \\ v_{bc} &= v_{bo} - v_{co} \\ v_{ca} &= v_{co} - v_{ao} \end{aligned} \tag{15.76}$$

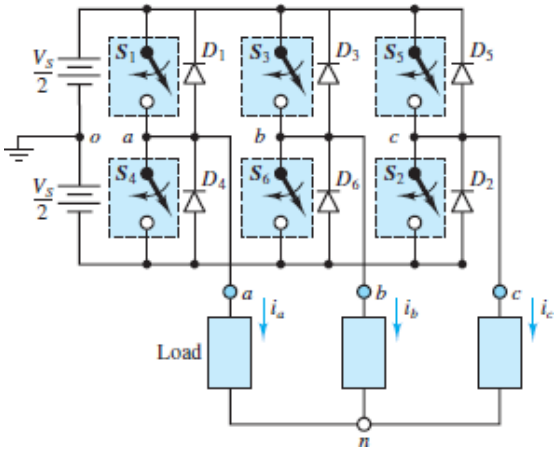


Figure 15.55 Three-phase voltage source inverter

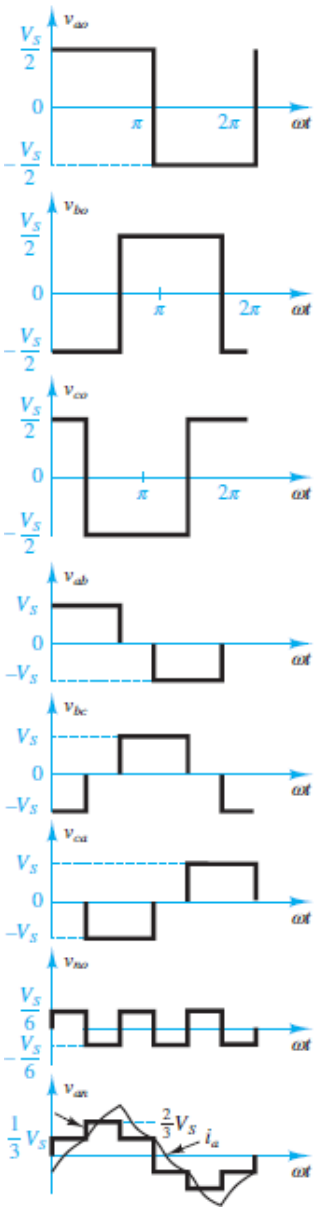


Figure 15.56 Three-phase voltage source inverter waveforms

and are shown in the second set of three diagrams in Figure 15.56. These are also phase-shifted by 120°. The pole voltages can be expressed in terms of the **load phase voltages** v_{an}, v_{bn}, and v_{cn}:

$$\begin{aligned} v_{ao} &= v_{an} - v_{no} \\ v_{bo} &= v_{bn} - v_{no} \\ v_{co} &= v_{cn} - v_{no} \end{aligned} \tag{15.77}$$

and since $v_{an} + v_{bn} + v_{cn} = 0$ is required for balanced operation, the following relationship can be derived for the DC **supply neutral** (o) to **load neutral** (n) voltage:

$$v_{no} = \frac{v_{ao} + v_{bo} + v_{co}}{3} \tag{15.78}$$

This voltage is also shown to be a square wave switching three times as fast as the inverter output voltage. Finally, the following relations can be used to obtain the phase voltages:

$$\begin{aligned} v_{an} &= v_{ao} - v_{no} = \tfrac{2}{3}v_{ao} - \tfrac{1}{3}(v_{bo} + v_{co}) \\ v_{bn} &= v_{bo} - v_{no} = \tfrac{2}{3}v_{bo} - \tfrac{1}{3}(v_{ao} + v_{co}) \\ v_{cn} &= v_{co} - v_{no} = \tfrac{2}{3}v_{bo} - \tfrac{1}{3}(v_{ao} + v_{bo}) \end{aligned} \tag{15.79}$$

Page 922

Only one phase voltage, v_{an}, is shown in the picture; however, it is straightforward to construct the other two phase voltages by using equation 15.79. Note that the load-phase voltage waveform shown in Figure 15.56 is a coarse stepwise approximation of a sinusoidal waveform; the corresponding load current i_a is a filtered version of the load voltage since the load is inductive in nature and is therefore somewhat smoothed with respect to the voltage waveform. The discontinuous nature of these waveforms creates a significant higher harmonic spectrum at frequencies that are integer multiples of the inverter output frequency; this is an unavoidable property of all inverters that employ switching circuits, but the problem can be reduced by using more complex switching schemes. Another major shortcoming of this AC supply is that if the DC supply is fixed, the amplitude of the inverter output is fixed.

The VSI circuit described in the foregoing paragraphs can provide a variable-frequency supply provided that the commutation frequency of the electronic switches can be varied. Thus, in general, it is necessary to also provide the capability for timing circuits that can provide variable switching rates; this is often accomplished with a microcontroller.

The limitations of the VSI of Figure 15.55 can be overcome with the use of more advanced switching schemes, such as *pulse-width modulation (PWM)* and *sinusoidal PWM*. The complexity of these schemes is beyond the scope of this book, and the interested reader is invited to explore a more advanced power electronics text to learn about advanced inverter circuits. However, it is possible

to significantly reduce the harmonic content of the inverter waveforms and to provide variable-frequency, variable-amplitude, three-phase supplies for AC motors by means of power switching circuits under microprocessor control.

EXAMPLE 15.15 Two-Quadrant Chopper

Problem

1. Determine the turn-on time of the chopper of Figure 15.50 in the motoring mode when n = 500 rpm and i_o = 90 A. Also determine the power absorbed by the motor armature winding, the power absorbed by the motor, and the power delivered by the source.
2. Determine the turn-on time of the chopper in the regenerative mode if n = 380 rpm and i_o = –90 A. Also determine the power absorbed by the motor armature winding, the power absorbed by the motor, and the power delivered by the source.

Solution

Known Quantities: Supply voltage; motor parameters; chopping frequency armature resistance and inductance.

Find: For each of the two cases: t_1, P_a, P_m, P_S.

Schematics, Diagrams, Circuits, and Given Data:

1. V_S = 120 V; $E_a = 0.1n$; R_a = 0.2 Ω; $1/T$ = chopping frequency = 300 Hz.
2. V_S = 120 V; $E_a = 0.1n$; R_a = 0.2 Ω; $L_S \rightarrow \infty$; $1/T$ = chopping frequency = 300 Hz.

Page 923

Assumptions: The switches in the chopper of Figure 15.50 act as ideal switches. Assume that the motor inductance is sufficiently small to be neglected in the calculations (i.e., assume a short-circuit).

Analysis:

1. *Analysis of motoring operation.* To analyze the motoring operation of the chopper, refer to Figure 15.45 and apply KVL to the motor side:

$$\langle v_o \rangle = R_a I_a + E_a = R_a \langle i_o \rangle + 0.1n = 0.2 \times 90 + 0.1 \times 500 = 68 \text{ V}$$

Use equation 15.70 to compute the duty cycle δ of the chopper:

$$\delta = \frac{t_1}{T} = \frac{\langle v_o \rangle}{V_S} = \frac{68}{120} = 0.567$$

Since the chopping frequency is 300 Hz, we can compute t_1:

$$t_1 = T\delta = \frac{0.567}{300} = 1.89 \text{ ms}$$

The power absorbed by the armature is

$$P_a = R_a I_a^2 = R_a \langle i_o \rangle^2 = 0.2 \times 90^2 = 1.62 \text{ kW}$$

The power absorbed by the motor is

$$P_m = E_a I_a = 0.1n \times \langle i_o \rangle = 0.1 \times 500 \times 90 = 4.5 \text{ kW}$$

The power delivered by the voltage supply is

$$P_S = \delta V_S \langle i_o \rangle = 0.567 \times 120 \times 90 = 6.12 \text{ kW}$$

2. *Analysis of regenerative operation.* To analyze the regenerative operation of the chopper, refer to Figure 15.47 and apply KVL to the motor side, noting that the current is reverse directed:

$$\langle v_o \rangle = R_a I_a + E_a = R_a \langle i_o \rangle + E_a = -90 \times 0.2 + 0.1 \times 380 = 20 \text{ V}$$

Use equation 15.74 and observe that the motor now acts as the source, and the supply voltage as the load.

$$V_S = \frac{1}{1-\delta}\langle v_o \rangle \quad \text{or} \quad 120 = \frac{1}{1-300t_1}20$$

This expression can be solved to find the duty cycle and t_1 for the step-up chopper.

$$\delta = \frac{5}{6} = 0.833 \quad \text{and} \quad t_1 = 2.8 \text{ ms}$$

The power absorbed by the armature is

$$P_a = R_a I_a^2 = R_a \langle i_o \rangle^2 = 0.2 \times (-90)^2 = 1.62 \text{ kW}$$

The power generated by the motor is

$$P_m = E_a I_a = 0.1n \times \langle i_o \rangle = 0.1 \times 380 \times 90 = 3.42 \text{ kW}$$

The power absorbed by the battery is

$$P_S = (1 - \delta) V_S \langle i_o \rangle = 0.167 \times 120 \times 90 = 1.8 \text{ kW}$$

Of course, the sum of the power absorbed by the armature and the battery is exactly equal to the power supplied by the motor.

Page 924

Comments:

1. In motoring operation, the sum of the motor and armature power losses is equal to the power supplied by the source; this result reflects the ideal (lossless) switches assumed in the example. In a practical chopper, the chopping circuit would absorb power; heat dissipation is therefore an important issue in the design of choppers.
2. In regenerative operation, the equivalent duty cycle is greater than 1. The motor supplies power to the rest of the circuit. However, the armature resistance still absorbs power, as must be true for any resistive load.
3. V_S could represent a battery pack in an electric vehicle, which would be recharged at the rate of 1.8 kW. The source of energy capable of producing this power is the inertial energy stored in the vehicle: when the vehicle decelerates, mechanical energy causes the electric motor to act as a generator, producing the 90-A current in the reverse direction.

Conclusion

This chapter introduces the most common classes of rotating electric machines. These machines, which can range in power from the milliwatt to the megawatt range, find common application in virtually every field of engineering, from consumer products to heavy-duty industrial applications. The principles introduced in this chapter can give you a solid basis from which to build upon.

Upon completing this chapter, you should have mastered the following learning objectives:

1. *Understand the basic principles of operation of rotating electric machines, their classification, and basic efficiency and performance characteristics.* Electric machines are defined in terms of their mechanical characteristics (torque–speed curves, inertia, friction and windage losses) and their electrical characteristics (current and voltage requirements). Losses and efficiency are an important part of the operation of electric machines, and it should be recognized that machines will suffer from electrical, mechanical, and magnetic core losses. All machines are based on the principle of establishing a magnetic field in the stationary part of the machine (stator) and a magnetic field in the moving part of the machine (rotor); electric machines can then be classified according to how the stator and rotor fields are established.
2. *Understand the operation and basic configurations of separately excited, permanent-magnet, shunt and series DC machines.* Direct-current machines, operated from a DC supply, are among the most common electric machines. The rotor (armature) circuit is connected to an external DC supply via a commutator. The stator electric field can be established by an external circuit (separately excited machines), by a permanent magnet (PM machines), or by the same supply used for the armature (self-excited machines).
3. *Analyze DC motors under steady-state and dynamic operation.* DC motors are commonly used in a variety of variable-speed applications (e.g., electric vehicles, servos) which require speed control; thus, their dynamics are also of interest.
4. *Analyze DC generators at steady state.* DC generators can be used to supply a variable direct current and voltage when propelled by a prime mover (engine, or other thermal or hydraulic machine).
5. *Understand the operation and basic configuration of AC machines, including the synchronous motor and generator and the induction machine.* AC machines require an alternating-current supply. The two principal classes of AC machines are the synchronous and induction types. Synchronous machines rotate at a predetermined speed, which is equal to the speed of a rotating magnetic field present in the stator, called the *synchronous speed.* Induction machines also operate based on a rotating Page 925magnetic field in the stator; however, the speed of the rotor is dependent on the operating conditions of the machine and is always less than the synchronous speed. Variable-speed AC machines require more sophisticated electric power supplies that can provide variable voltage/current and variable frequency. As

the cost of power electronics is steadily decreasing, variable-speed AC drives are becoming increasingly common.

6. Understand the basic operation of power converters, and in particular of DC-DC converters (choppers) used to control DC machines, and of DC-AC converters (inverters), used to control AC machines.

HOMEWORK PROBLEMS

Section 15.1: Rotating Electric Machines

15.1 The power rating of a motor can be modified to account for different ambient temperature, according to the following table:

Ambient temperature	**30°C**	**35°C**	**40°C**
Variation of rated power	+8%	+5%	0
Ambient temperature	**45°C**	**50°C**	**55°C**
Variation of rated power	–5%	–12.5%	–25%

A motor with P_e = 10 kW is rated up to 85°C. Find the actual power for each of the following conditions:

a. Ambient temperature is 50°C.

b. Ambient temperature is 30°C.

15.2 The speed-torque characteristic of an induction motor has been empirically determined as follows:

Speed (r/min)	**1,470**	**1,440**	**1,410**	**1,300**	**1,100**
Torque (N-m)	3	6	9	13	15
Speed (r/min)	**900**	**750**	**350**	**0**	
Torque (N-m)	13	11	7	5	

The motor will drive a load requiring a starting torque of 4 N-m and increase linearly with speed to 8 N-m at 1,500 r/min.

a. Find the steady-state operating point of the motor.

b. Equation 15.68 predicts that the motor speed can be regulated in the face of changes in load torque by adjusting the stator voltage. Find the change in voltage required to maintain the speed at the operating point of part a if the load torque increases to 10 N-m.

Section 15.2: Direct-Current Machines

15.3 Calculate the force exerted by each conductor, 6 in long, on the armature of a DC motor when it carries a current of 90 A and lies in a field the density of which is 5.2×10^{-4} Wb/in^2.

15.4 In a DC machine, the air gap flux density is 4 Wb/m^2. The area of the pole face is 2 cm × 4 cm. Find the flux per pole in the machine.

Section 15.3: Direct-Current Motors

15.5 A 220-V shunt motor has an armature resistance of 0.32 Ω and a field resistance of 110 Ω. At no load the armature current is 6 A and the speed is 1,800 r/min. Assume that the flux does not vary with load, and calculate

a. The speed of the motor when the line current is 62 A (assume a 2-V brush drop).

b. The speed regulation of the motor.

15.6 A 50-hp, 550-V shunt motor has an armature resistance, including brushes, of 0.36 Ω. When operating at rated load and speed, the armature takes 75 A. What resistance should be inserted in the armature circuit to obtain a 20 percent speed reduction when the motor is developing 70 percent of rated torque? Assume that there is no flux change.

15.7 A shunt DC motor has a shunt field resistance of 400 Ω and an armature resistance of 0.2 Ω. The motor nameplate rating values are 440 V, 1,200 r/min, 100 hp, and full-load efficiency of 90 percent. Find

a. The motor line current.

b. The field and armature currents.

c. The counter-emf at rated speed.

d. The output torque.

15.8 A 240-V series motor has an armature resistance of 0.42 Ω and a series-field resistance of 0.18 Ω. If the speed is 500 r/min when the current is 36 A, what will be the motor speed when the load reduces the line current to 21 A? (Assume a 3-V brush drop and that the flux is proportional to the current.)

15.9 A 220-V DC shunt motor [see Figure 15.14(b)] has an armature resistance of 0.2 Ω and a rated armature current of 50 A. Find

a. The voltage generated in the armature.

b. The power developed.

15.10 A 550-V series motor takes 112 A and operates at 820 r/min when the load is 75 hp. If the effective armature-circuit resistance is 0.15 Ω, calculate the horsepower output of the motor when the current drops to 84 A, assuming that the flux is reduced by 15 percent.

15.11 A 200-V DC shunt motor has the following parameters:

$$R_a = 0.1\ \Omega \qquad R_f = 100\ \Omega$$

When running at 1,100 r/min with no load connected to the shaft, the motor draws 4 A from the line. Find E and the rotational losses at 1,100 r/min (assuming that the stray-load losses can be neglected).

15.12 A 230-V DC shunt motor has the following parameters:

$$R_a = 0.5\ \Omega \qquad R_f = 75\ \Omega$$
$$P_{rot} = 500\ \text{W} \qquad \text{at } 1{,}120\ \text{r/min}$$

When loaded, the motor draws 46 A from the line. Find

a. The speed, P_{dev}, and T_{sh}.

b. If L_f = 25 H, L_a = 0.008 H, and the terminal voltage has a 115-V change, find $i_a(t)$ and $\omega_m(t)$.

15.13 A 200-VDC shunt motor with an armature resistance of 0.1 Ω and a field resistance of 100 Ω draws a line current of 5 A when running with no load at 955 r/min. Determine the motor speed, the motor efficiency, the total losses (i.e., rotational and I^2R losses), and the load torque T_{sh} that will result when the motor draws 40 A from the line. Assume rotational power losses are proportional to the square of shaft speed.

15.14 A 50-hp, 230-V shunt motor has a field resistance of 17.7 Ω and operates at full load when the line current is 181 A at 1,350 r/min. To increase the speed of the motor to 1,600 r/min, a resistance of 5.3 Ω is "cut in" via the field rheostat; the line current then increases to 190 A. Calculate

a. The power loss in the field and its percentage of the total power input for the 1,350 r/min speed.

b. The power losses in the field and the field rheostat for the 1,600 r/min speed.

c. The percent losses in the field and in the field rheostat at 1,600 r/min.

15.15 A 10-hp, 230-V shunt-wound motor has a rated speed of 1,000 r/min and full-load efficiency of 86 percent. Armature circuit resistance is 0.26 Ω; field-circuit resistance is 225 Ω. If this motor is operating under rated load and the field flux is very quickly reduced to 50 percent of its normal value, what will be the effect upon counter-emf, armature current, and torque? What effect will this change have upon the operation of the motor, and what will be its speed when stable operating conditions have been regained?

15.16 The machine of Example 15.5 is to be used in a series connection. That is, the field coil is connected in series with the armature, as shown in Figure P15.16. The machine is to be operated under the same conditions as in Example 15.5, that is, $n = 120$ r/min and $I_a = 8$ A. In the operating region, $\phi = kI_f$ and $k = 200$. The armature resistance is 0.2 Ω, and the resistance of the field winding is negligible.

a. Find the number of field winding turns necessary for full-load operation.

b. Find the torque output for the following speeds:

1. $n' = 2n$
2. $n' = 3n$
3. $n' = n/2$
4. $n' = n/4$

c. Plot the speed–torque characteristic for the conditions of part b.

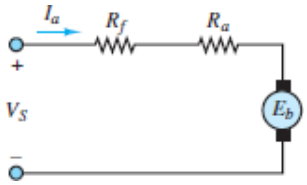


Figure P15.16

15.17 With reference to Example 15.7, assume that the load torque applied to the PM DC motor is zero. Determine the speed response of the motor speed to a step change in input voltage. Derive expressions for Page 927the natural frequency and damping ratio of the second-order system. What determines whether the system is over- or underdamped?

15.18 A motor with polar moment of inertia J develops torque according to the relationship $T = a\omega + b$. The motor drives a load defined by the torque–

speed relationship $T_L = c\omega^2 + d$. If the four coefficients are all positive constants, determine the equilibrium speeds of the motor-load pair, and whether these speeds are stable.

15.19 Assume that a motor has known friction and windage losses described by the equation $T_{FW} = b\omega$. Sketch the T-ω characteristic of the motor if the load torque T_L is constant, and the T_L-ω characteristic if the motor torque is constant. Assume that T_{FW} at full speed is equal to 30 percent of the load torque.

15.20 A PM DC motor is rated at 6 V, 3,350 r/min and has the following parameters: $r_a = 7\ \Omega$, $L_a = 120$ mH, $k_T = 7 \times 10^{-3}$ N-m/A, $J = 1 \times 10^{-6}$ kg-m^2. The no-load armature current is 0.15 A.

a. In the steady-state no-load condition, the magnetic torque must be balanced by an internal damping torque; find the damping coefficient b. Now sketch a model of the motor, write the dynamic equations, and determine the transfer function from armature voltage to motor speed. What is the approximate 3-dB bandwidth of the motor?

b. Now let the motor be connected to a pump with inertia $J_L = 1 \times 10^{-4}$ kg-m^2, damping coefficient $b_L = 5 \times 10^{-3}$ N-m-s, and load torque $T_L = 3.5 \times 10^{-3}$ N-m. Sketch the model describing the motor-load configuration, and write the dynamic equations for this system; determine the new transfer function from armature voltage to motor speed. What is the approximate 3-dB bandwidth of the motor/pump system?

15.21 A PM DC motor with torque constant k_{PM} is used to power a hydraulic pump; the pump is a positive displacement type and generates a flow proportional to the pump velocity: $q_p = k_p\omega$. The fluid travels through a conduit of negligible resistance; an accumulator is included to smooth out the pulsations of the pump. A hydraulic load (modeled by a fluid resistance R) is connected between the pipe and a reservoir (assumed at zero pressure). Sketch the motor-pump circuit. Derive the dynamic equations for the system, and determine the transfer function between motor voltage and the pressure across the load.

15.22 The shunt motor in Figure P15.22 is characterized by a field coefficient $k_f = 0.12$ V-s/A-rad, such that the back emf is given by the expression $E_b = k_f I_f \omega$ and the motor torque by the expression $T = k_f I_f I_a$. The motor drives

an inertia/viscous friction load with parameters $J = 0.8$ kg-m^2 and $b = 0.6$ N-m-s/rad. The field equation may be approximated by $V_S = R_f I_f$. The armature resistance is $R_a = 0.75\ \Omega$, and the field resistance is $R_f = 60\ \Omega$. The system is perturbed around the nominal operating point $V_{S0} = 150$ V, $\omega_0 = 200$ rad/s, and $I_{a0} = 186.67$ A.

a. Derive the dynamic system equations in *symbolic form.*

b. Linearize the equations you obtained in part a.

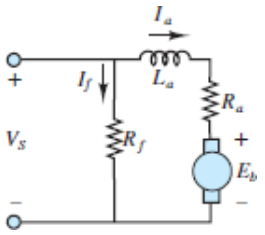


Figure P15.22

15.23 A PM DC motor is rigidly coupled to a fan; the fan load torque is described by the expression $T_L = 5 + 0.05\omega + 0.001\omega^2$, where torque is in newton-meters and speed in radians per second. The motor has $k_a\phi = k_T\phi = 2.42$; $R_a = 0.2\ \Omega$ and the inductance is negligible. If the motor voltage is 50 V, what is the speed of rotation of the motor and fan?

15.24 A separately excited DC motor has the following parameters:

$R_a = 0.1\ \Omega$	$R_f = 100\ \Omega$	$L_a = 0.2$ H
$L_f = 0.02$ H	$K_a = 0.8$	$K_f = 0.9$

An inertial load has $J = 0.5$ kg-m^2 and $b = 2$ N-m-s/rad. No external load torque is applied.

a. Sketch a diagram of the system and derive the (three) differential equations.

b. Sketch a simulation block diagram of the system (you should have three integrators).

c. Code the diagram, using Simulink.

d. Run the following simulations:

Armature control. Assume a constant field with V_f = 100 V; now simulate the response of the system when the armature voltage changes in step fashion from 50 to 75 V. Save and plot the current and angular speed responses.

Field control. Assume a constant armature voltage with V_a = 100 V; now simulate the response of the system when the field voltage changes in step Page 928fashion from 75 to 50 V. This procedure is called *field weakening.* Save and plot the current and angular speed responses.

15.25 Determine the transfer functions from *input voltage* to *angular velocity* and from *load torque* to *angular velocity* for a PM DC motor rigidly connected to an inertial load. Assume resistance and inductance parameters R_a, L_a let the armature constant be k_a. Assume ideal energy conversion, so that $k_a = k_T$. The motor has inertia J_m and damping coefficient b_m, and it is rigidly connected to an inertial load with inertia J and damping coefficient b. The load torque T_L acts on the load to oppose the magnetic torque.

15.26 Assume that the coupling between the motor and the inertial load of Problem 15.25 is flexible (e.g., a long shaft). This can be modeled by adding a torsional spring between the motor inertia and the load inertia. Now we can no longer lump together the two inertias and damping coefficients as if they were one; we need to write separate equations for the two inertias. In total, there will be three equations in this system: the motor electrical equation, the motor mechanical equation (J_m and B_m), and the load mechanical equation (J and B).

a. Sketch a diagram of the system.

b. Use free-body diagrams to write each of the two mechanical equations. Set up the equations in matrix form.

c. Compute the transfer function from input voltage to load speed, using the method of determinants.

15.27 A wound DC motor is connected in both a shunt and a series configuration. Assume generic resistance and inductance parameters R_2, R_2, L_2, L_2; let the field magnetization constant be k_f and the armature constant be k_a. Assume ideal energy conversion, so that $k_a = k_T$. The

motor has inertia J_m and damping coefficient b_m, and it is rigidly connected to an inertial load with inertia J and damping coefficient b.

a. Sketch a system-level diagram of the two configurations that illustrates both the mechanical and electrical systems.

b. Write an expression for the torque–speed curve of the motor in each configuration.

c. Write the differential equations of the motor-load system in each configuration.

d. Determine whether the differential equations of each system are linear; if one (or both) is (are) nonlinear, could they be made linear with some simple assumption? Explain clearly under what conditions this would be the case.

15.28 Derive the differential equations describing the electrical and mechanical dynamics of a shunt-connected DC motor, shown in Figure P15.28, and draw a simulation block diagram of the system. The motor constants are k_a, k_T = armature and torque reluctance and k_f = field flux.

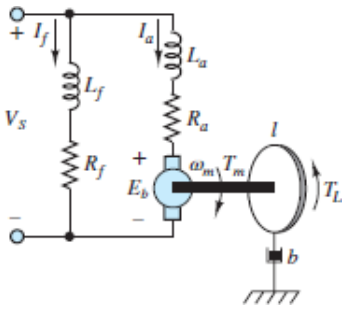


Figure P15.28

15.29 Derive the differential equations describing the electrical and mechanical dynamics of a series-connected DC motor, shown in Figure P15.29, and draw a simulation block diagram of the system. The motor constants are k_a, k_T = armature and torque reluctance and k_f = field flux.

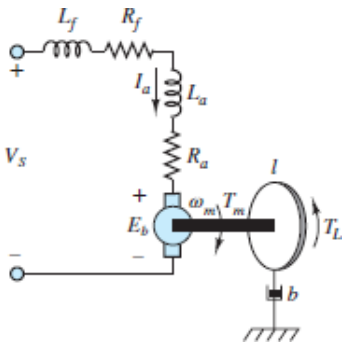


Figure P15.29

15.30 Develop a Simulink simulator for the shunt-connected DC motor of Problem 15.28. Assume the following parameter values: $L_a = 0.15$ H; $L_f = 0.05$ H; $R_a = 1.8\ \Omega$; $R_f = 0.2\ \Omega$; $k_a = 0.8$ V-s/rad; $k_T = 20$ N-m/A; $k_f = 0.20$ Wb/A; $b = 0.1$ N-m-s/rad; $J = 1$ kg-m^2.

15.31 Develop a Simulink simulator for the series-connected DC motor of Problem 15.29. Assume the following parameter values: $L = L_a + L_f = 0.2$ H; $R = R_a + R_f = 2\ \Omega$; $k_a = 0.8$ V-s/rad; $k_T = 20$ N-m/A; $k_f = 0.20$ Wb/A; $b = 0.1$ N-m-s/rad; $J = 1$ kg-m^2.

Page 929

Section 15.4: Direct-Current Generators

15.32 A 120-V, 10-A shunt generator has an armature resistance of 0.6 Ω. The shunt field current is 2 A. Determine the voltage regulation of the generator.

15.33 A 20-kW, 230-V separately excited generator has an armature resistance of 0.2 Ω and a load current of 100 A. Find

a. The generated voltage when the terminal voltage is 230 V.

b. The output power.

15.34 A 10-kW, 120-VDC series generator has an armature resistance of 0.1 Ω and a series field resistance of 0.05 Ω. Assuming that it is delivering rated current at rated power, find (a) the armature current and (b) the generated voltage.

15.35 The armature resistance of a 30-kW, 440-V shunt generator is 0.1 Ω. Its shunt field resistance is 200 Ω. Find

a. The power developed at rated load.

b. The load, field, and armature currents.

c. The electric power loss.

15.36 A four-pole, 450-kW, 4.6-kV shunt generator has armature and field resistances of 2 and 333 Ω. The generator is operating at the rated speed of 3,600 r/min. Find the no-load voltage of the generator and terminal voltage at half load.

15.37 A 30-kW, 240-V generator is running at half load at 1,800 r/min with an efficiency of 85 percent. Find the total losses and input power.

15.38 A self-excited DC shunt generator is delivering 20 A to a 100-V line when it is driven at 200 rad/s. The magnetization characteristic is shown in Figure P15.38. It is known that R_a = 1.0 Ω and R_f = 100 Ω. When the generator is disconnected from the line, the drive motor speeds up to 220 rad/s. What is the terminal voltage?

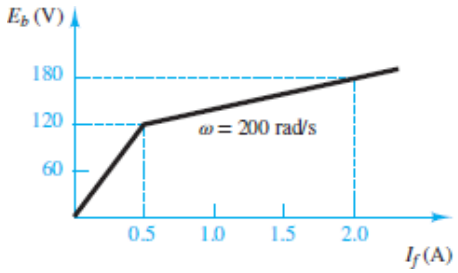


Figure P15.38

Section 15.6: The Alternator (Synchronous Generator)

15.39 An automotive alternator is rated 500 VA and 20 V. It delivers its rated voltamperes at a power factor of 0.85. The resistance per phase is 0.05 Ω, and the field takes 2 A at 12 V. If the friction and windage loss is 25 W and the core loss is 30 W, calculate the percent efficiency under rated conditions.

15.40 It has been determined by test that the synchronous reactance X_s and armature resistance r_a of a 2,300-V, 500-VA, three-phase synchronous generator are 8.0 and 0.1 Ω, respectively. If the machine is operating at rated load and voltage at a power factor of 0.867 lagging, find the generated voltage per phase and the torque angle.

Section 15.7: The Synchronous Motor

15.41 A non–salient pole, Y-connected, three-phase, two-pole synchronous machine has a synchronous reactance of 7 Ω and negligible resistance and rotational losses. One point on the open-circuit characteristic is given by $V_o = 400$ V (phase voltage) for a field current of 3.32 A. The machine is to be operated as a motor, with a terminal voltage of 400 V (phase voltage). The armature current is 50 A, with power factor 0.85, leading. Determine E_b, field current, torque developed, and power angle δ.

15.42 A factory load of 900 kW at 0.6 power factor lagging is to be increased by the addition of a synchronous motor that takes 450 kW. At what power factor must this motor operate, and what must be its kilovoltampere input if the overall power factor is to be 0.9 lagging?

15.43 A non–salient pole, Y-connected, three-phase, two-pole synchronous generator is connected to a 400-V (line to line), 60-Hz, three-phase line. The stator impedance is $0.5 + j1.6$ Ω (per phase). The generator is delivering rated current (36 A) at unity power factor to the line. Determine the power angle for this load and the value of E_b for this condition. Sketch the phasor diagram, showing $\mathbf{E}_b$, $\mathbf{I}_S$, and $\mathbf{V}_S$.

15.44 A non–salient pole, three-phase, two-pole synchronous motor is connected in parallel with a three-phase, Y-connected load so that the per-phase equivalent circuit is as shown in Figure P15.44. The parallel combination is connected to a 220-V (line to line), 60-Hz, three-phase line. The load current $\mathbf{I}_L$ is 25 A at a power factor of 0.866 inductive. The motor has $X_S = 2$ Ω and is operating with $I_f = 1$ A and $T = 50$ N-m at a power angle of –30°. (Neglect all Page 930losses for the motor.) Find $\mathbf{I}_S$, P_{in} (to the motor), the overall power factor (i.e., angle between $\mathbf{I}_1$ and $\mathbf{V}_S$), and the total power drawn from the line.

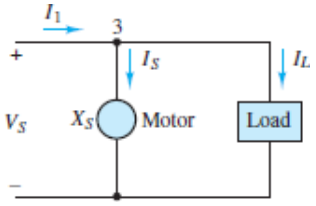


Figure P15.44

15.45 A four-pole, three-phase, Y-connected, non–salient pole synchronous motor has a synchronous reactance of 10 Ω. This motor is connected to a $230\sqrt{3}$ V (line to line), 60-Hz, three-phase line and is driving a load such that T_{shaft} = 30 N-m. The line current is 15 A, leading the phase voltage. Assuming that all losses can be neglected, determine the power angle δ and E for this condition. If the load is removed, what is the line current, and is it leading or lagging the voltage?

15.46 A 10-hp, 230-V, 60-Hz, three-phase, Y-connected synchronous motor delivers full load at a power factor of 0.8 leading. The synchronous reactance is 6 Ω, the rotational loss is 230 W, and the field loss is 50 W. Find

a. The armature current.

b. The motor efficiency.

c. The power angle.

Neglect the stator winding resistance.

15.47 A 2,000-hp, unity power factor, three-phase, Y-connected, 2,300-V, 30-pole, 60-Hz synchronous motor has a synchronous reactance of 1.95 Ω per phase. Neglect all losses. Find the maximum power and torque.

15.48 A 1,200-V, three-phase, Y-connected synchronous motor takes 110 kW (exclusive of field winding loss) when operated under a certain load at 1,200 r/min. The back emf of the motor is 2,000 V. The synchronous reactance is 10 Ω per phase, with negligible winding resistance. Find the line current and the torque developed by the motor.

15.49 The per-phase impedance of a 600-V, three-phase, Y-connected synchronous motor is $5 + j50$ Ω. The motor takes 24 kW at a leading power factor of 0.707. Determine the induced voltage and the power angle of the motor.

Section 15.8: The Induction Motor

15.50 A 74.6-kW, three-phase, 440-V (line to line), four-pole, 60-Hz induction motor has the following (per-phase) parameters referred to the stator circuit (see Figure 15.36):

$$R_S = 0.06\,\Omega \qquad X_S = 0.3\,\Omega \qquad X_m = 5\,\Omega$$
$$R_R = 0.08\,\Omega \qquad X_R = 0.3\,\Omega$$

The no-load power input is 3,240 W at a current of 45 A. Determine the line current, input power, developed torque, shaft torque, and efficiency at $s = 0.02$.

15.51 A 60-Hz, four-pole, Y-connected induction motor is connected to a 400-V (line to line), three-phase, 60-Hz line. The equivalent circuit parameters are

$$R_S = 0.2\ \Omega \qquad R_R = 0.1\ \Omega$$
$$X_S = 0.5\ \Omega \qquad X_R = 0.2\ \Omega$$
$$X_m = 20\ \Omega$$

When the machine is running at 1,755 r/min, the total rotational and stray-load losses are 800 W. Determine the slip, input current, total input power, mechanical power developed, shaft torque, and efficiency.

15.52 A three-phase, 60-Hz induction motor has eight poles and operates with a slip of 0.05 for a certain load. Determine

a. The speed of the rotor with respect to the stator.

b. The speed of the rotor with respect to the stator magnetic field.

c. The speed of the rotor magnetic field with respect to the rotor.

d. The speed of the rotor magnetic field with respect to the stator magnetic field.

15.53 A three-phase, two-pole, 400-V (per phase), 60-Hz induction motor develops 37 kW (total) of mechanical power P_m at a certain speed. The rotational loss at this speed is 800 W (total). (Stray-load loss is negligible.)

a. If the total power transferred to the rotor is 40 kW, determine the slip and the output torque.

b. If the total power into the motor P_{in} is 45 kW and R_S is 0.5 Ω, find I_S and the power factor.

15.54 The nameplate speed of a 25-Hz induction motor is 720 r/min. If the speed at no load is 745 r/min, find

a. The slip.

b. The percent regulation.

15.55 The nameplate of a squirrel cage four-pole induction motor has the following information: 25 hp, 220 V, three-phase, 60 Hz, 830 r/min, 64-A line current. If the motor draws 20,800 W when operating at full load, calculate

a. Slip.

b. Percent regulation if the no-load speed is 895 r/min.

c. Power factor.

d. Torque.

e. Efficiency.

15.56 A 60-Hz, four-pole, Y-connected induction motor is connected to a 200-V (line to line), three-phase, 60-Hz line. The equivalent circuit parameters are

$R_S = 0.48\ \Omega$	Rotational loss torque = 3.5 N-m
$X_S = 0.8\ \Omega$	$R_R = 0.42\ \Omega$ (referred to stator)
$X_m = 30\ \Omega$	$X_R = 0.8\ \Omega$ (referred to stator)

The motor is operating at slip s = 0.04. Determine the input current, input power, mechanical power, and shaft torque (assuming that stray-load losses are negligible).

15.57 a. A three-phase, 220-V, 60-Hz induction motor runs at 1,140 r/min. Determine the number of poles (for minimum slip), the slip, and the frequency of the rotor currents.

b. To reduce the starting current, a three-phase squirrel cage induction motor is started by reducing the line voltage to $V_s/2$. By what factor are the starting torque and the starting current reduced?

15.58 A six-pole induction motor for vehicle traction has a 50-kW input electric power rating and is 85 percent efficient. If the supply is 220 V at 60 Hz, compute the motor speed and torque at a slip of 0.04.

15.59 An AC induction machine has six poles and is designed for 60-Hz, 240-V (rms) operation. When the machine operates with 10 percent slip, it produces 60 N-m of torque.

a. The machine is now used in conjunction with a friction load that opposes a torque of 50 N-m. Determine the speed and slip of the machine when used with the above-mentioned load.

b. If the machine has an efficiency of 92 percent, what minimum rms current is required for operation with the load of part a?

(*Hint:* You may assume that the speed–torque curve is approximately linear in the region of interest.)

15.60 A blocked-rotor test was performed on a 5-hp, 220-V, four-pole, 60-Hz, three-phase induction motor. The following data were obtained: $V = 48$ V, $I = 18$ A, $P = 610$ W. Calculate

a. The equivalent stator resistance per phase R_S.

b. The equivalent rotor resistance per phase R_R.

c. The equivalent blocked-rotor reactance per phase X_R.

15.61 Calculate the starting torque of the motor of Problem 15.60 when it is started at

a. 220 V

b. 110 V

The starting torque equation is

$$T = \frac{m}{\omega_e} \cdot V_S^2 \cdot \frac{R_R}{(R_R + R_S)^2 + (X_R + X_S)^2}$$

15.62 A four-pole, three-phase induction motor drives a turbine load. At a certain operating point the machine has 4 percent slip and 87 percent efficiency. The motor drives a turbine with torque–speed characteristic given by $T_L = 20 + 0.006\omega^2$. Determine the torque at the motor-turbine shaft and the total power delivered to the turbine. What is the total power consumed by the motor?

15.63 A four-pole, three-phase induction motor rotates at 1,700 r/min when the load is 100 N-m. The motor is 88 percent efficient.

a. Determine the slip at this operating condition.

b. For a constant-power, 10-kW load, determine the operating speed of the machine.

c. Sketch the motor and load torque–speed curves on the same graph. Show numerical values.

d. What is the total power consumed by the motor?

15.64 Find the speed of the rotating field of a six-pole, three-phase motor connected to (a) a 60-Hz line and (b) a 50-Hz line, in revolutions per minute and radians per second.

15.65 A six-pole, three-phase, 440-V, 60-Hz induction motor has the following model impedances:

$$R_S = 0.8\,\Omega \qquad X_S = 0.7\,\Omega$$
$$R_R = 0.3\,\Omega \qquad X_R = 0.7\,\Omega$$
$$X_m = 35\,\Omega$$

Calculate the input current and power factor of the motor for a speed of 1,200 r/min.

Page 932

15.66 An eight-pole, three-phase, 220-V, 60-Hz induction motor has the following model impedances:

$$R_S = 0.78\,\Omega \qquad X_S = 0.56\,\Omega \qquad X_m = 32\,\Omega$$
$$R_R = 0.28\,\Omega \qquad X_R = 0.84\,\Omega$$

Find the input current and power factor of this motor for $s = 0.02$.

15.67 A nameplate is given in Example 15.2. Find the rated torque, rated voltamperes, and maximum continuous output power for this motor.

15.68 A three-phase induction motor, at rated voltage and frequency, has a starting torque of 140 percent and a maximum torque of 210 percent of full-load torque. Neglect stator resistance and rotational losses and assume constant rotor resistance. Determine

a. The slip at full load.

b. The slip at maximum torque.

c. The rotor current at starting as a percentage of full-load rotor current.

15.69 A 60-Hz, four-pole, three-phase induction motor delivers 35 kW of mechanical (output) power. At a certain operating point the machine has 4 percent slip and 87 percent efficiency. Determine the torque delivered to the load and the total electric (input) power consumed by the motor.

15.70 A four-pole, three-phase induction motor rotates at 16,800 rev/min when the load is 140 N-m. The motor is 85 percent efficient.

a. Determine the slip at this operating condition.

b. For a constant-power, 20-kW load, determine the operating speed of the machine.

c. Sketch the motor and load torque–speed curves for the load of part b on the same graph. Show numerical values.

15.71 An AC induction machine has six poles and is designed for 60-Hz, 240-V (rms) operation. When the machine operates with 10 percent slip, it produces 60 N-m of torque.

a. The machine is now used in conjunction with an 800-W constant power load. Determine the speed and slip of the machine when used with the above-mentioned load.

b. If the machine has an efficiency of 89 percent, what minimum rms current is required for operation with the load of part a?

(*Hint:* You may assume that the speed torque curve is approximately linear in the region of interest.)

Section 15.9: Electric Motor Drives

15.72 The DC-DC converter of Figure 15.45 is used to control the speed of a DC motor. Let the supply voltage be 120 V and the armature resistance of the motor be 0.15 Ω. The motor back-emf constant is 0.05 V/rpm, and the switching frequency is 250 Hz. Assume that the motor current is free of ripple and equal to 125 A at 120 rpm.

a. Determine the duty cycle of the converter δ and the converter on time t_1.

b. Determine the average power absorbed by the motor.

c. Determine the apparent power supplied by the source.

15.73 The circuit of Figure 15.49 is used to provide regenerative braking in a traction motor. The motor constant is 0.25 V/rpm, and the supply voltage is 550 V. The armature resistance is R_a = 0.15 Ω. The motor speed is 1,000 rpm, and the motor current is 200 A.

a. Determine the duty cycle δ of the converter.

b. Determine the average power fed back to the battery.

15.74 For the two-quadrant converter of Figure 15.50 assume that thyristors S_1 and S_2 are on for t_1 and off for $T - t_1$, where T is the switching period. Derive an expression for the average output voltage $v_{o_{av}}$ in terms of the supply voltage V_S and the duty cycle δ.

15.75 A boost converter is powered by an ideal 200-V battery pack. The load voltage waveform consists of rectangular pulses that are high (on) for 0.5 ms out of a total period of 3.0 ms. Calculate the average and rms values of the converter supply voltage.

15.76 A buck converter connected to a 100-V battery pack supplies an *RL* load, where $R = 0.5\ \Omega$ and $L = 1$ mH. The thyristor switch is on for 1 ms, and the period of the switching waveform is 3 ms. Calculate the average value of the load voltage and the power supplied by the battery.

15.77 The converter of Problem 15.76 is used to supply a separately excited DC motor with $R_a = 1\ \Omega$ and $L_a = 2$ mH. At the lowest speed of operation, the back-emf E_a is equal to 15 V. What is the average value of the load current and voltage when the switching period is 4 ms and the duty cycle is 0.5?

Page 933

15.78 A separately excited DC motor with $R_a = 0.33\ \Omega$ and $L_a = 15$ mH is controlled by a DC-DC converter in the range of 0 to 2,000 rpm. The DC supply is 220 V. If the load torque is constant and requires an average armature current of 25 A, calculate the range of duty cycles required if the motor armature constant is $K_a\phi = 0.00167$ V-s/rad.

15.79 A separately excited DC motor is rated at 10 kW, 240 V, 1,000 rpm, and is supplied by a single-phase controlled bridge rectifier. The power supply is sinusoidal and rated at 240 V, 60 Hz. The motor armature resistance is 0.42 Ω, and the motor constant is $K_a = 2$ V-s/rad. Calculate the speed, power factor, and efficiency for SCR firing angles α of 0° and 20° if the load torque is constant. Assume that additional inductance is present to ensure continuous conduction.

15.80 A separately excited DC motor is rated at 10 kW, 300 V, 1,000 rpm, and is supplied by a three-phase controlled bridge rectifier. The power supply is sinusoidal and rated at 220 V, 60 Hz. The motor armature resistance is 0.2 Ω, and the motor constant is $K_a = 1.38$ V-s/rad. The motor delivers rated

power at α = 0°. Calculate the speed, power factor, and efficiency for a firing angle α = 30° if the load torque is constant. Assume that additional inductance is present to ensure continuous conduction.

15.81 Sketch the current through the load R_o in the switched-mode power supply of Figure P15.81.

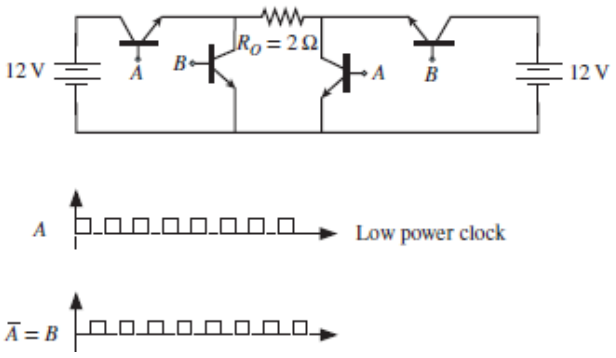


Figure P15.81

15.82 In the switched-mode power supply of Figure P15.82, sketch the load voltage signal v_o.

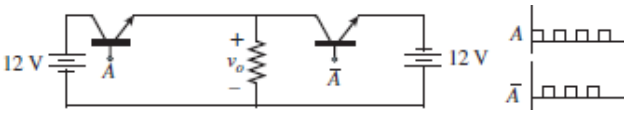


Figure P15.82

15.83 The switched mode power supply of Figure P15.83 will convert DC to three-phase AC. Sketch timing diagrams for the three low-power clock inputs A, B, and C to generate a balanced three-phase source. Also sketch the current in the neutral return wire. Assume the period of the cycle is normalized.

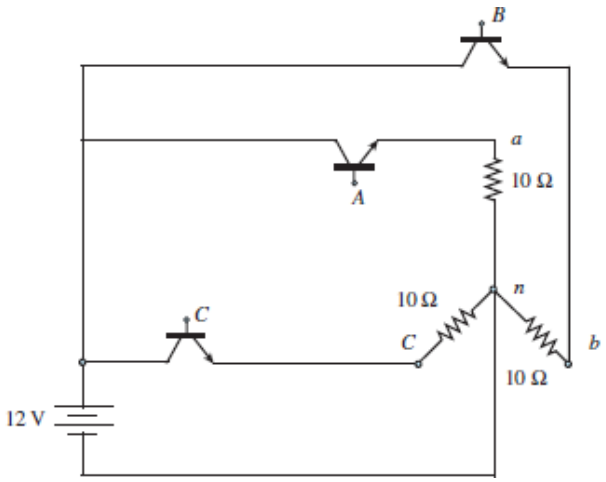


Figure P15.83

15.84 A DC-to-DC converter can also be thought of, and analyzed as, a DC transformer! Figure P15.84 shows a particular DC-to-DC converter configuration that converts the 1.2 V of a Ni-Cd battery cell to a desired 12 VDC supply. Using the usual transformer "reflecting theorems" (see Chapter 13), determine the power supplied by the 1.2-V source, and to the 10-Ω load.

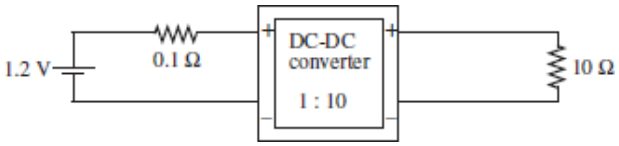


Figure P15.84

15.85 Shown in Figure P15.85 is a "charge pump" circuit for a switched-mode power supply (with all transistors acting in the switched mode). A 555-timer chip drives the two inputs, which are the timer's clock (CLK), and its inverse, clock-not ($\overline{CLK}$). $\overline{CLK}$ is low Page 934whenever *CLK* is high, and conversely. Assuming that the frequency of the clock is relatively high, determine the voltage across a high resistance load *R*.

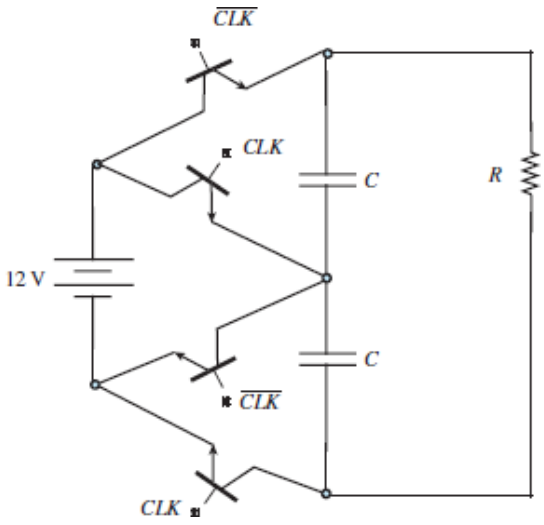


Figure P15.85

15.86 Sketch the low-power periodic signals, A, B, and C versus time that drive the high current power transistors in the (DC)-to-(three-phase-AC) switched-mode power supply of Figure P15.86, so that the load sees a balanced three-phase square-wave source.

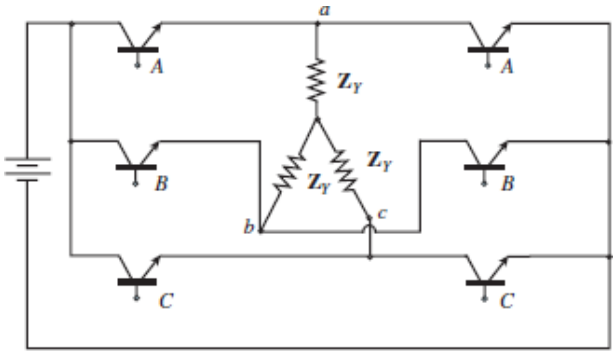


Figure P15.86

15.87 In the switched-mode power supply of Figure P15.87, sketch the load voltage signal v_o.

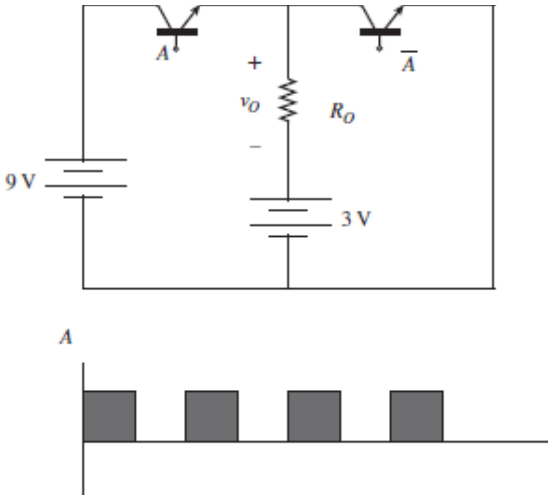


Figure P15.87

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]Note that the abbreviation *rpm* although certainly familiar to the reader, is not a standard unit, and its use should be discouraged.

C H A P T E R

16

SPECIAL-PURPOSE ELECTRIC MACHINES

The objective of this chapter is to introduce the operating principles and performance characteristics of a number of special-purpose electric machines that find widespread engineering application in a variety of fields, ranging from robotics to vehicle propulsion, aerospace, and automotive control. In Chapters 14 and 15, you were introduced to the operating principles of the major classes of electric machines: DC machines, synchronous machines, and induction motors. The machines discussed in this chapter operate according to the essential principles described earlier but are also characterized by unique features that set them apart from the machines described in Chapter 15. The first of these special-purpose machines is the brushless DC motor. Next, we discuss stepping motors, illustrating a very natural match between electromechanical devices and digital logic. The switched reluctance motor is presented next. A discussion of universal motors and single-phase induction motors follows, with a brief description of the types of electronic drives used to supply power to these machines. The discussion of the electronic drives ties the electromechanics material with the subject of power electronics introduced in Chapter 17. Finally, in the last section motor selection and applications are discussed, along with coverage of design and performance specifications.

The machines introduced in this chapter are used in many applications requiring fractional horsepower, or the ability to accurately control position, velocity, or torque.

Learning Objectives

Students will learn to...

1. Understand the basic principles of operation of brushless DC motors, and the tradeoffs between these and brush-type DC motors. *Section 16.1*.
2. Understand the operation and basic configurations of step motors as well as step sequences for the different classes of step motors. *Section 16.2*.
3. Understand the operating principles of switched reluctance machines *Section 16.3*.
4. Classify and analyze single-phase AC motors, including the universal motor and various types of single-phase induction motors, using simple circuit models. *Section 16.4*.
5. Outline the selection process for an electric machine, given an application; perform calculations related to load inertia, acceleration, efficiency, and thermal characteristics. *Section 16.5*.

16.1 BRUSHLESS DC MOTORS

In spite of its name, the brushless DC motor is actually not a DC motor, but a permanent-magnet synchronous machine; the name is actually due not to the construction of the machine, but to the fact that its operating characteristics resemble those of a shunt DC motor with constant field current. This characteristic can be obtained by providing the motor with a power supply whose electrical frequency determines the mechanical frequency of rotation of the rotor. To generate a source of variable frequency, use is made of DC-to-AC converters (inverters), consisting of banks of transistors that are switched on and off at a frequency corresponding to the rotor speed; thus, the inverter converts a DC source to an AC source of variable frequency. As far as the user is concerned, then, the source of excitation of a brushless DC motor is DC although the current that actually flows through the motor windings is AC. In

effect, the brushless DC motor is a synchronous motor in which the torque angle δ is kept constant by an appropriate excitation current.

Brushless DC motors require measurement of the position of the rotor to determine its speed of rotation, and to generate a supply current at the same frequency. This function is accomplished by means of a position-sensing arrangement that usually consists either of a magnetic Hall-effect position sensor, which senses the passage of each pole in the rotor, or of an optical encoder similar to the encoders discussed in Chapter 11.

Figure 16.1(a) depicts the appearance of a brushless DC motor. Note how the multiphase winding is similar to that of the synchronous motor of Chapter 15. Figure 16.1(b) depicts the construction of a typical brushless DC servomotor. The brushless motor consists of a stator with a multiphase winding, usually three phase; a permanent-magnet rotor; and a rotor position sensor. It is interesting to observe that since the commutation is performed electronically by switching the current to the motor—rather than by brushes, as in DC motors—the brushless motor can be produced in many different configurations, including, for example, very flat ("pancake") motors. Page 16-3Figure 16.1 shows the classical configuration of inside rotor, outside stator. For simple machines, it is also possible to resort to an outside rotor, with greater ease of magnet attachment and inherently smoother rotation, but with inferior thermal characteristics, since a stator encased within the rotor structure cannot be cooled efficiently.

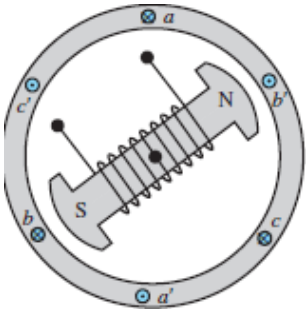


Figure 16.1(a) Two-pole brushless DC motor with three-phase stator winding

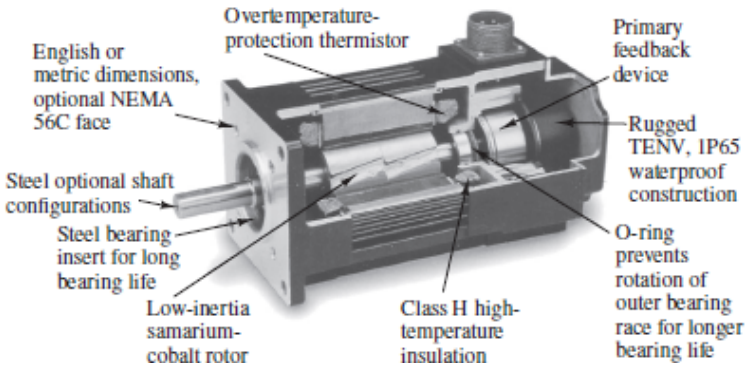


Figure 16.1(b) A typical brushless DC servomotor (*Courtesy of Kollmorgen Corporation*.)

In conventional DC motors, the supply voltage is limited by brush wear and sparking that can occur at the commutator, often resulting in the need for transformers to step down the supply voltage. In brushless DC motors, on the other hand, such a concern does not arise because the commutation is performed electronically without the need for brushes. Further, since, in general, the armature (load-carrying winding) is on the stator and thus the losses are concentrated in the stator, liquid cooling (if required) is feasible and does not involve excessive complexity. You will recall that in a conventional DC motor the armature is on the rotor, and therefore auxiliary liquid cooling is very difficult to implement.

Another important advantage of brushless DC motors is that by sealing the stator, submersible units can be built. In addition to these operational advantages, note that these motors are also characterized by easier construction: The construction of the stator in a brushless DC motor is similar to that in traditional induction motors and is therefore suitable for automated production. The windings may also be fitted with temperature sensors, providing the possibility of additional thermal protection.

The permanent-magnet rotor is typically made either of rare-earth magnets (Sm-Co) or of ceramic magnets (ferrites). Rare-earth magnets have outstanding magnetic properties, but they are expensive and in limited supply, and therefore the more commonly employed materials are ceramic magnets. Rare-earth magnet motors can be a cost-effective solution—since they allow much greater fluxes to be generated by a given supply current—in applications where high speed, high efficiency, and small size are important. Brushless DC motors can be rated up to 250 kW at 50,000 r/min. The rotor position sensor must be

designed for operation inside the motor, and must withstand the backlash, vibrations, and temperature range typical of motor operation.

Brushless DC motors do require a position-sensing device, though, to permit proper switching of the supply current. Recall that the brushless DC motor replaces the cumbersome mechanical commutation arrangement with electronic switching of the supply current. The most commonly used position-sensing devices are *position encoders* and **resolvers**. The resolver, shown in Figure 16.2, is a rotating machine that is mechanically coupled to the rotor of the brushless motor and consists of two Page 16-4stator and two rotor windings; the stator windings are excited by an AC signal, and the resulting rotor voltages are proportional to the sine and cosine of the angle of rotation of the rotor, thus providing a signal that can be directly related to the instantaneous position of the rotor. The resolver has two major disadvantages: First, it requires a separate AC supply; second, the resolver output must be appropriately decoded to obtain a usable position signal. For these reasons, *angular position encoders* (see Chapter 11) are often used. You will recall that such encoders provide a digital signal directly related to the position of a rotating shaft. Their output can therefore be directly used to drive the current supply for a brushless motor.

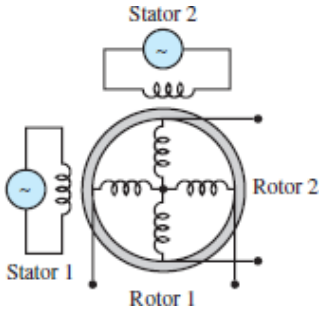


Figure 16.2 Resolver

To understand the operation of the brushless DC motor, it is useful to make an analogy with the operation of a permanent-magnet (PM) DC motor. As discussed in Chapter 15, in a permanent-magnet DC motor, a fixed magnetic field generated by the permanent magnets interacts with the field induced by the currents in the rotor windings, thus creating a mechanical torque. As the rotor turns in response to this torque, however, the angle between the stator and rotor fields is reduced, so that the torque would be nullified within a rotation of 90 electrical degrees. To sustain the torque acting on the rotor, permanent-magnet DC motors incorporate a commutator, fixed to the rotor shaft. The commutator switches the supply current to the stator so as to maintain a constant angle δ =

90° between interacting fields. Because the current is continually switched between windings as the rotor turns, the current in each stator winding is actually alternating, at a frequency proportional to the number of motor magnetic poles and the speed.

The basic principle of operation of the brushless DC motor is essentially the same, with the important difference that the supply current switching takes place electronically, instead of mechanically. Figure 16.3 depicts a transistor switching circuit capable of switching a DC supply so as to provide the appropriate currents to a three-phase rotor winding. The electronic switching device consists of a rotor position sensor, fixed on the motor shaft, and an electronic switching module that can supply each stator winding. Diagrams of the phase-to-phase back emf's and the switching sequence of the inverter are shown in Figure 16.4. The back emf Page 16-5waveforms shown in Figure 16.4 are called *trapezoidal*; the total back emf of the inverter is obtained by piecewise addition of the motor phase voltages and is a constant voltage, proportional to motor speed. You should visually verify that the addition of the three phase voltages of Figure 16.4 leads to a constant voltage. The brushless DC motor (BLDC) is therefore similar to a standard permanent-magnet DC motor, and it can be described by the following simplified equations:

$$V = k_a \omega_m + R_w I \tag{16.1}$$

$$T = k_T I \tag{16.2}$$

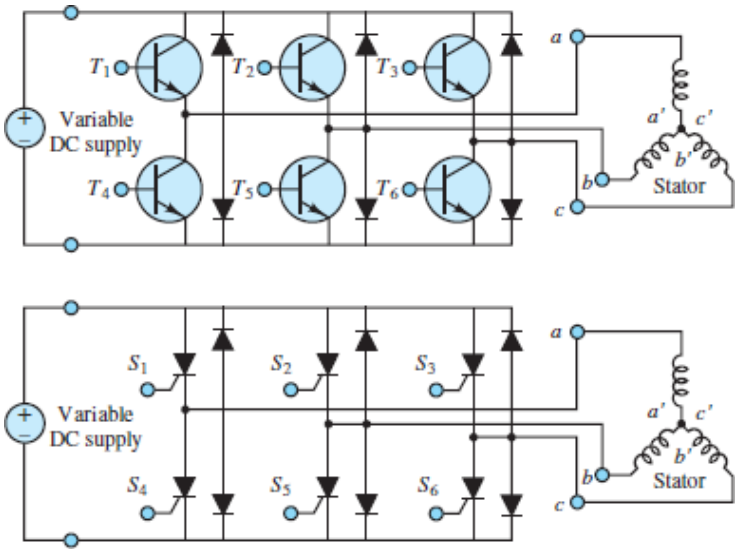


Figure 16.3 Transistor and SCR drives for a brushless DC motor

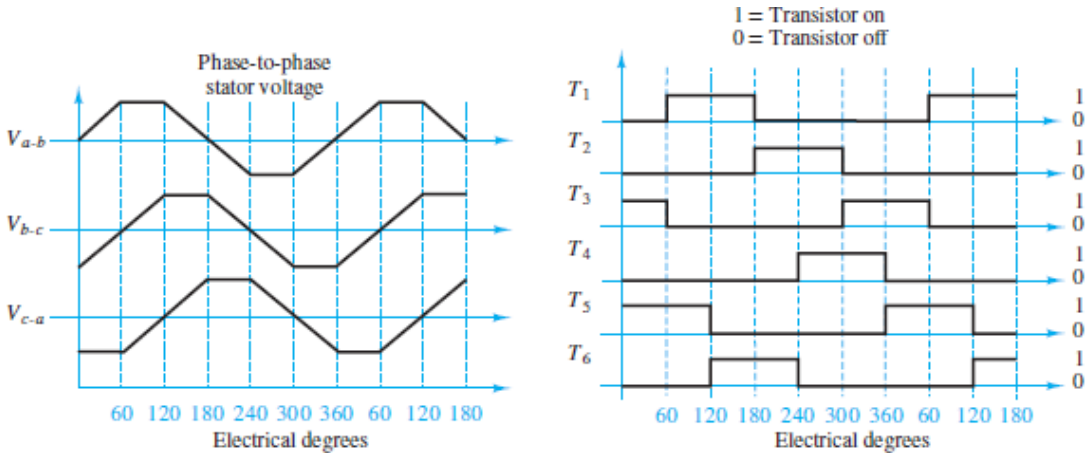


Figure 16.4 Phase voltages and transistor switching sequence for the brushless DC motor drive of Figure 16.3

where

$$k_a = k_T$$

and where

V = motor voltage
k_a = armature constant
ω_m = mechanical speed
R_w = winding resistance
T = motor torque
k_T = torque constant
I = motor (armature) current

The speed and torque of a brushless DC motor can therefore be controlled with any variable-speed DC supply, such as one of the supplies discussed in Chapter 17. Further, since the brushless motor has intrinsically higher torque and lower inertia than its DC counterpart, its response speed is superior to that obtained from traditional DC motors. Figure 16.5 depicts the (a) torque–speed and (b) efficiency curves of a commercially produced brushless DC motor.

One important difference between the conventional DC motor and the brushless motor, however, is due to the coarseness of the electronic switching compared with the mechanical switching of the brush-type DC motor (recall the discussion of Page 16-6torque ripple due to the commutation effect in DC motors in Chapter 15). In practice, one cannot obtain the exact trapezoidal emf of Figure 16.4 by means of the transistor switching circuit of Figure 16.3, and a voltage ripple results as a consequence, leading to a torque ripple in the motor. Additional phase windings on the stator could solve the problem, at the expense of further complexity in the drive electronics, since the switching sequence would be more complex. Thus, brushless motors suffer from an inherent tradeoff between torque ripple and drive complexity.

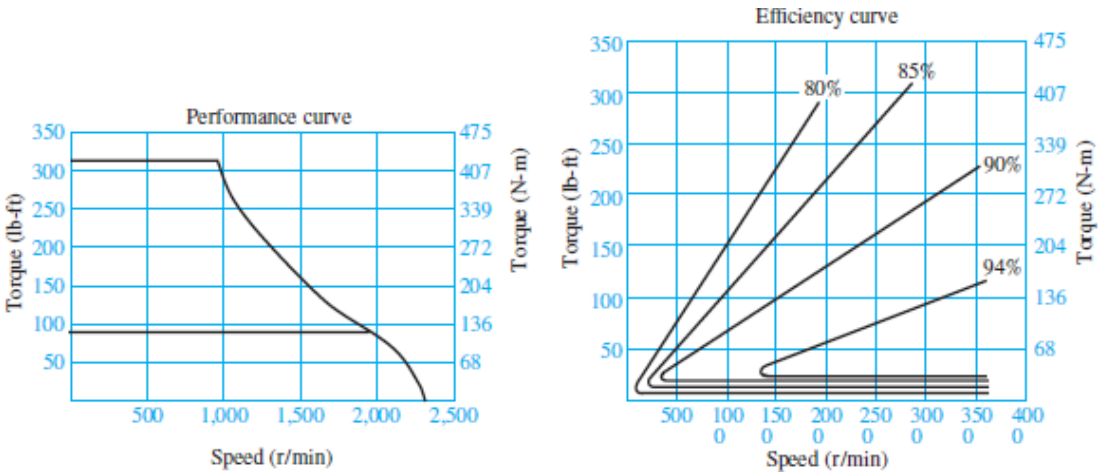


Figure 16.5 Performance and efficiency characteristics of brushless DC motor (*Courtesy Pacific Scientific*)

Among other applications, brushless DC motors find use in the design of servo loops in control systems, for example, in computer disk drives, and in propulsion systems for electric vehicles. The comparisons between the conventional DC motor and the brushless DC motor are summarized in the following table:

Conventional DC motors

Advantages

1. Controllability over a wide range of speeds.
2. Capability of rapid acceleration and deceleration.
3. Convenient control of shaft speed and position by servo amplifiers.

Disadvantage

1. Commutation (through brushes) causing wear, electrical noise, and sparking.

Brushless DC motors

Advantages

1. Controllability over a wide range of speeds.
2. Capability of rapid acceleration and deceleration.
3. Convenient control of shaft speed and position.
4. No mechanical wear or sparking problem due to commutation.
5. Better heat dissipation capabilities.

Disadvantage

1. Need for more complex power electronics than the brush-type DC motor for equivalent power rating and control range.

EXAMPLE 16.1 Sinusoidal Torque Generation in Brushless DC Motors

Problem

Show that the use of sinusoidal currents in a brushless DC motor can result in a ripple-free torque.

Solution

Known Quantities: Coil (phase) currents.

Find: Total output torque T.

Schematics, Diagrams, Circuits, and Given Data: $I_{m1} = I_m \sin\theta$; $I_{m2} = I_m \cos\theta$.

Assumptions: The field coil is wound in a two-phase circuit; each winding is sinusoidally spaced. Sinusoidal currents can be generated by suitable power electronics circuits.

Analysis: Using equation 16.2, we determine that the torques generated by the currents in each of the two coils of the two-phase stator are

$$T_1 = k_T I_{m1} \sin\theta$$
$$T_2 = k_T I_{m2} \cos\theta$$

The sinusoidal form of the torques is due to the sinusoidal distribution of the stator windings in each phase, which are spaced 90° out of phase with one another so as to produce sine-cosine components, and is shown in Figure 16.6.

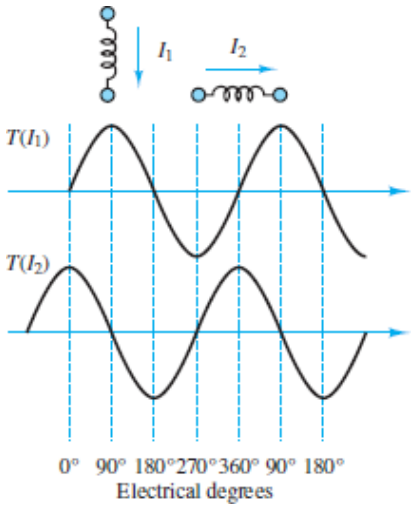


Figure 16.6 Sinusoidal torque generation circuit and current waveforms for a brushless DC motor

The net torque produced by the motor is the sum of the two phase torques:

$$T = T_1 + T_2 = k_T I_{m1} \sin\theta + k_T I_{m2} \cos\theta = k_T[(I_m \sin\theta) \sin\theta + (I_m \cos\theta) \cos\theta]$$
$$= k_T I_m(\sin^2\theta + \cos^2\theta) = k_T I_m$$

Thus, the torque generated by the motor is constant, or ripple free.

Page 16-8

Comments: Note that this scheme requires two features: sinusoidally spaced two-phase windings and sinusoidal phase currents. It is also very important that both the windings and the currents be exactly 90° out of phase.

EXAMPLE 16.2 Selecting a Trapezoidal Speed Profile to Match a Desired Motion Profile

Problem

Determine the trapezoidal speed profile required to move a load 0.5 m in 5 s. Analyze the motion of the motor.

Solution

Known Quantities: Desired load motion profile.

Find: Required trapezoidal speed profile.

Schematics, Diagrams, Circuits, and Given Data: The motor covers 0.5 m in 100 revolutions. Trapezoidal profile characteristics as shown in Figure 16.7.

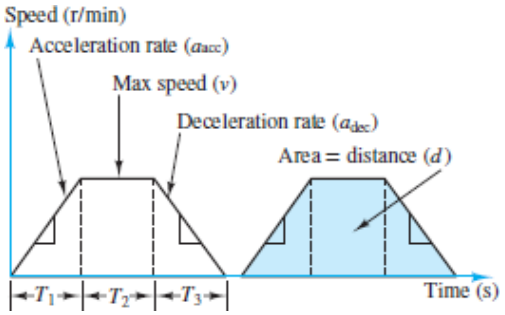


Figure 16.7 Trapezoidal profile

Assumptions: Assume a trapezoidal speed profile and that the motor will accelerate for 1 s and decelerate for 1 s.

Analysis: Define the following quantities:

d = motor travel (r)
v = motor speed (r/s)
T_1 = acceleration time (s)
T_2 = time at maximum speed (s)
T_3 = deceleration time (s)
a = acceleration or deceleration rate (r/s^2)

From the above definitions, we can calculate the maximum rotational velocity of the motor as follows. For constant acceleration, the expressions for the motor displacement and velocity are

$$d = \tfrac{1}{2}at^2 \qquad \text{and} \qquad v = d' = at$$

From the above expressions, we can relate the maximum velocity to the acceleration and deceleration rates:

$$a_{acc} = \frac{v}{T_1} \qquad a_{dec} = \frac{v}{T_3}$$

Page 16-9

Now we can write an expression for the total motor travel (100 revolutions):

$$d = \frac{1}{2}a_{acc}T_1^2 + vT_2 + \frac{1}{2}a_{dec}T_3^2 = \frac{1}{2}\frac{v}{T_1}T_1^2 + vT_2 + \frac{1}{2}\frac{v}{T_3}T_3^2$$
$$= \frac{1}{2}vT_1 + vT_2 + \frac{1}{2}vT_3 = v\left(\frac{1}{2}T_1 + T_2 + \frac{1}{2}T_3\right)$$

Note that the above expression is quite general and could be used also for asymmetric profiles. Using the given numbers, we calculate the maximum velocity to be

$$v = \frac{d}{\frac{1}{2}T_1 + T_2 + \frac{1}{2}T_3} = \frac{100 \text{ r}}{(0.5 + 3 + 0.5) \text{ s}} = 25 \text{ r/s}$$

which corresponds to 25 × 60 = 1,500 r/min.

Comments: The results derived in this example are very useful—trapezoidal speed profiles are very common in servomotors.

16.2 STEPPING MOTORS

Stepping, or stepper, motors convert digital information to mechanical motion. The principles of operation of stepping motors have been known since the 1920s; however, their application has seen a dramatic rise with the increased use of digital computers. Stepping motors, as the name suggests, rotate in distinct steps, and their position can be controlled by means of logic signals. Typical applications of stepping motors are in-line printers, positioning of heads in magnetic disk drives, and any other situation where continuous or stepwise displacements are required.

Stepping motors can generally be classified in one of three categories: variable-reluctance, permanent-magnet, and hybrid types. The principles of operation of each of these devices bear a definite resemblance to those of devices already encountered in the book. Stepping motors have a number of special features that make them particularly useful in practical applications. Perhaps the most important feature of a stepping motor is that the angle of rotation of the motor is directly proportional to the number of input pulses; further, the angle error per step is very small and does not accumulate. Stepping motors are also capable of rapid responses—starting, stopping, and reversing commands—and can be driven directly by digital signals. Another important feature is a self-holding capability that makes it possible for the rotor to be held in the stopped position without the use of brakes. Finally, a wide range of rotating speeds—proportional to the frequency of the pulse signal—may be attained in these motors.

Figure 16.8 depicts the general appearance of three types of stepping motors. The **permanent-magnet rotor stepping motor**, seen in Figure 16.8(a), permits a nonzero holding torque when the motor is not energized. Depending on the construction of the motor, it is typically possible to obtain step angles of 7.5, 11.25, 15, 18, 45, or 90°. The angle of rotation is determined by the number of stator poles, as is illustrated in Example 16.3. The **variable-reluctance stepping motor**, seen in Figure 16.8(b), has an iron multipole rotor and a laminated wound stator, and it rotates when the teeth on the rotor are attracted to the electromagnetically energized stator teeth. The rotor inertia of a variable-reluctance stepping motor is low, and the response is very quick, but the allowable load inertia is small. When Page 16-10the windings are not energized,

the static torque of this type of motor is zero. Generally, the step angle of the variable-reluctance stepping motor is 15° (see Examples 16.4 and 16.5).

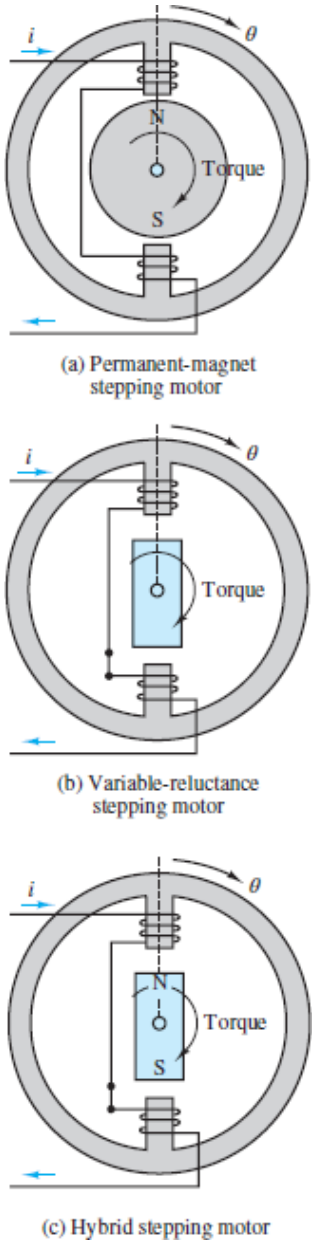


Figure 16.8 Stepping motor configurations

The **hybrid stepping motor**, seen in Figure 16.8(c), is characterized by a multitoothed stator and rotor, the rotor having an axially magnetized concentric

magnet around its shaft. It can be seen that this configuration is a mixture of the variable-reluctance (VR) and permanent-magnet types. This type of motor generally has high accuracy and high torque and can be configured to provide a step angle as small as 1.8°. Figure 16.9(a) through (e) depicts the construction of a VR step motor.

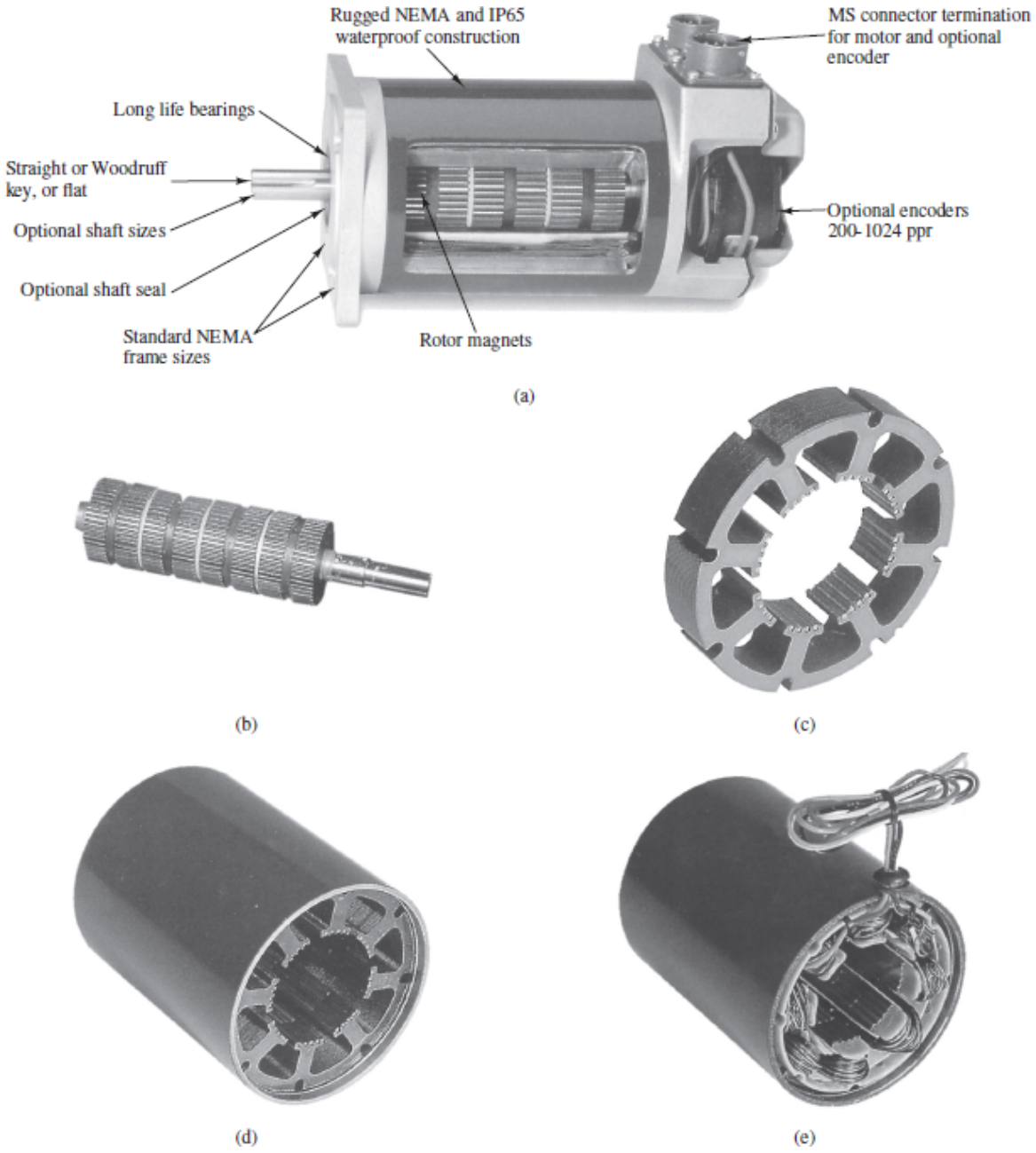


Figure 16.9 VR stepper motor: (a) complete motor assembly; (b) PM rotor; (c) stator cross section; (d) fully assembled stator; (e) stator

with windings (*Courtesy of Kollmorgen Corporation*.)

Page 16-11

For any of these configurations, the principle of operation is essentially the same: When the coils are energized, magnetic poles are generated in the stator, and the rotor will align in accordance with the direction of the magnetic field developed in the stator. By reversing the phase of the currents in the coils, or by energizing only some of the coils (this is possible in motors with more than two stator poles), the alignment of the stator magnetic field can take one of a discrete number of positions; if the currents in the coils are pulsed in the appropriate sequence, the rotor will advance in a step-by-step fashion. Thus, this type of motor can be very useful whenever precise incremental motion must be attained. As mentioned earlier, typical applications are in printer wheels, computer disk drives, and plotters. Other applications are found in the control of the position of valves (e.g., control of the throttle valve in an engine, or of a hydraulic valve in a fluid power system) and in drug-dispensing apparatus for clinical applications.

Examples 16.3 to 16.6 illustrate the operation of a four-pole, two-phase permanent-magnet stepping motor and of a similar motor of the variable-reluctance type. The operation of these motors is representative of all stepping motors.

EXAMPLE 16.3 Analysis of Two-Phase, Four-Pole Step Motor

Problem

Determine the full-step single-phase, full-step two-phase, and half-step current excitation sequences for the PM step motor of Figure 16.10.

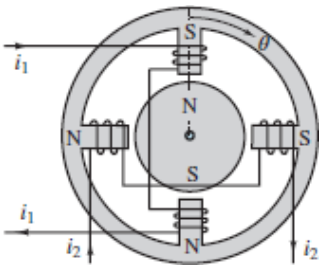


Figure 16.10 Two-phase four-pole PM stepper motor

Solution

Known Quantities: Phase currents.

Find: Full-step sequence for the motor.

Assumptions: The motor currents at the start of the sequence are $i_1 > 0$ and $i_2 = 0$.

Analysis: With the initial currents assumed (phase 1 energized), the motor will be at rest if the rotor is in the position shown in Figure 16.10. A single-phase sequence consists of turning on each of the two coils in sequence, reversing the polarity of the currents every other time. Then the PM rotor will align with the stator poles according to the polarity of the magnetic field generated by each coil's pole pair. For example, if coil 1 is turned off and coil 2 is turned on with a positive current polarity, the rotor will rotate clockwise by 90°. Table 16.1 depicts the (bipolar) sequence of coil currents and the corresponding motor position.

If both coils are activated, it is possible to cause the rotor to align between stator poles, also in increments of 90°, but shifted in phase by 45° with respect to the single-phase stepping sequence. Table 16.2 illustrates this stepping sequence.

Table 16.1 Full-step, single-phase sequence

i_1	i_2	θ
+	0	0°
0	+	90°
−	0	180°
0	−	270°
+	0	0°

Table 16.2 Full-step, two-phase sequence

i_1	i_2	θ
+	+	45°
−	+	135°
−	−	225°
+	−	315°
+	+	45°

Finally, if one combines the two sequences (easily accomplished, since the current commands for the two sequences are distinct), it is possible to obtain increments of 45°. Table 16.3 depicts the half-step sequence. Any finer resolution would require an increase in the number of windings and teeth in the stator.

Table 16.3 Half-step sequence

i_1	i_2	θ
+	0	0°
+	+	45°
0	+	90°
−	+	135°
−	0	180°
−	−	225°
0	−	270°
+	−	315°
+	0	0°

Comments: The simplicity of the electronic controls required by this type of machine is one of the very attractive features of step motors.

Page 16-12

EXAMPLE 16.4 Analysis of Variable-Reluctance Step Motor Problem

Determine the current excitation sequences required to achieve 45° steps in the VR step motor of Figure 16.11.

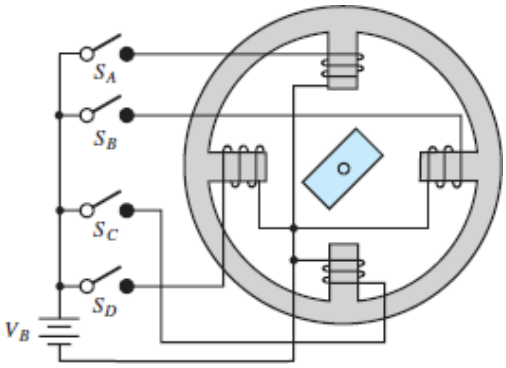


Figure 16.11 Two-phase, four-pole VR stepping motor

Solution

Known Quantities: Phase currents.

Find: Current excitation sequence for 45° steps.

Assumptions: The motor currents at the start of the sequence are $i_1 > 0$ and $i_2 = 0$.

Analysis: The operation of the variable-reluctance step motor (with a salient pole rotor) is simpler than that of the PM type because the rotor is not magnetically polarized, and therefore it is not necessary to have bipolar currents to achieve the desired rotor motion. The stator of Figure 16.10 is excited by direct currents supplied by a single (unipolar) voltage supply. The switches shown in the figure could be controlled by a logic circuit similar to the ones described in Chapters 11 and 12. Note that four separate coils are used.

Page 16-13

Figure 16.12 depicts how the first three steps of the sequence could be achieved. These are summarized in Table 16.4.

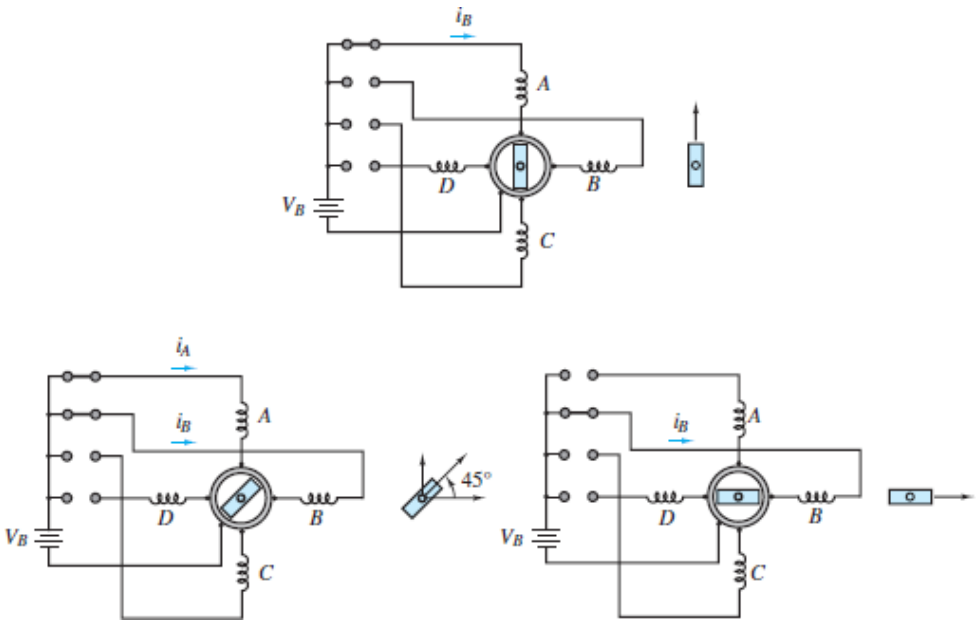


Figure 16.12 Two-phase, four-pole VR motor positioning sequence

Table 16.4 Current excitation sequence for VR step motor

S_A	S_B	S_C	S_D	Rotor position
1	0	0	0	0°
1	1	0	0	45°
0	1	0	0	90°
0	1	1	0	135°
0	0	1	0	180°
0	0	1	1	225°
0	0	0	1	270°
1	0	0	1	315°
1	0	0	0	360°

Comments: Note that the circuit required to drive this circuit is even simpler than the one required by the PM step motor.

EXAMPLE 16.5 Step Angle Determination of VR Step Motor

Problem

Determine an expression for the step angle of a VR step motor based on the number of teeth on the rotor and stator and on the number of phases.

Solution

Known Quantities: Number of rotor and stator teeth; number of phases.

Page 16-14

Find: Step angle.

Schematics, Diagrams, Circuits, and Given Data: t = number of teeth = 4; m = number of phases = 3.

Analysis: The number of steps in a revolution N is given by the product of the number of teeth and the number of phases (e.g., in Example 16.4 it is equal to 2 teeth × 4 phases = 8 steps). Thus, $N = tm$.

The step angle increment, or *resolution*, is equal to $\Delta\theta = 360°/N$. For the motor described in this example,

$$\Delta\theta = \frac{360°}{N} = \frac{360°}{tm} = \frac{360°}{12 \times 3} = 10°$$

EXAMPLE 16.6 Torque Equation of Step Motor

Problem

Calculate the torque generated by a step motor.

Solution

Known Quantities: t = number of teeth per phase; L = axial length of rotor; g = rotor-to- stator radial air gap; r = rotor radius; $\mathcal{F} = \text{mmf}$ developed across the two air gaps (in series) through which a line of flux must pass in one phase. Expression for the motor torque.

Find: Torque developed by the motor.

Schematics, Diagrams, Circuits, and Given Data: $t = 16$ (48 steps, three-phase excitation); $L = 6.35 \times 10^{-3}$ m; $g = 6.35 \times 10^{-5}$ m; $r = 1.29 \times 10^{-2}$ m; $\mathcal{F} = 720$ A-turns.

$$T = 0.314 \times 10^{-6} \frac{tL(r + g/2)\mathcal{F}^2}{g} \text{ N-m}$$

Analysis: Using the expression given above gives

$$\begin{aligned} T &= 0.314 \times 10^{-6} \frac{tL(r + g/2)\mathcal{F}^2}{g} \\ &= 0.314 \times 10^{-6} \frac{16 \times 6.35 \times 10^{-3}(1.29 \times 10^{-2} + 3.175 \times 10^{-5})(720^2)}{6.35 \times 10^{-5}} \\ &= 3.37 \text{ N-m} \end{aligned}$$

CHECK YOUR UNDERSTANDING

Determine the smallest increment in angular position that can be achieved with a PM stepper motor with six stator teeth and three-phase current excitation.

Answer: $\Delta\theta = 20^\circ$

CHECK YOUR UNDERSTANDING

Express the stepping sequence of the variable-reluctance stepping motor of Example 16.4 as a four-digit binary sequence.

Answer:

CHECK YOUR UNDERSTANDING

Express the torque in Example 16.6 in units of pound-inches.

Answer: 29.82 lb-in

From the preceding examples, you should now have a feeling for the operation of variable-reluctance and PM stepping motors. The hybrid configuration is characterized by multitooth rotors that are made of magnetic materials, thus providing a variable-reluctance geometry in conjunction with a permanent-magnet rotor.

An ideal torque–speed characteristic for a stepper motor is shown in Figure 16.13. Two distinct modes of operation are marked on the curve: the **locked-step mode** and the **slewing mode**. In the first mode, the rotor comes to rest (or at least decelerates) between steps; this is the mode commonly used to achieve a given rotor position. In the locked-step mode, the rotor can be started, stopped, and reversed. The slewing mode, on the other hand, does not allow stopping or reversal of the rotor although the rotor still advances in synchronism with the stepping sequence, as described in the preceding examples. The slewing mode can achieve higher continuous speeds than the locked-step mode.

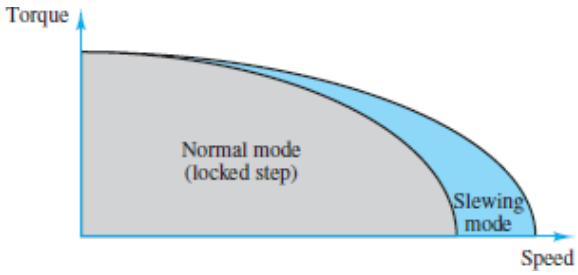


Figure 16.13 Ideal torque–speed characteristic of a stepping motor

Page 16-16

The power supply, or driver, required by a stepping motor is shown in block diagram form in Figure 16.14; it includes a DC power supply, to provide the required current to drive the motor, in addition to logic and switching circuits to provide the appropriate inputs at the right time. One of the important considerations in driving a stepping motor is the excitation mode, which can be one phase or two phase. The driver is the circuit that arranges, distributes, and amplifies pulse trains from the logic circuit determining the stepping sequence; the driver excites each winding of the stepping motor at specified times. In the **one-phase excitation mode**, current is supplied to one phase at a time, with the advantages of low power consumption and good step-angle accuracy. Input signal pulses and the change in the condition of each phase excitation are shown in Figure 16.15. In the **two-phase excitation mode**, current is simultaneously provided to two phases.

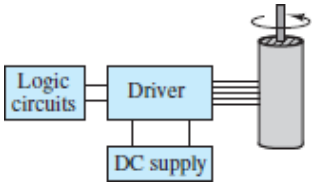
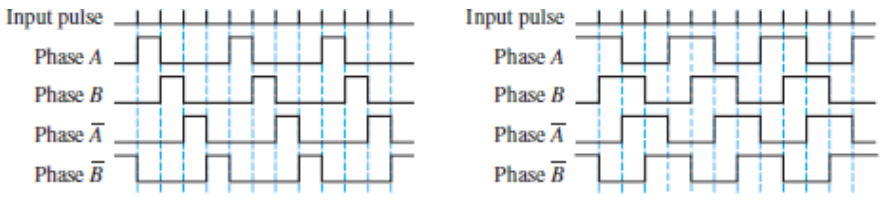


Figure 16.14 Power supply for stepping motor



Input signal pulses and the change in phase excitation of the one-phase excitation mode

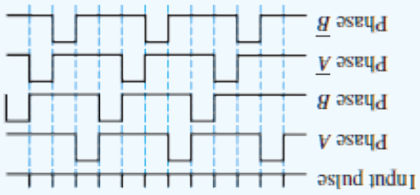
Input signal pulses and the change in phase excitation of the two-phase excitation mode

Figure 16.15 One- and two-phase excitation waveforms for stepper motors

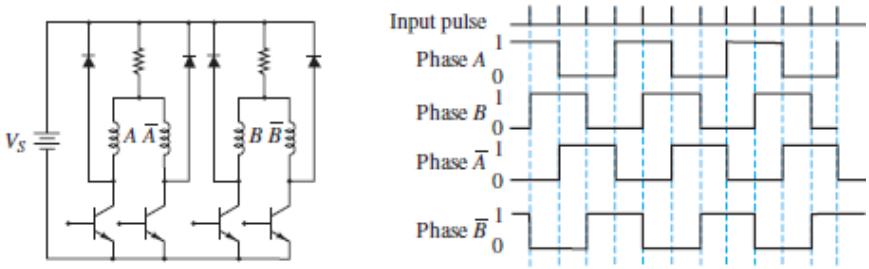
CHECK YOUR UNDERSTANDING

Derive the excitation waveforms corresponding to the direction of rotation opposite to that caused by the stepping sequence shown in Figure 16.15.

Answer:



In addition to the classification of the excitation by phase, stepping motor drives are classified according to whether the drive supplies are unipolar or bipolar, that is, whether they can carry current in one or two directions. Unipolar excitation is clearly simpler although in the case of the two-phase excitation mode, only one-half of the motor windings are used, with a resulting decrease in performance. Figure 16.16 shows a circuit diagram of a unipolar drive and the sequence of phase excitation.



Simplified unipolar drive circuit diagram

Input signal pulses and the change in phase excitation of unipolar drive

Figure 16.16 Unipolar drive for stepper motor

When a bipolar drive is used, motor windings are used effectively, because of the bidirectional exciting current; when operated in this mode, a stepping motor Page 16-17can generate a large output torque at low speed compared with

the unipolar drive. Figure 16.17 shows two versions of the bipolar drive. The first requires two power supplies, one for each polarity, while the second requires only one power supply but needs four switching transistors per phase to reverse the polarity.

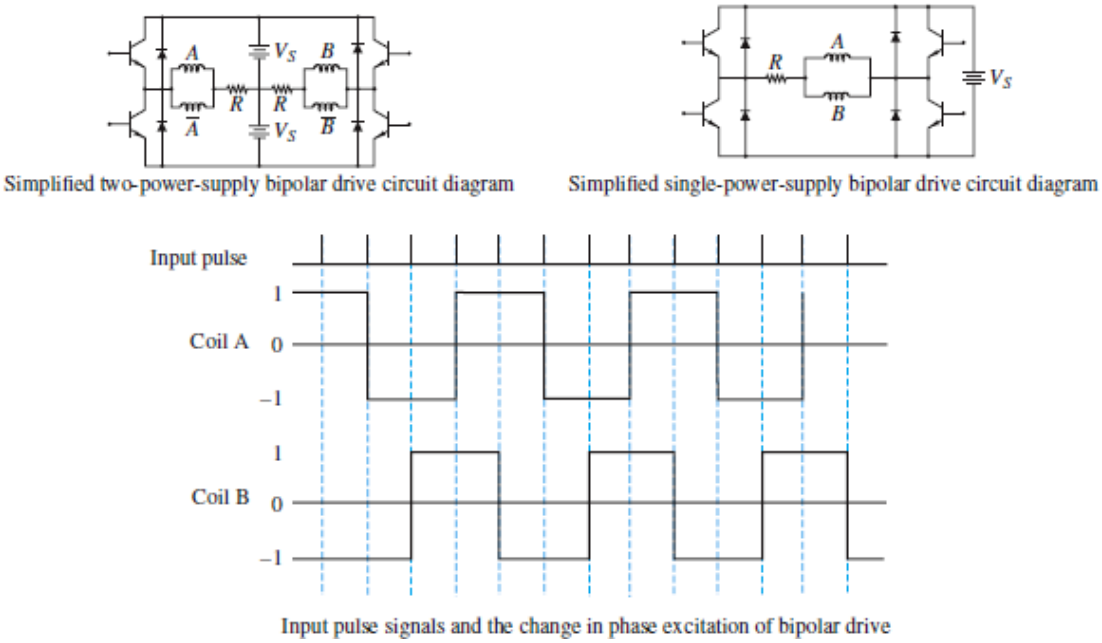


Figure 16.17 Bipolar drive for stepper motors

16.3 SWITCHED RELUCTANCE MOTORS

The switched reluctance (SR) machine is the simplest electric machine that permits variable-speed operation. Today, this machine finds increasingly common application in variable-speed drives for industrial applications and in traction drives for automotive propulsion.

Figure 16.18 depicts the simplest configuration of a reluctance machine and illustrates how the reluctance and inductance of the machine change as a function of position. Note that the magnetic circuit consists only of iron and air —no permanent magnets are required! Note also that the rotor is a salient pole iron element, Page 16-18which is the lowest-cost rotor that can be manufactured. When a current is supplied to the coil, the rotor will experience a torque seeking to align it with the magnetic poles of the stator; when $\theta = 0$, the torque is zero and the rotor will no longer move, having reached its minimum reluctance position. Note that minimum reluctance corresponds to minimum stored energy in the system. Thus, the torque in the motor is developed because

of the change in reluctance with rotor position. This principle makes the reluctance machine different from all other (AC or DC) machines discussed so far. Note also that this machine is one of a few machines, along with the induction motor and VR step motor, to be *singly excited*, that is, to have a single source of magnetic field (whether generated by a coil or by a permanent magnet). One can think of the basic reluctance machine as a salient pole synchronous machine without any field excitation.

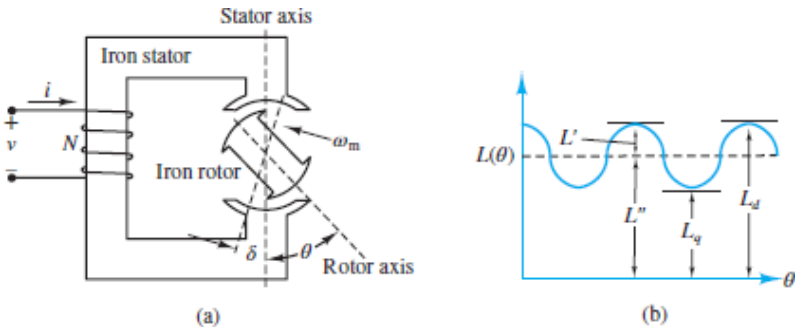


Figure 16.18 (a) Basic reluctance machine and (b) inductance variation as a function of position

The *switched* reluctance machine is a special variation of the simple reluctance machine shown in Figure 16.18 that relies on continuous switching of currents in the stator to guarantee motion of the rotor. It is also a true reluctance machine in that it has *salient poles* both in the rotor and in the stator. The configuration of a typical SR machine is shown in Figure 16.19. Note that Page 16-19the configuration of the SR machine is very similar to that of a VR step motor, discussed in Section 16.2. The primary difference between the two is that the SR machine is designed for continuous and not stepped (discrete) motion. The advent of low-cost power semiconductors, especially GTOs, IGBTs, and power MOSFETs (see Chapter 17) has made it possible to reliably control SR machines. With reference to Figure 16.19, you can see that the stator of an SR machine is wound through slots, with simple solenoid-type windings, and is similar to that of an induction or synchronous AC machine. This stator can be excited by any multiphase source, such as the three-phase sources described in Chapter 15. The SR machine is excited by discrete current pulses that must be timed with respect to the relative position of the rotor poles to the

stator poles, thus requiring **position feedback**. The speed of the rotor is determined by the switching frequency of the stator coil currents.

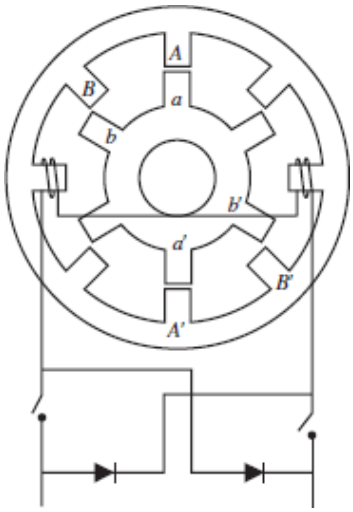


Figure 16.19 Configuration of switched reluctance machine

Operating Principles of SR Machine

Torque production in an SR machine depends on the variation in stored magnetic energy as a function of position. Consider the simple reluctance machine of Figure 16.18, and assume that the variation in winding inductance with rotor position is sinusoidal. The inductance will vary at twice the excitation frequency, because of the number of poles:

$$L(\theta) = L'' + L' \cos 2\theta \tag{16.3}$$

We determine the torque generated by the machine, given the excitation current:

$$i(t) = I_m \sin(\omega t) \tag{16.4}$$

The magnetic stored energy (see Chapter 14) is given by

$$W_m = \tfrac{1}{2} L(\theta) i^2(t) \tag{16.5}$$

and the flux linkage is

$$\lambda(\theta) = L(\theta) i(t) \tag{16.6}$$

We know that the torque can be written as follows:

$$T_m = -\frac{\partial W_m}{\partial \theta} + i\frac{\partial \lambda}{\partial \theta} = -\frac{1}{2}i^2\frac{\partial L}{\partial \theta} + i^2\frac{\partial L}{\partial \theta} = \frac{1}{2}i^2\frac{\partial L}{\partial \theta} \tag{16.7}$$

Given the known sinusoidal current and inductance variations, we can write the torque expression as

$$T_m = -I_m^2 L' \sin(2\theta)\sin^2(\omega t) \tag{16.8}$$

It can be shown, with the use of trigonometric identities, that if the rotor rotates at angular velocity ω_m, such that $\theta = \omega_m t - \theta_0$ (with θ_0 equal to the rotor position at $t = 0$), the torque of the SR machine will be nonzero only if the frequency of the sinusoidal stator current is $\omega = \omega_m$. If the electrical Page 16-20frequency is synchronous with the mechanical frequency, then the average torque is given by

$$\langle T_m \rangle = \tfrac{1}{4}I_m^2 L' \sin(2\theta_0) = \tfrac{1}{8}I_m^2 (L_d - L_q)\sin(2\theta_0) \tag{16.9}$$

We can draw some conclusions from this simplified analysis of the SR machine:

1. The reluctance machine develops an average torque only at one particular (synchronous) speed $\omega = \omega_m$. Thus, the reluctance machine is a synchronous machine.
2. The torque developed by the machine is proportional to $L_d - L_q$ and is therefore dependent on the amplitude of the variation in inductance (or reluctance); thus, this torque is called **reluctance torque**. The values L_d and L_q are called **direct axis inductance** and **quadrature axis inductance**, respectively.
3. The torque varies with angle θ_0, which is therefore equivalent to the *torque angle* δ defined in Chapter 15 for synchronous machines. The maximum torque occurs at $\theta_0 = \pi/4$ and is called the **pull-out torque**.

The above equations have been derived for a continuous reluctance machine; a switched reluctance machine has discontinuous currents, and will therefore have nonsinusoidal reluctance (inductance) variations and a

discontinuous torque. Figure 16.20 depicts the typical appearance of the L and T_m curves for an SR machine. It can be shown that the magnetic torque generated by an SR machine may be expressed as

$$\langle T_m \rangle = \frac{1}{4\pi}(KmP)\tilde{I}^2(L_{\max} - L_{\min}) \qquad \text{SR machine torque} \tag{16.10}$$

where $\tilde{I}$ is the root-mean-square (rms) current, P is the number of pulses per revolution, m is the number of phases, $L_{\max}$ and $L_{\min}$ are the maximum and minimum inductances seen by the exciting coils, and K is a physical constant. Note that the rms value in equation 16.10 includes also the higher harmonics.

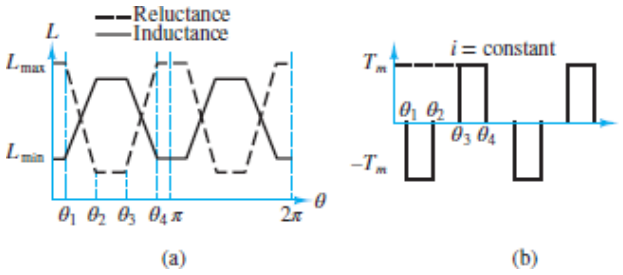


Figure 16.20 Inductance and torque variation in switched reluctance machine

Page 16-21

16.4 SINGLE-PHASE AC MOTORS

In Chapter 15, two types of AC machines were discussed: synchronous and induction. In the discussion of these devices, especially in their motor applications, three-phase excitation was assumed; however, in many practical applications—and especially in small household appliances and small industrial motors—three-phase sources are not readily available, and it would be desirable to use single-phase excitation. Unfortunately, single-phase power does not lend itself to the generation of a rotating magnetic field: single-phase currents in the winding of an AC machine lead to a magnetic field that pulsates in amplitude but does not rotate in space. Thus, it would not be possible to use the AC

machines described in Chapter 15 if only single-phase power were available. This section discusses the construction and the operating and performance characteristics of single-phase AC motors. The discussion focuses mainly on the **universal motor** and **single-phase induction motors**.

Fractional-horsepower (as opposed to **integral-horsepower**) **motors** represent by far the major share of all electric motors. Many fractional-horsepower motors are designed for single-phase use, since single-phase AC power is readily available practically anywhere. Many applications are related to household appliances: refrigerator compressors, air conditioners, fans, electric tools, washer and dryer motors, and others. For the rest of this chapter, we shall examine qualitatively the principle of operation of single-phase motors and look at a few applications. The variety of designs for practical single-phase motors is such that it would not be possible to present the detailed principles of operation for all common types. However, it is hoped that the introduction provided in this chapter will help you in decoding the manufacturer's specifications for a given motor, and in making a preliminary selection for a given application.

Fractional-Horsepower Motors

A *small motor*, as defined by the American Standards Association (ASA) and the National Electrical Manufacturers Association (NEMA; see Chapter 15), is a "motor built in a frame smaller than that having a continuous rating of 1 hp, open type, at 1700 to 1800 r/min." Small motors are generally considered *fractional-horsepower motors*. However, since the determination is based on frame size and on a given speed range, the classification of a motor is not always obvious. Let us give two examples.

1. Consider a a $\frac{3}{4}$-hp, 1,200 r/min motor. This motor is not considered a fractional-horsepower motor, because of its frame size. If the same frame size were used for an 1,800 r/min motor, it would produce a rating of more than 1 hp. Thus, it is considered an integral-horsepower motor of

$$0.75 \text{ hp} \times \frac{1{,}800}{1{,}200} = 1.125 \text{ hp}$$

In other words, since the motor is capable of integral-horsepower performance at speeds of 1,700 to 1,800 r/min, it is classified as an

integral-horsepower motor.

2. Consider now a 1.25-hp, 3,600 r/min motor. This motor is classified as a fractional-horsepower motor, in spite of the fact that its power output is actually greater than 1 hp. If the same motor were used at a speed of 1,800 r/min, it would produce a rating of less than 1 hp:

$$1.25 \times \frac{1{,}800}{3{,}600} = 0.625 \text{ hp}$$

Thus, we see once again that some attention must be paid to the speed of operation of the motor in determining its classification. The term *fractional horsepower* relates more to the physical size of the machine than to the actual power output rating.

Page 16-22

The Universal Motor

If it were possible to operate a DC motor from a single-phase AC supply, a wide range of simple applications would become readily available. Recall that the direction of the torque produced by a DC machine is determined by the direction of current flow in the armature conductors and by the polarity of the field; torque is developed in a DC machine because the commutator arrangement permits the field and armature currents to remain in phase, thus producing torque in a constant direction. A similar result can be obtained by using an AC supply, and by connecting the armature and field windings in series, as shown in Figure 16.21. A series DC motor connected in this configuration can therefore operate on a single-phase AC supply, and it is referred to as a **universal motor**. An additional consideration is that, because of the AC excitation, it is necessary to reduce AC core losses by laminating the stator; thus, the universal motor differs from the series DC motor discussed in Chapter 15 in its construction. As shown in Figure 16.21, the load current is sinusoidal and therefore reverses direction each half-cycle; however, the torque generated by the motor is always in the same direction, resulting in a pulsating torque, with nonzero average value. Typical torque–speed curves for AC and DC operation of a universal motor are shown in Figure 16.22.

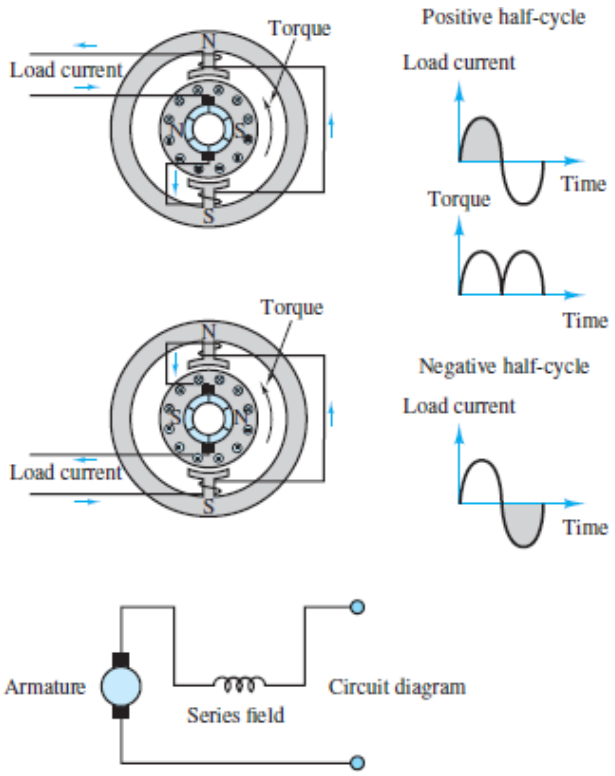


Figure 16.21 Operation and circuit diagram of a universal motor

As in the case of a DC series motor, the best method for controlling the speed of a universal motor is to change its (rms) input voltage. The higher the rms input voltage, the greater the resulting speed of the motor. Approximate torque–speed characteristics of a universal motor as a function of voltage are shown in Figure 16.23.

Page 16-23

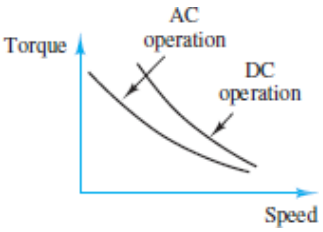


Figure 16.22 Torque–speed curve of a universal motor

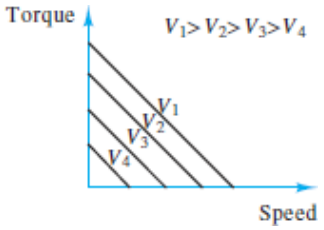


Figure 16.23 Torque–speed characteristics of a universal motor

EXAMPLE 16.7 Analysis of Universal Motor

Problem

Find the following quantities for a universal motor:

1. Back emf
2. Power output
3. Shaft torque
4. Motor efficiency

Solution

Known Quantities: Motor operating data and circuit parameters.

Find: $\mathbf{E}_b$; P_{out}; T_{out}; η (efficiency).

Schematics, Diagrams, Circuits, and Given Data: Motor operating data: 120 V; 60 Hz; two poles; 800 r/min; 17.85 A (full load); pf = 0.912 (lagging). Circuit parameters: $R_f = 0.65\ \Omega$; $X_f = 1.2\ \Omega$; $R_a = 1.36\ \Omega$; $X_a = 1.6\ \Omega$.

Assumptions: Use the circuit model for the series motor described in Chapter 15. The rotational losses amount to 80 W.

Analysis: The circuit model for the series machine is shown in Figure 16.24. We shall use this model with the understanding that all currents and voltages are now phasors.

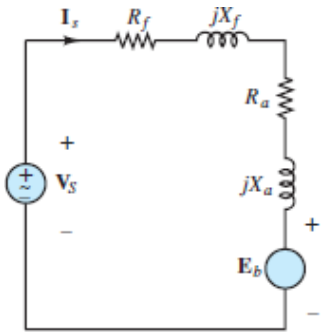


Figure 16.24 Equivalent circuit of a universal motor

1. *Back emf computation.* To determine the back emf, we need to calculate the voltage across the armature coil and subtract it from the supply voltage:

$$\begin{aligned}\mathbf{E}_b &= \mathbf{V}_S - \mathbf{I}_S(R_f + jX_f + R_a + jX_a)\\ &= \mathbf{V}_S - (I_S\angle\theta)(R_f + jX_f + R_a + jX_a)\end{aligned}$$

The impedance angle is the only unknown quantity, and it may be found from the power factor:

$$\text{pf} = \cos(\theta) = 0.912 \text{ (lagging)} \qquad \theta = \arccos(0.912) = -24.22°$$

Thus:

$$\begin{aligned}\mathbf{E}_b &= \mathbf{V}_S - (I_S\angle\theta)(R_f + jX_f + R_a + jX_a)\\ &= 120 - (17.85\angle -24.22°)(0.65 + j1.2 + 1.36 + j1.6)\\ &= 73.57\angle -24.8° \text{ V}\end{aligned}$$

Page 16-24

2. *Output power calculation.* The total power developed by the motor is equal to the product of the back emf and the series current:

$$P_{\text{total}} = E_b I_S = 73.56 \times 17.85 = 1{,}313.22 \text{ W}$$

The mechanical (output) power of the motor is the difference between the total power and the rotational losses:

$$P_{\text{out}} = P_{\text{total}} - P_{\text{rot}} = 1{,}313.22 - 80 = 1{,}233.22 \text{ W}$$

3. *Output (shaft) torque calculation.* The output torque is equal to the ratio of output power to shaft speed:

$$T_{out} = \frac{P_{out}}{\omega} = \frac{1{,}233.15 \text{ W}}{(2\pi \times 800/60) \text{ rad/s}} = 14.72 \text{ N-m}$$

4. *Efficiency calculation.* The efficiency of the motor is defined as the ratio of output power to input power:

$$\eta = \frac{P_{out}}{P_{in}} = \frac{P_{out}}{V_S I_S \cos\theta} = \frac{1{,}233.15}{1{,}953.5} = 63.12\%$$

Comments: Note that the analysis of this machine is very similar to that of the series DC motor, except for the use of phasors. It is very important to notice that in calculating the input power, one has to consider the power factor of the motor to obtain the *real power.*

EXAMPLE 16.8 Universal Motor Torque Expression

Problem

Compute an expression for the average torque generated by a universal motor, based on the circuit diagram of Figure 16.24.

Solution

Known Quantities: Circuit model of motor.

Find: Expression for average torque T_{av}.

Assumptions: The motor operates in the linear region of the magnetization curve.

Analysis: With reference to Chapter 15, we know that the flux produced in a series motor by the series current $i_S(t)$ is $\phi = k_S i_S(t)$. The instantaneous torque produced by the machine is given by

$$T(t) = k_T \phi(t) i_S(t)$$

If the source waveform has period $\tau = 2\pi/\omega$, we can calculate the average power by integrating the instantaneous torque over one period:

$$T_{av}(t) = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} k_T k_S i_S^2(t')\,dt' = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} k_T k_S I_S^2(\sin^2 \omega t')\,dt' = \tfrac{1}{2} k_T k_S I_S^2$$

where I_S is the peak value of the (series) armature current.

Comments: The series motor can produce a nonzero average torque when excited by an alternating current because of the quadratic nature of the instantaneous torque. A permanent-magnet DC machine, which has a linear torque–current relationship, would generate zero average torque if driven from an AC supply.

Page 16-25

Single-Phase Induction Motors

A typical single-phase induction motor bears close resemblance to the polyphase squirrel cage induction motor discussed in Chapter 15, the major difference being in the configuration of the stator winding. A simplified schematic diagram of such a motor, with a single winding, is shown in Figure 16.25; the winding is typically distributed around the stator so as to produce an approximately sinusoidal mmf.

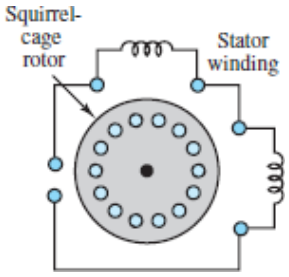


Figure 16.25 Single-phase induction motor

Assume that the mmf for a practical motor can be generated so as to approximate the following function:

$$\mathcal{F} = F_{\max} \cos(\omega t) \cos(\theta_m)$$

This function can be written as the sum of two components, as follows:

$$\mathcal{F}^+ = \tfrac{1}{2}F_{\max}\cos(\theta_m - \omega t)$$
$$\mathcal{F}^- = \tfrac{1}{2}F_{\max}\cos(\theta_m + \omega t)$$

These two components may be interpreted as representing two mmf waves traveling in opposite directions around the stator. Each mmf produces torque according to the induction principles described in Chapter 15; however, the two components are equal and opposite, and no net torque results if the rotor is at rest. The resulting mmf is pulsating (i.e., changing in amplitude), but not rotating in space, as it would be in a polyphase stator. If the rotor is made to turn in either direction, however, the two mmf's will not be equal any longer because the motion of the rotor will induce an additional mmf, which will add to one of the two mmf's and subtract from the other. Thus, a net torque will be established, causing the motor to continue its rotation in the same direction in which it was started. In particular, if the rotor is started in the forward direction, the forward mmf $\mathcal{F}^+$ will be greater than the backward mmf, and the motor will continue to rotate in the forward direction.

Figure 16.26 depicts an equivalent circuit for the single-phase induction motor *with stationary rotor*, where the parameters in the circuit are defined as follows:

R_S = resistance of stator winding

X_S = leakage reactance of stator winding

X_m = magnetizing reactance of stator winding

X_R = leakage reactance of rotor referred to stator at standstill

R_R = leakage resistance of rotor referred to stator at standstill

E_b = voltage induced in stator winding by (stationary) pulsating flux in air gap

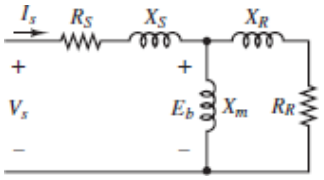


Figure 16.26 Circuit model for single-phase induction motor with rotor at standstill

Figure 16.27 depicts the equivalent circuit for the same motor with the rotor rotating with slip s. Note that the circuit is asymmetric, because of the different air gap flux forward and backward components E_f and E_b, respectively. The factors of 0.5 come from the resolution of the pulsating stator mmf into forward and backward components. Note further that the reflected rotor impedance is asymmetric because of the presence of the slip parameter in the expression for the reflected rotor resistance. Further, the circuit model also confirms that the forward induced voltage E_f must be greater than the backward voltage E_b since the slip is always less than 1.

Page 16-26

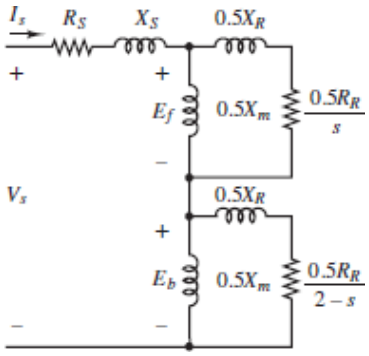


Figure 16.27 Circuit model for single-phase induction motor with rotor in motion

It can be shown that the torque components in the forward and backward directions are given by

$$T_f = \frac{P_f}{\omega_s} \tag{16.11}$$

and:

$$T_b = \frac{P_b}{\omega_s} \tag{16.12}$$

where ω_s is the synchronous speed and:

$$P_f = I_s^2 R_f \tag{16.13}$$

Here, R_f is the resistive component of the forward field impedance; also:

$$P_b = I_s^2 R_b \tag{16.14}$$

where R_b is the resistive component of the backward field impedance. Since the torque produced by the backward field is in the opposite direction to that produced by the forward field, the net torque will consist of the difference between the two:

$$T = T_f - T_b = \frac{I_s^2(R_f - R_b)}{\omega_s} \tag{16.15a}$$

The mechanical power developed by the motor is

$$\begin{aligned} P_{\text{mech}} &= T\omega_m = T\omega_s(1-s) = (P_f - P_b)(1-s) \\ &= I_s^2(R_f - R_b)(1-s) \end{aligned} \tag{16.15b}$$

EXAMPLE 16.9 Slip in a Single-Phase Induction Motor

Problem

Find the slip of the field in the forward and backward directions for a single-phase induction machine.

Solution

Known Quantities: Motor operating characteristics.

Page 16-27

Find: Forward slip s_f; backward slip s_b.

Schematics, Diagrams, Circuits, and Given Data: Motor operating characteristics: 115 V; 60 Hz; four poles; 1,710 r/min.

Analysis: We first determine the synchronous speed of the motor:

$$n_s = \frac{120f}{p} = \frac{120 \times 60}{4} = 1{,}800 \text{ r/min}$$

The slip in the forward direction (direction of rotation of the motor) can now be computed:

$$s_f = \frac{n_s - n}{n_s} = \frac{1{,}800 - 1{,}710}{1{,}800} = 0.05 = 5\%$$

The slip in the backward direction can be computed as follows, with reference to Figure 16.27:

$$s_b = 2 - s_f = 2 - 0.05 = 1.95$$

EXAMPLE 16.10 Analysis of Single-Phase Induction Motor

Problem

Find the input current and generated torque for a single-phase induction motor.

Solution

Known Quantities: Motor operating characteristics and circuit parameters.

Find: Motor input (stator) current I_S; motor torque T.

Schematics, Diagrams, Circuits, and Given Data: Motor operating data: $\frac{1}{4}$ hp; 110 V; 60 Hz; four poles. Circuit parameters: $R_S = 1.5\ \Omega$; $X_S = 2\ \Omega$; $R_R = 3\ \Omega$; $X_R = 2\ \Omega$; $X_m = 50\ \Omega$; $s = 0.05$.

Assumptions: The motor is operated at rated voltage and frequency.

Analysis: With reference to the equivalent circuit of Figure 16.27, you can easily show that the impedance seen by the back emf E_b is much smaller than that seen by the forward emf E_f. This corresponds to stating that the backward component of the magnetizing impedance (which is in parallel with the backward component of the rotor impedance) is much larger than the backward component of the rotor impedance and can therefore be neglected as shown below.

$$0.5Z_b = 0.5\frac{jX_m[R_R/(2-s)+jX_R]}{R_R/(2-s)+j(X_m+X_R)}$$
$$\approx 0.5\left(\frac{R_R}{2-s}+jX_R\right)$$

This approximation is generally true for values of slip less than 0.15. An approximate circuit based on this simplification is shown in Figure 16.28. There, the impedance seen by the backward emf is given by

$$0.5Z_b = 0.5\left(\frac{R_R}{2-s}+jX_R\right) = 0.5(1.538+j2) = 0.5(R_b+jX_b)$$

Page 16-28

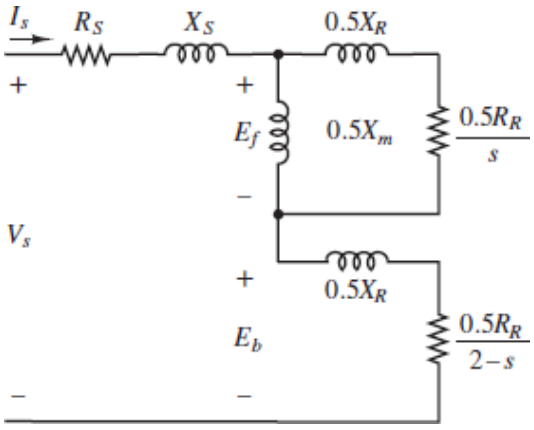


Figure 16.28 Approximate circuit model for single-phase induction motor

The impedance seen by the forward emf is, on the other hand, given by the exact expression:

$$0.5Z_f = 0.5\frac{jX_m(R_R/s+jX_R)}{R_R/s+j(X_m+X_R)}$$
$$= 0.5\frac{j50(60+j2)}{60+j(50+2)} = 11.9+j14.69 = 0.5R_f+j0.5X_f$$

If we let $Z_S = R_S + jX_S = 1.5 + j2\ \Omega$, we can write an expression for the total impedance of the motor as follows:

$$Z = Z_S + 0.5Z_f + 0.5Z_b = 14.169 + j17.69 = 22.66\angle 51.3^\circ\ \Omega$$

Knowing the total series impedance, we can calculate the stator current:

$$\mathbf{I}_S = \frac{\mathbf{V}_S}{Z} = \frac{110\text{ V}}{22.66\angle 51.3°\,\Omega} = 4.85\angle -51.3°\text{ A}$$

We can now calculate the power absorbed by the motor by separately computing the real power absorbed in the forward and backward fields:

$$P_f = I_S^2 \times 0.5 R_f = (4.85)^2 \times 11.9 = 279.9\text{ W}$$
$$P_b = I_S^2 \times 0.5 R_b = (4.85)^2 \times 0.769 = 18.1\text{ W}$$

The net power is the difference between the two components; thus, $P = P_f - P_b$ = 261.8 W, and the torque developed by the motor is equal to the ratio of the power to the motor speed. The synchronous speed can be computed to be

$$\omega_s = \frac{4\pi f}{p} = 188.5\text{ rad/s}$$

and if we assume negligible rotational losses, we have

$$T = \frac{P}{\omega} = \frac{P}{(1-s)\times\omega_s} = \frac{261.8\text{ W}}{0.95 \times 188.5\text{ rad/s}} = 1.46\text{ N-m}$$

Comments: Note that the power factor of the motor is pf = cos (–51.3°) = 0.625. Such low power factors are typical of single-phase motors.

EXAMPLE 16.11 Analysis of Single-Phase Induction Motor Problem

Find the following quantities for the single-phase machine of Example 16.10:

1. Output torque
2. Output power

3. Efficiency

Solution

Known Quantities: Motor operating characteristics.

Find: Motor torque T; output power P_{out}; efficiency η.

Schematics, Diagrams, Circuits, and Given Data: Motor operating data: $\frac{1}{4}$ hp; 110 V; 60 Hz; four poles; $s = 0.05$.

Assumptions: The motor is operated at rated voltage and frequency. The combined rotational and core losses are $P_{rot} + P_{core} = 30$ W.

Analysis:

1. *Output power calculation.* The motor generated power is the difference between the forward and backward components, as explained in Example 16.10. Thus:

$$P = P_f - P_b = 261.8 \text{ W}$$

The motor power is the difference between the generated power and the losses:

$$P_{out} = P - P_{loss} = 261.8 - 30 = 231.8 \text{ W}$$

2. *Shaft torque calculation.* The shaft speed is

$$\omega = (1 - s) \times \omega_s = (1 - s) \times \frac{4\pi f}{p} = 179 \text{ rad/s}$$

and the torque is

$$T = \frac{P_{out}}{\omega} = \frac{231.8 \text{ W}}{179 \text{ rad/s}} = 1.29 \text{ N-m}$$

3. *Efficiency calculation.* To calculate the overall efficiency of the motor, we must account for three loss mechanisms: mechanical losses, core losses, and electrical losses. The first two are given as a lumped number; the electrical losses can be computed by calculating the I^2R losses in the stator and forward and backward circuits:

$$P_{S\,\text{loss}} = I_S^2 R_S = (4.85)^2 \times 1.5 = 35.3\ \text{W}$$
$$P_{R_f\,\text{loss}} = sP_f = 0.05 \times 279.9 = 14\ \text{W}$$
$$P_{R_b\,\text{loss}} = (2-s)P_b = 1.95 \times 18.1 = 35.3\ \text{W}$$
$$P_{\text{elec}} = P_{S\,\text{loss}} + P_{R_f\,\text{loss}} + P_{R_b\,\text{loss}} = 84.6\ \text{W}$$

The efficiency can be calculated according to the following expression:

$$\eta = 1 - \frac{\sum \text{losses}}{P_{\text{out}} + \sum \text{losses}} = 1 - \frac{P_{\text{rot}} + P_{\text{core}} + P_{\text{elec}}}{P_{\text{out}} + P_{\text{rot}} + P_{\text{core}} + P_{\text{elec}}}$$
$$= 1 - \frac{(30 + 84.6)\ \text{W}}{(2131.8 + 30 + 84.6)\ \text{W}} = 0.949 = 94.9\%$$

Page 16-30

Comments: Note that the overall efficiency of this machine is fairly low. Multiphase AC machines can achieve significantly higher efficiencies.

CHECK YOUR UNDERSTANDING

(a) What is the zero-speed torque for a single-phase induction motor? (b) If a starting torque is applied to the machine, what final speed will the machine reach?

Answer: (a) Zero, without a start winding; (b) $n = (1 - s)n_s$, where $n_s = 120f/p$ and s = slip (determined by shaft load).

The equations and circuit models in the preceding examples suggest that a single-phase induction motor is capable of sustaining a torque, and of reaching its operating speed, once it is started by external means. However, because the magnetic field in a single-phase winding is stationary, a single-phase motor is not self-starting. The speed–torque characteristic of a typical single-phase induction motor shown in Figure 16.29 shows that the starting torque for this motor is zero. The curve also shows that the motor can operate in either

direction, depending on the direction of the initial starting torque, which must be provided by separate means.

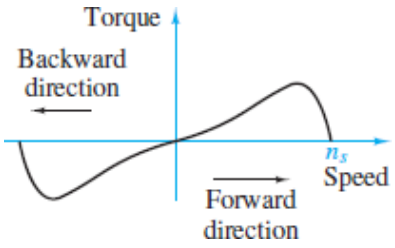


Figure 16.29 Torque–speed curve of a single-phase induction motor

Classification of Single-Phase Induction Motors

Thus far, we have not mentioned how the initial starting torque can be provided to a single-phase motor. In practice, single-phase motors are classified by their starting and running characteristics, and several methods exist to provide nonzero starting torque. The aim of this section is to classify single-phase motors by describing their configuration on the basis of the method of starting. For each class of motor, a torque–speed characteristic is also described.

Split-Phase Motors

Split-phase motors are constructed with two separate stator windings, called **main** and **auxiliary windings**; the axes of the two windings are actually at 90° with respect to each other, as shown in Figure 16.30. The auxiliary winding current is designed to be out of phase with the main winding current, as a result of the different reactances of the two windings. Different winding reactances can be attained by having a different ratio of resistance to inductance, for example, by increasing the resistance of the auxiliary winding. In particular, the auxiliary winding current $\mathbf{I}_{aux}$ leads the main winding current $\mathbf{I}_{main}$. The net effect is that the motor sees a two-phase (unbalanced) current that results in a rotating magnetic field, as in any polyphase stator arrangement. Thus, the motor has a nonzero starting torque, as shown in Figure 16.31. Once the motor has started, a centrifugal switch is used to disconnect the auxiliary winding, since a single winding is sufficient to sustain the motion of the rotor. The switching action permits the use of relatively high-resistance windings, since these are not used during normal operation, and therefore Page 16-31one need not be concerned with the losses associated with a higher-resistance winding. Figure 16.31 also depicts the combined effect of the two modes of operation of the split-phase motor.

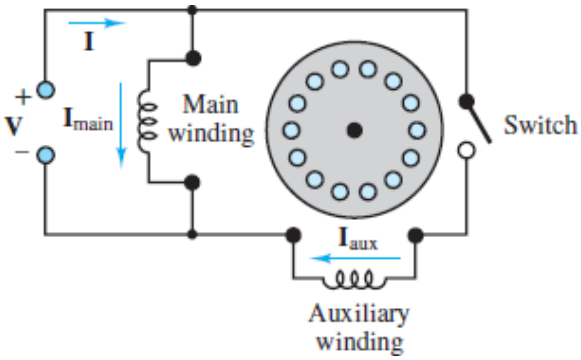


Figure 16.30 Split-phase motor

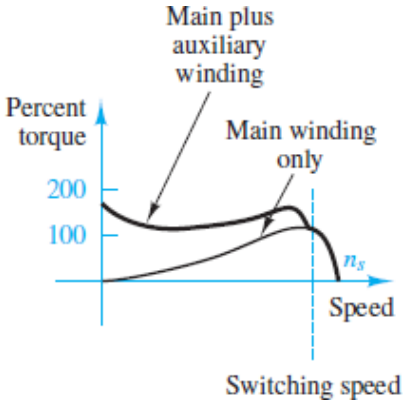


Figure 16.31 Torque–speed curve of split-phase motor

Split-phase motors have appropriate characteristics (at very low cost) for fans, blowers, centrifugal pumps, and other applications in the range of $\frac{1}{20}$ to $\frac{1}{2}$.

Capacitor-Type Motors

Another way to obtain a phase difference between two currents, and thus, a rotating magnetic field is by the addition of a capacitor. Motors that use this arrangement are termed **capacitor-type motors**. These motors make different use of capacitors to provide starting or running capabilities, or a combination of the two. The **capacitor-start motor** is essentially identical to the split-phase motor, except for the addition of a capacitor in series with the auxiliary winding, as shown in Figure 16.32. The addition of the capacitor changes the reactance of the auxiliary circuit in such a way as to cause the auxiliary current to lead the main current. The advantage of using the capacitor to achieve a phase split is that greater starting torque may be obtained than with the split-phase arrangement. A centrifugal switching arrangement is used to disconnect the

auxiliary winding above a certain speed, in the neighborhood of 75 percent of synchronous speed.

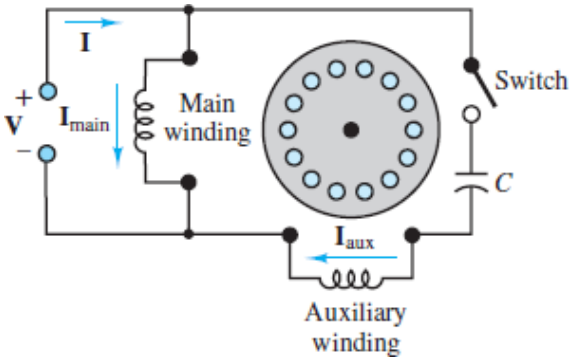


Figure 16.32 Capacitor-start motor

Figure 16.33 depicts the torque–speed characteristic of a capacitor-start motor. Because of their higher starting torque, these motors are very useful in Page 16-32connection with loads that present a high static torque. Examples of such loads are seen in compressors, pumps, and refrigeration and air-conditioning equipment.

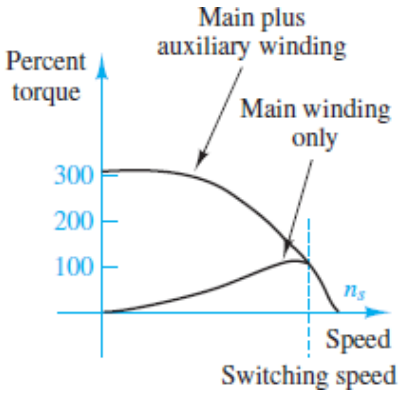


Figure 16.33 Torque–speed curve for a capacitor-start motor

It is also possible to use the capacitor-start motor without the centrifugal switch, leading to a simpler design. Motors with this design are called **permanent split-capacitor motors**; they offer a compromise between running and starting characteristics. A typical torque–speed curve is shown in Figure 16.34.

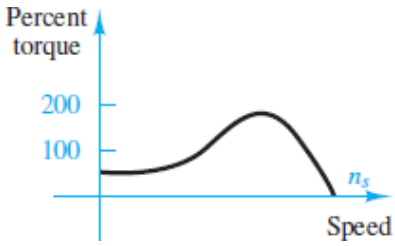


Figure 16.34 Torque–speed curve for a permanent split-capacitor motor

A further compromise can be achieved by using two capacitors—one to obtain a permanent phase split and the resulting improvement in running characteristics, the other to improve the starting torque. A small capacitance is sufficient to improve the running performance, while a much larger capacitor provides the temporary improvement in starting torque. A motor with this design is called a **capacitor-start capacitor-run motor**; its schematic diagram is shown in Figure 16.35. Its torque–speed characteristic is similar to that of a capacitor-start motor.

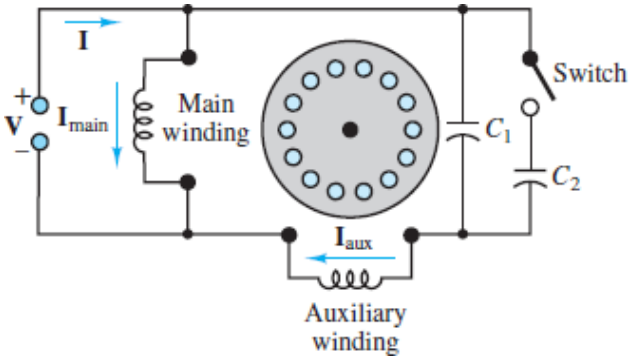


Figure 16.35 Capacitor-start capacitor-run motor

Shaded-Pole Motors

The last type of single-phase induction motor discussed in this chapter is the **shaded-pole motor**. This type of motor operates on a different principle from the motors discussed thus far. The stator of a shaded-pole motor has a salient pole construction, as shown in Figure 16.36, that includes a shading coil consisting of a copper band wound around part of each pole. The flux in the shaded portion of the pole lags behind the flux in the unshaded part, achieving an effect similar to a rotation of the flux in the direction of the shaded part of the pole. This flux rotation in effect produces a rotating field that enables the motor

to have a starting torque. This construction technique is rather inexpensive and is used in motors up to about $\frac{1}{20}$ hp.

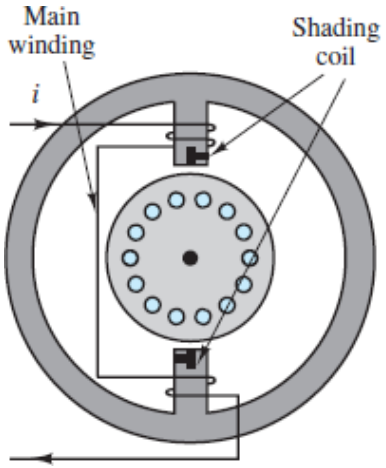


Figure 16.36 Shaded-pole motor

A typical torque–speed characteristic for a shaded-pole motor is given in Figure 16.37.

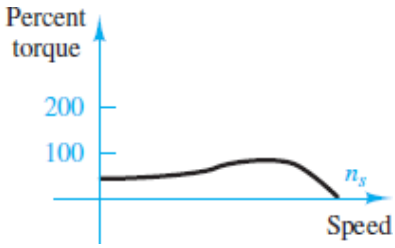


Figure 16.37 Torque–speed curve of a shaded-pole motor

LO

Summary of Single-Phase Motor Characteristics

Four basic classes of single-phase motors are commonly used:

1. Single-phase induction motors are used for the larger home and small business tasks, such as furnace oil burner pumps, or hot water or hot air circulators. Refrigerator compressors, lathes, and bench-mounted circular saws are also powered with induction motors.

2. Shaded-pole motors are used in the smaller sizes for quiet, low-cost applications. The size range is from $\frac{1}{30}$ hp (24.9 W) to $\frac{1}{2}$ hp (373 W), particularly for fans and similar drives in which the starting torque is low.
3. Universal motors will operate on any household AC frequency or on direct current without modification or adjustment. They can develop very high speed while loaded and very high power for their size. Vacuum cleaners, sewing machines, kitchen food mixers, portable electric drills, portable circular saws, and home motion-picture projectors are examples of applications of universal motors.
4. The capacitor-type motor finds its widest field of application at low speeds (below 900 r/min) and in ratings from $\frac{3}{4}$ hp (0.5595 kW) to 3 hp (2.238 kW) at all speeds, especially in fan drives.

EXAMPLE 16.12 Analysis of Capacitor-Start Motor

Problem

With reference to Figure 16.32, find the required starting capacitance.

Solution

Known Quantities: Motor operating characteristics; motor circuit parameters.

Find: Starting capacitance C.

Schematics, Diagrams, Circuits, and Given Data: Motor operating data: $\frac{1}{3}$ hp; 120 V; 60 Hz. Circuit parameters: $R_{main} = 4.5\ \Omega$; $X_{main} = 3.7\ \Omega$; $R_{aux} = 9.5\ \Omega$; $X_{aux} = 3.5\ \Omega$.

Analysis: The purpose of the starting capacitor is to cause the auxiliary winding current $\mathbf{I}_{aux}$ at standstill to lead the main winding current $\mathbf{I}_{main}$ by 90°. The 90° phase lead will provide the maximum starting torque. Figure 16.38 shows the

phasor diagram for these two currents and the voltage. The impedance angle of the main winding is

$$\theta_{\text{main}} = \arctan\left(\frac{X_{\text{main}}}{R_{\text{main}}}\right) = \arctan\left(\frac{3.7\ \Omega}{4.5\ \Omega}\right) = 39.4^\circ$$

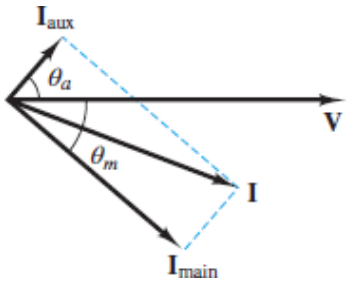


Figure 16.38 Starting phasor diagram for capacitor-start motor

Knowing that the desired phase shift between the main and auxiliary impedance angles is –90°, we compute the impedance angle of the auxiliary winding:

$$\theta_{\text{aux}} = 39.4^\circ - 90^\circ = -50.6^\circ$$

The minus sign indicates that $\mathbf{I}_{\text{aux}}$ leads the terminal voltage. The required capacitance can now be calculated from the relationship:

$$\arctan\left(\frac{X_{\text{aux}} - X_C}{R_{\text{aux}}}\right) = -50.6^\circ$$

$$X_C = -R_{\text{aux}} \times \tan(-50.6^\circ) + X_{\text{aux}} = -9.5 \times (-1.21) + 3.5 = 15.07\ \Omega$$

Page 16-34

and we can compute the desired capacitance as

$$C = \frac{1}{\omega X_C} = \frac{1}{377 \times 15.07} = 176 \times 10^{-6}\,\text{F} = 176\ \mu\text{F}$$

EXAMPLE 16.13 Split-Phase Motor Nameplate Analysis

Problem

The table below depicts a split-phase motor nameplate. Determine the following quantities, using nameplate data:

1. Rated slip
2. Synchronous speed
3. Rated torque

Solution

Known Quantities: Nameplate data.

Find: s; ω_S; T.

Schematics, Diagrams, Circuits, and Given Data:

Thermal protected				Split-phase fan and blower motor			
MOD	4k800	HP	$\frac{1}{3}$	RPM	1,725	KVA CODE	N
V	115			A.	5.5		
FR	48Y	Hz	60	PH	1	DUTY	CONT.
INS.CL.	B	MAX. AMB	40 C	S. F.	1.35	BRG	SLEEVE

Analysis: An explanation of the nameplate for a typical electric motor was given in Chapter 15. This example focuses on a few specific items of interest in the case of a split-phase motor. As you can see, the nameplate directly indicates the split-phase motor classification. Following the hertz designation is the phase information. AC systems may have one, two, or three phases. Single-phase and three-phase systems are the most common.

The code letter following KVA CODE indicates the locked-rotor kilovoltamperes per horsepower, as explained in *NEMA Motor and Generator Standards*, NEMA Publication No. MG 1-10.37. The symbol "N" means that this motor has a maximum locked-rotor kilovoltamperes per horsepower of 12.5. Since the motor is rated at $\frac{1}{3}$ hp, the maximum locked-rotor kilovoltampere value is 12.5/3 = 4.167. The maximum locked-rotor ampere value at 115 V will be 4.167 kVA/115 V = 36.23 A.

A large percentage of fractional-horsepower motors are now provided with built-in thermal protection. The use of such protection will also be indicated in the motor nameplate, for example, here by "Thermal Protected."

Bearing is abbreviated as "BRG." Fractional-horsepower motors normally use one of two types of bearings: sleeve or ball.

Page 16-35

A variety of additional information may appear on the nameplate. This may include instructions for connecting the motor to a source of supply, reversing the direction of rotation, lubricating the motor, or operating it safely.

For the machine in this example, the synchronous speed is

$$n_s = 1{,}800 \text{ r/min}$$

The slip at rated speed is

$$s = \frac{n_s - n}{n_s} = \frac{1{,}800 - 1{,}725}{1{,}800} = 0.042$$

The power is

$$P = \tfrac{1}{3}\text{hp} = \tfrac{1}{3} \times 746 \frac{\text{W}}{\text{hp}} = 248.7 \text{ W}$$

The rated torque is

$$T = \frac{K \times P}{n}$$

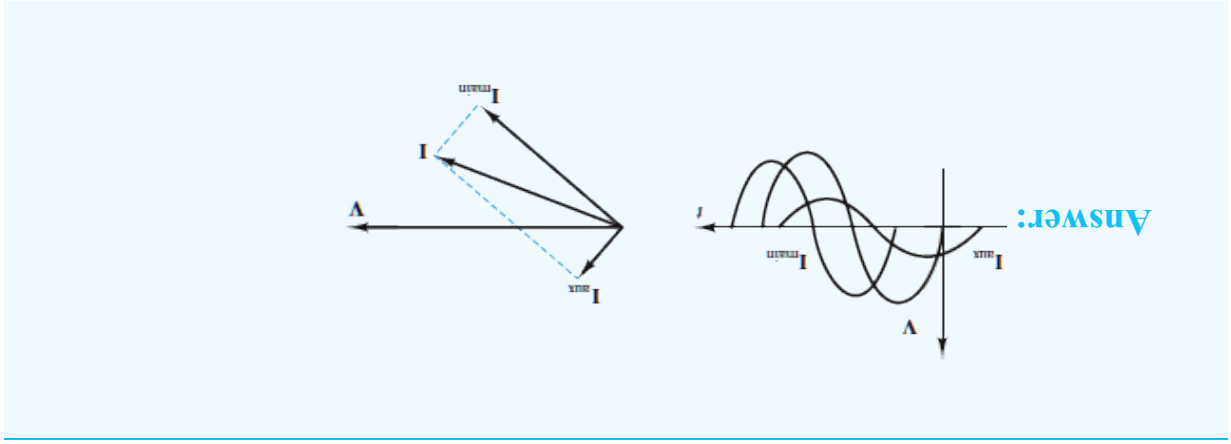
where the constant K is given by

$$K = \begin{cases} 0.97376 & T \text{ in meter-kilograms} \\ 9.549 & T \text{ in newton-meters} \end{cases}$$

$$T = 9.549 \times \frac{248.7}{1{,}725} = 1.377 \text{ N-m}$$

CHECK YOUR UNDERSTANDING

Draw the starting phasor diagram relating **V**, **I**, $\mathbf{I}_{main}$, and $\mathbf{I}_{aux}$ for the circuit of Figure 16.36, and sketch the time-domain waveforms.



16.5 MOTOR SELECTION AND APPLICATION

The objectives of this section are to outline the selection process of a motor for application to an electric drive and to summarize the characteristics of the most common drive motors, with emphasis on fractional-horsepower applications. An Page 16-36electric motor should satisfy a set of precise requirements to be considered for a specific application. These include

1. Starting characteristics (torque and current).
2. Acceleration characteristics (dependent on the load).
3. Efficiency at rated load.
4. Overload capability.
5. Electrical and thermal safety.
6. Cost.

These requirements suggest that the specific details of the application should be clear in the designer's mind. For example, the nature of the load, the available electrical supplies, and the ambient conditions should be carefully investigated before the motor selection process is initiated. Once the application environment is known, it is usually possible to narrow the selection of a drive motor to a few choices. In this section we provide some insight into the motor selection process.

Motor Performance Calculations

To better understand the motor selection process, it is important to review some of the basic ideas underlying the motion of the motor and of the load. For rotational systems, the relationships of Table 16.5 hold. Figure 16.39 summarizes the various types of load profiles that are likely to be encountered in practical applications. These include constant-torque loads; viscous friction-type loads, where torque is proportional to speed; loads in which the torque is proportional to a power of speed (e.g., fans, pumps); and constant-power loads, where torque is inversely proportional to speed.

Table 16.5 Equations of motion and definitions of variables

Equations of motion	Definition of terms
$\omega_2 = \omega_1 + \alpha t$	ω_1 = initial velocity (rad/s)
$\theta = \omega_1 t + \frac{1}{2}\alpha t^2$	ω_2 = final velocity (rad/s)
$\omega_2^2 = \omega_1^2 + 2\alpha\theta$	α = acceleration (rad/s^2)
$P = T\omega$	θ = angular displacement (rad)
$T = Fr = J\alpha$	n = speed of rotation (r/min)
$W = T\theta$	P = output power (W)
	W = work (J)
$\omega = \frac{2\pi n}{60} = 0.105n$	F = force (N)
	T = torque (N-m)
$J = mk^2$	J = polar moment of inertia (kg-m^2)
	r = radius of arm (m)
	k = radius of gyration (m)
	m = mass (kg)

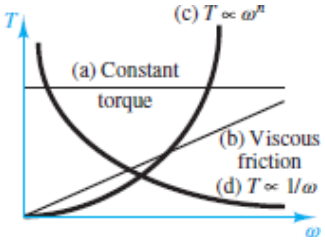


Figure 16.39 Typical load torque–speed curves

Reflected Load Inertia Calculations

To calculate the motor requirements, one must compute the required torque referenced to the motor output shaft. Since gearing systems are often employed, the Page 16-37inertias of all rotating components must be referred to the motor shaft. Using the terminology of Table 16.5, we then conclude that the reflected load torque at the motor shaft T_r is related to the load torque T_L by the relationship:

$$T_r = \frac{\omega_L}{\omega_r} T_L \tag{16.16}$$

where $\omega_r = \omega_m$ is the motor shaft speed, and the ratio of load speed to motor speed is equal to the gear ratio:

$$\frac{\omega_L}{\omega_r} = \frac{n_L}{n_r} \tag{16.17}$$

If we equate the kinetic energy on the motor side to that on the load side, we can also derive an expression for the reflected load inertia:

$$J_L \omega_L^2 = J_r \omega_r^2 \tag{16.18}$$

or:

$$J_r = J_L \left(\frac{n_L}{n_r}\right)^2 \quad \text{Reflected inertia} \tag{16.19}$$

Thus, the reflected inertia seen by the motor at the shaft is equal to the load inertia times the square of the gear ratio. Note that this is a mechanical impedance transformation similar to that used in the case of transformers. For all practical purposes, one can think of a gearing system as a mechanical transformer that, in the ideal case, conserves power. Under this ideal gearing assumption, it can be shown that the acceleration of the load is given by

$$\alpha = \frac{T_{m\ \text{peak}}}{(\omega_r/\omega_L)[J_m + J_L/(\omega_r/\omega_L)^2]} \tag{16.20}$$

where the numerator on the right-hand side is the peak torque the motor can produce and J_m is the motor inertia. If one wished to determine what gear ratio

were required to obtain maximum acceleration of the load, equation 16.20 would be differentiated and set equal to zero, to obtain

$$\frac{\omega_r}{\omega_L} = \sqrt{\frac{J_L}{J_m}} \tag{16.21}$$

This expression implies that the load inertia should be made equivalent to the motor inertia by appropriate gearing, in order to obtain the best acceleration. Substituting equation 16.21 in 16.20, we can show that the maximum acceleration is given by

$$\alpha_{\max} = \frac{1}{2}\frac{T_{m\ \text{peak}}\,\omega_L}{J_m \omega_r} \qquad \text{Maximum acceleration} \tag{16.22}$$

Equations 16.16 to 16.22 are very useful in sizing a motor and in determining whether any gearing will be necessary to achieve the desired performance.

Torque Definitions

Although definitions of various torques were introduced in Chapter 15, it is useful to briefly review those definitions in light of the preceding subsection. In sizing a motor, it is important to ensure that the motor is capable of overcoming static friction at start-up, to accelerate to the desired speed in an acceptable fashion, and to handle any overloads that may occur. The following definitions will help in the analysis:

1. **Locked-rotor**, or **static, torque:** The minimum torque the motor will develop at rest for all angular positions under rated conditions.
2. **Breakdown torque:** The maximum torque a motor will develop under rated conditions without an abrupt drop in speed.

3. **Full-load torque:** The torque necessary to produce rated power output at full-load speed.
4. **Acceleration torque:** At any specified speed, acceleration torque, that is, the torque available for acceleration, is $T_{acc} = T_m - T_L - T_F$, where T_m is the motor torque, T_L is the load torque, and T_F is the frictional load torque.

Clearly, for the motor to accelerate to full-speed operation, the motor torque at standstill must exceed the total static-load torque. When the motor torque–speed curve intersects the load torque–speed curve, then a balanced operating condition has been reached.

Acceleration Calculations

The equation that defines the acceleration characteristics of a motor-load pair is

$$T_m - T_L = J\frac{d\omega}{dt} \tag{16.23}$$

where T_L is the total load torque. From this equation we can calculate the time required to accelerate the load from a speed ω_1 to a speed ω_2 as follows:

$$t = J\int_{\omega_1}^{\omega_2} \frac{d\omega}{T_m - T_L} \quad \text{s} \tag{16.24}$$

or, in units of revolutions per minute:

$$t = \frac{2\pi J_T}{60T}(n_1 - n_2) \quad \text{s} \qquad \text{Acceleration} \tag{16.25}$$

where T is the net torque (motor torque minus load torque) and J_T is the total system inertia (motor inertia plus reflected load inertia).

Efficiency Calculations

The efficiency of a motor is the ratio of the mechanical power output to the electric power input, that is, the effectiveness of the electromechanical energy

conversion. We have already discussed the sources of loss in Chapter 15 and classified them Page 16-39as electrical losses, magnetic losses, and mechanical losses; refer to Section 15.1 for these definitions. The efficiency of a motor η is defined by

$$\eta = 1 - \frac{P_{\text{loss}}}{P_{\text{input}}} = 1 - \frac{P_{\text{loss}}}{VI} \qquad \text{Efficiency} \tag{16.26}$$

Thermal Calculations

The calculation of the temperature rise and thermal dissipation in a motor can be quite complex and depends very much on the motor construction. For the purpose of illustration, we briefly discuss only the thermal characteristics of a DC motor and perform some thermal calculations for this type of machine.

Thermal dissipation is one of the most important limiting factors in the operation of DC machines. We assume that most power losses take place in the armature (a reasonable assumption, since most of the electric power flows through the armature circuit), and we use the thermal-electrical system analogy where the thermal difference $\Delta\theta°$ is given by

$$\Delta\theta° = I_a^2 R_a \times R_T \tag{16.27}$$

and where I_a is the armature current, R_a the armature resistance, and R_T the thermal resistance of the rotor. The thermal time constant of the motor is then defined to be the time (in seconds) taken by the armature to reach 63 percent of the temperature rise corresponding to a given constant power dissipation. Now, the maximum continuous torque the motor can develop is related to the power dissipation because the motor torque is proportional to the armature current:

$$T_{\max} = K_T I_{a\,\max} = K_T \sqrt{\frac{P_{\text{diss}}}{R_{\max}}} = K_T \sqrt{\frac{\Delta\theta°}{R_T R_{\max}}} \tag{16.28}$$

where P_{diss} is the dissipated power and $R_{\max}$ the rotor resistance at the maximum temperature, R_T is the rotor thermal resistance at ambient temperature, and $\Delta\theta$ is the temperature rise. The temperature rise of copper can

be determined from the known resistance of the wound rotor by computing the maximum temperature as follows:

$$\theta^{\circ}_{\max} = \frac{R_{\max}}{R_T}(\theta^{\circ}_{\text{ambient}} + 235) - 235 \qquad \text{Temperature rise} \tag{16.29}$$

and by computing $\Delta\theta^{\circ} = \theta^{\circ}_{\max} - \theta^{\circ}_{\text{ambient}}$, it is possible to use equation 16.28 to determine the maximum acceptable torque.

Conversely, to ensure that a given motor can operate within its thermal limits, one can calculate an average rms current requirement I_{rms}, consisting of the acceleration, deceleration, and running current, and use it to compute the temperature as

$$\Delta\theta^{\circ} = I^2_{\text{rms}} R_a R_T \tag{16.30}$$

Motor Selection

The range of electric motor applications is so broad that it is difficult to establish precise rules for motor selection. The differences between applications such as vehicle traction, robot motion, micromotors, disk drives, manufacturing machines, and pump systems, for example, are so many that it is virtually impossible to specify what the best motor would be unless the application and its environment were clearly specified. The aim of this subsection is simply to outline a procedure that can help in narrowing down the choice of a suitable drive motor to a few most likely candidates.

The first step in selecting a motor is the analysis of the requirements imposed by the application; these can be divided into three groups: (1) motor requirements, (2) load requirements, and (3) control requirements. Table 16.6 summarizes the important considerations for each of these.

Table 16.6 Motor selection requirements

Motor requirements	Load requirements	Control requirements
Operating speed	Determine worst-case operating conditions	Available power (AC, DC)
Life span and maintenance	Dynamic acceleration, full-load and overload conditions	Motor operating voltage and current
Torque characteristics	Starting conditions	Open- or closed-loop
Mechanical aspects (size, weight, noise level, environment)	Transients	Forward/reverse operation
Applicable standards (e.g., radio-frequency interference)	Need for gearing, selection of optimum gear ratio	Motoring and/or braking
Overload characteristics	Frictional characteristics	Torque, position, or speed control
Thermal dissipation characteristics		Accuracy of speed or position control
		Controller complexity and cost

On the basis of the requirements listed in Table 16.6, one can undertake the task of selecting a motor for a specific application.

Motion Requirements

The first step in the drive selection process is to understand the application-driven specifications, such as the type of motion, duty cycle, required acceleration and gearing system, and type of control that may be required (position, velocity, torque).

Motor Sizing

The second step in the drive selection process concerns the sizing of the motor itself. This is done first by calculating the maximum speed; next, the reflected inertia of the load and drive components is calculated. From the inertia calculations, the peak torque required by the application can be calculated. The maximum speed and torque requirements thus obtained will narrow the field significantly. Next, one should determine the appropriate constants for each of the candidate motors; these include, in general, inertias, resistances (electrical and thermal), and torque and back emf constants. With these constants it becomes possible to determine that the motor can operate within its thermal specifications by calculating the **temperature rise** of Page 16-41the machine in operation. This, of course, can be a greater limitation during certain portions of the motion cycle, for example, during a hard acceleration.

6Defining the Power Requirements

Calculate the peak voltage and current to determine the supply requirements.

Choosing a Transmission

Although we have been assuming that the mechanical drive system was known beforehand, so that the reflected inertia and peak torque could be calculated, there are many issues that need to be investigated in establishing the drive system, for example, the effect that elastic couplings might have in creating mechanical resonances; noise and vibration characteristics; and backlash due to gearing system imperfections.

Summary

It should be apparent from this brief discussion that the process of selecting an electromechanical drive is quite complex, and it requires a good understanding of many aspects of engineering, including heat transfer, kinematics, dynamics, electronics, systems, and, of course, electromechanics. We hope that this brief introduction will provide the motivation to pursue further studies in this subject area.

Conclusion

This chapter introduces a number of special-purpose electric machines that find widespread application in industry. The operating principles and analysis methods used in this chapter build directly on the foundations of Chapters 14 and 15.

Upon completing this chapter on electric machines, you should have mastered the following learning objectives:

1. *Understand the basic principles of operation of brushless DC motors and the tradeoffs between these and brush-type DC motors.* The brushless DC motor is a PM synchronous motor in which the mechanical commutation of conventional DC motors is replaced by electronic commutation. Brushless DC motors can be made quite compact, and they find application in vehicle propulsion and motion control.
2. *Understand the operation and basic configurations of step motors, and understand step sequences for the different classes of step motors.* Stepping motors—of the variable-reluctance, PM, or hybrid type—permit fine angular displacement control by moving in fixed, discrete steps. Typical applications are seen in robotics and control systems.
3. *Understand the operating principles of switched reluctance machines.* Switched reluctance machines are gaining more widespread acceptance because of their simplicity, since they require no permanent magnets and have very simple stator windings and rotor construction. Possible

applications of switched reluctance machines include low-cost industrial applications and vehicular propulsion.

4. *Classify and analyze single-phase AC motors, including the universal motor and various types of single-phase induction motors, using simple circuit models.* The universal motor is very similar in construction to a DC motor, but can operate on AC supplies; its speed can be controlled by electronic circuits of modest complexity. Thus, the universal motor finds common application in both variable- and fixed-speed appliances, such as power drills and vacuum cleaners, respectively. Induction motors can operate on a single-phase AC supply as a means is provided to establish a starting torque. Various techniques are commonly employed, such as split-phase, capacitor-start, and shaded-pole construction. Page 16-42The different types are characterized by differing torque–speed characteristics that make the single-phase induction motor a very versatile device. This is probably the most commonly employed electric machine.
5. *Outline the selection process for an electric machine, given an application; perform calculations related to load inertia, acceleration, efficiency, and thermal characteristics.* The selection of the appropriate motor for a given application should take into consideration many factors, including cost and packaging, the nature of the load, the performance specifications, and thermal considerations.

HOMEWORK PROBLEMS

Section 16.1: Brushless DC Motors

16.1 It is found that λ_m = 0.1 V-s for a permanent-magnet, six-pole, two-phase synchronous machine. Calculate the amplitude (peak value) of the open-circuit phase voltage measured when the rotor is turned at 60 r/sec.

16.2 A four-pole, two-phase brushless dc motor is driven by a mechanical source at n = 3,600 r/min. The open-circuit voltage across one of the phases is 50 V rms.

a. Calculate λ.

b. The mechanical source is removed, and the following voltages are applied: $v_a = \sqrt{2}\,25\cos\theta$ and $v_b = \sqrt{2}\,25\sin\theta$, where $\theta = \omega_e t$. Calculate the no-load rotor speed ω in radians per second.

16.3 With reference to Example 16.2, we wish to shorten the trapezoidal speed profile cycle time by accelerating the motor to a maximum speed of 1,800 r/min. If we still allow 1 s for acceleration and deceleration, how long will the cycle time be?

16.4 With reference to the triangular speed profile of Figure P16.4, determine the speed profile required to move a load 0.5 m (or 100 r) in 3 s.

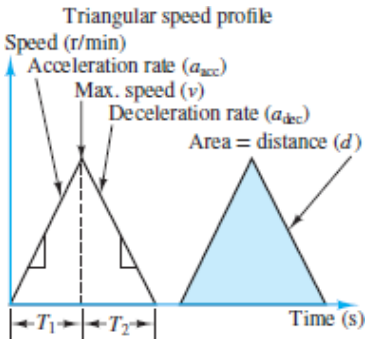


Figure P16.4

Section 16.2: Stepping Motors

16.5 With reference to Example 16.4, design a logic circuit that uses the logic design principles of Chapters 11 and 12 to achieve the step sequence given in Table 16.4.

(*Hint:* Use a counter and logic gates.)

16.6 A permanent-magnet stepper motor has six poles and a bipolar supply (i.e., the current into each coil pair can be either positive or negative). Figure 16.10 depicts a four-pole stepper motor as an example; the motor described in this problem has two additional poles. The spacing between the poles is uniform. Determine the size of the smallest achievable step in degrees.

16.7 Derive the dynamic equation for a stepping motor coupled to a load. The motor moment of inertia is J_m, the load moment of inertia is J_L, the viscous damping coefficient is D, and motor friction torque is T_f.

16.8 Sketch the rotor-stator configuration of a hybrid stepper motor capable of 18° steps.

(*Hint:* The rotor will have five teeth.)

16.9 Use a binary counter and logic gates to implement the stepping motor binary sequence of the second Check Your Understanding following Example 16.6.

16.10 A two-phase permanent-magnet stepper motor has 50 rotor teeth. When the rotor is driven by an external mechanical source at $\omega = 100$ rad/s, the measured open-circuit phase voltage is 25 V peak to peak. Calculate λ. If $i_a = 1$ A and $i_b = 0$, express the developed torque. Assume the winding resistance is 0.1 Ω.

16.11 The schematic diagram of a four-phase, two-pole permanent-magnet stepper motor is shown in Figure P16.11. The phase coils are excited in sequence by means of a logic circuit. Find

a. The logic schedule for full stepping of this motor.

b. The displacement angle of the full step.

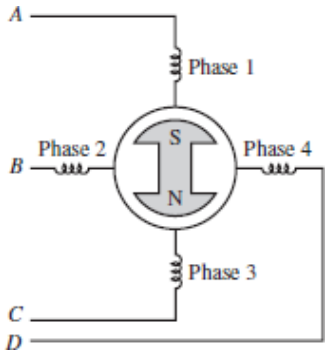
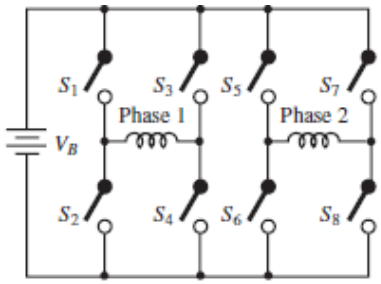


Figure P16.11

16.12 A permanent-magnet stepper motor is designed to provide a full-step angle of 15°. Find the number of stator and rotor poles.

16.13 A bridge driver scheme for a two-phase stepping motor is shown in Figure P16.13. Find the excitation sequences of the bridge operation (fill in the blanks of the table).



Clock state	Reset	1	2	3	4	5	6	7	8
S_1									
S_2									
S_3									
S_4									
S_5									
S_6									
S_7									
S_8									

Figure P16.13

16.14 A permanent-magnet stepper motor with a 15° step angle is used to directly drive a 0.100-in lead screw. Determine

a. The resolution of the stepper motor in steps per revolution.

b. The distance the lead screw travels (in inches) for each 15° step of the stepper motor.

c. The number of full 15° steps required to move the lead screw and the stepper motor shaft through 17.5 revolutions.

d. The shaft speed (r/min) when the stepping frequency is 220 pulses per second.

Section 16.4: Single-Phase AC Motors

16.15 Determine whether the following motors are integral- or fractional-horsepower motors:

a. $\frac{3}{4}$ hp, 900 r/min

b. $1\frac{1}{2}$ hp, 3,600 r/min

c. $\frac{3}{4}$ hp, 1,800 r/min

d. $1\frac{1}{2}$ hp, 6,000 r/min

16.16 The spatial fluctuation of the stator mmf $\mathcal{F}_1$ is expressed as

$$\mathcal{F}_1 = F_{1(\text{peak})} \cos\theta$$

where θ is the electrical angle measured from the stator coil axis and $F_{1(\text{peak})}$ is the instantaneous value of the mmf wave at the coil axis and is proportional to the instantaneous stator current. If the stator current is a cosine function of time, the instantaneous value of the spatial peak of the pulsating mmf wave is

$$F_{1(\text{peak})} = F_{1(\text{max})} \cos\omega t$$

where $F_{1(\text{max})}$ is the peak value corresponding to maximum instantaneous current. Derive the expression for $\mathcal{F}_1$, and verify that for a single-phase winding, both forward and backward components are present.

16.17 A 200-V, 60-Hz, 10-hp, single-phase induction motor operates at an efficiency of 0.86 and a power factor of 0.9. What capacitor should be placed in parallel with the motor so that the feeder supplying the motor will operate at unity power factor?

16.18 A 230-V, 50-Hz, two-pole, single-phase induction motor is designed to run at 3 percent slip. Find the slip in the opposite direction of rotation. What is the speed of the motor in the normal direction of rotation?

16.19 Determine the amount of time (in seconds) it will take for a stepper motor with a 15° step angle, operating in one-phase excitation mode, to rotate through 28 rad when the pulse rate is 180 pps. *Note:* $t = \theta/\omega$.

16.20 A $\frac{1}{4}$-hp, 110-V, 60-Hz, 110-V, 60-Hz, four-pole capacitor-start motor has the following parameters:

$$R_S = 2.02\,\Omega \qquad X_S = 2.8\,\Omega$$
$$R_R = 4.12\,\Omega \qquad X_R = 2.12\,\Omega$$
$$X_m = 66.8\,\Omega \qquad s = 0.05$$

Find

a. The stator current.

b. The mechanical power.

c. The rotor speed.

Page 16-44

16.21 A $\frac{1}{4}$-hp, four-pole, 110-V, 60-Hz, single-phase induction motor has the following data:

$$R_S = 1.86\,\Omega \qquad X_S = 2.56\,\Omega$$
$$R_R = 3.56\,\Omega \qquad X_R = 2.56\,\Omega$$
$$X_m = 53.5\,\Omega \qquad s = 0.05$$

Find the mechanical power output.

16.22 A single-phase, 115-V, 60-Hz, four-pole induction motor has the following parameters:

$$R_S = 0.5\,\Omega \qquad X_S = 0.4\,\Omega$$
$$R_R = 0.25\,\Omega \qquad X_R = 0.4\,\Omega$$
$$X_m = 35\,\Omega$$

Find the input current and developed torque when the motor speed is 1,730 r/min.

16.23 The no-load test of a single-phase induction motor is made by running the motor without load at rated voltage and rated frequency. Derive the equivalent circuit of a single-phase induction motor for the no-load test.

(*Hint:* The no-load slip is very small.)

16.24 Derive the equivalent circuit of a single-phase induction motor for the locked-rotor test. Neglect the magnetizing current.

16.25 The design for a $\frac{1}{8}$-hp, two-pole, 115-V universal motor gives the effective resistances of the armature and series field as 4 and 6 Ω, respectively. The output torque is 0.17 N-m when the motor is drawing rated current of 1.5 A (rms) at a power factor of 0.88 at rated speed. Find

a. The full-load efficiency.

b. The rated speed.

c. The full-load copper losses.

d. The combined windage, friction, and iron losses.

e. The motor speed when the rms current is 0.5 A, neglecting phase differences and saturation.

16.26 A 240-V, 60-Hz, two-pole universal motor operates at a speed of 12,000 r/min on full load and draws a current of 6.5 A at 0.94 power factor lagging. The series field-winding impedance is $4.55 + j3.2\ \Omega$, and the armature circuit impedance is $6.15 + j9.4\ \Omega$. Find

a. The back emf of the motor.

b. The mechanical power developed by the motor.

c. The power output if the rotational loss is 65 W.

d. The efficiency of the motor.

16.27 A single-phase motor is drawing 20 A from a 400-V, 50-Hz supply. The power factor is 0.8 lagging. What value of capacitor connected across the circuit will be necessary to raise the power factor to unity?

16.28 A three-phase induction motor is required to operate from a single-phase source. One possible connection is shown in Figure P16.28. Will the motor work? Explain why or why not.

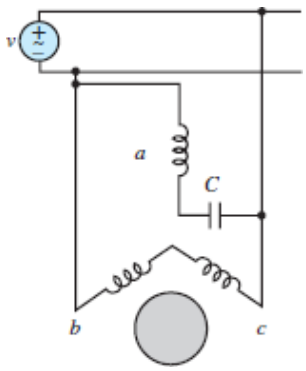


Figure P16.28

16.29 In performing a brake-load test upon a $\frac{1}{4}$-hp capacitor-start motor with its output adjusted to rated value, the following data were obtained: $E = 115$ V, $I = 3.8$ A, $P = 310$ W, rotation speed = 1,725 r/min. Calculate

a. Efficiency.

b. Power factor.

c. Torque in pound-inches.

Section 16.5: Motor Selection and Application

16.30 What type of motor would you select to perform the following tasks? Justify your selection.

a. Vacuum cleaner

b. Refrigerator

c. Air conditioner compressor

d. Air conditioner fan

e. Variable-speed sewing machine

f. Clock

g. Electric drill

h. Tape drive

i. *X*-*Y* plotter

16.31 A 5-hp, 1,150-r/min shunt motor has its speed controlled by means of a tapped field resistor, as shown in Figure P16.31. With the tap at position 3, determine the speed of the motor and the torque available at the maximum permissible load.

Page 16-45

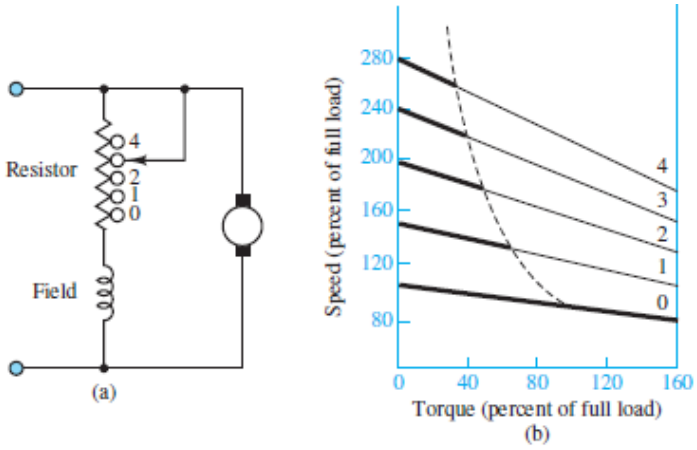


Figure P16.31

Figure P16.31

16.32 Which single-phase motor would you choose for the following applications?

a. Inexpensive analog electric clock.

b. Bathroom ventilator fan.

c. Escalator which must start under all load conditions.

d. Kitchen blender.

e. Table-model circular saw operating at about 3,500 r/min.

f. Handheld circular saw operating at 15,000 r/min.

g. Water pump.

16.33 The power required to drive a fan varies as the cube of the speed. If a motor driving a shaft-mounted fan is loaded to 100 percent of its horsepower rating on the top speed connection, find the horsepower output in percent of rating:

a. At a speed reduction of 20 percent.

b. At a speed reduction of 30 percent.

c. At a speed reduction of 50 percent.

16.34 An industrial plant has a load of 800 kW at a power factor of 0.8 lagging. It is desired to purchase a synchronous motor of sufficient capacity to deliver a load of 200 kW and also serve to correct the overall plant power factor to 0.92. Assuming that the synchronous motor has an efficiency of 91 percent, determine its kilovoltampere input rating and the power factor at which it will operate.

16.35 An electric machine is controlled so that its torque–speed characteristics exhibit a constant-torque region and a constant-power region as shown in Figure P16.35. The average efficiency of the electric drive (combination of machine, plus power electronics, plus control electronics) is 87 percent. The machine torque is constant from 0 to 2,500 r/min, and is equal to 150 N-m. The constant-power region is valid from 2,500 to 6,000 r/min. The machine drives a constant-torque load requiring 75 N-m.

a. Determine the operating speed of the machine.

b. Determine the electric power needed to operate the machine.

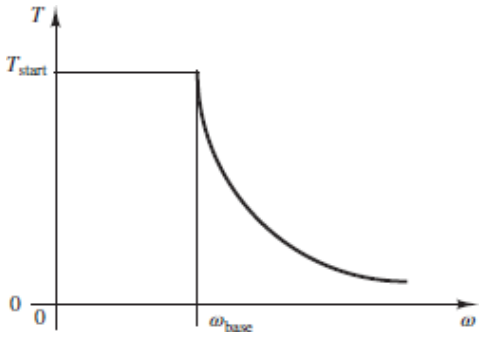


Figure P16.35

16.36 A PM synchronous (brushless DC) motor (wound stator, PM rotor, as shown in Figure P16.36) needs to be rated in terms of its thermal dissipation characteristics. To help in developing a rating, you are asked to write the dynamic equations linking the electrical dynamics to the thermal dynamics. Write the differential equations describing the electrothermomechanical dynamics of the system. You may make the following assumptions:

- All heat is generated in the stator by the stator current (i.e., the heat generated in the rotor is negligible). The rotor and stator are at the same temperature, and you may assume specific heat c.
- The stator is highly thermally conductive, and the dominant heat-transfer term is convection. You may assume an overall thermal resistance R_t from stator to air. The motor thermal mass is m.
- The motor generates torque according to the equation $T_m = kI_S$, and the back emf is equal to $E_b = k\omega$.

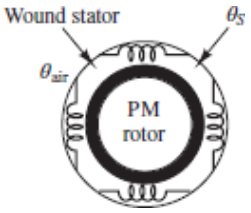


Figure P16.36 Thermoelectromechanical system. Electrical subsystem parameters: R_S, L_S, k (motor constant), $V_S(t)$, $I_S(t)$. Mechanical subsystem parameters: inertia and damping coefficient, J, b. Thermal subsystem parameters: thermal resistance, specific heat, mass, R_t, c, m.

16.37 A wound separately excited DC motor (wound stator and rotor, as shown in Figure P16.37) needs to be rated in terms of its thermal dissipation characteristics. To help in developing a rating, you are asked to write the dynamic equations linking the electrical dynamics to the thermal dynamics. Write the differential equations describing the electrothermomechanical dynamics of the system. (*Hint:* Start with the thermal equations.)

Use the following assumptions:

a. Heat is generated in the stator and in the rotor by the respective currents.

b. The stator and rotor are highly thermally conductive, and the dominant heat-transfer term is convection through the air gap and to ambient.

c. Heat storage in the air gap is negligible, and the air gap is infinitely thin.

d. The motor generates torque according to the equation $T_m = k_T I_a$, and the back emf is equal to $E_b = k_a \omega$.

e. The stator and rotor each act as a lumped thermal mass.

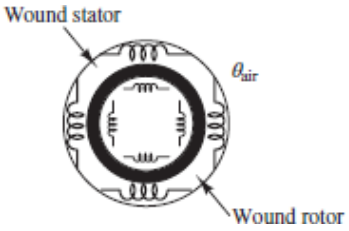


Figure P16.37 Thermoelectromechanical system. Electrical subsystem parameters: R_f, L_f, R_a, L_a (motor field and armature electrical parameters), k_f, k_a, k_T (motor field and armature constants), $V_S(t)$, $V_f(t)$, $I_a(t)$, $I_f(t)$. Mechanical subsystem parameters: load inertia, damping coefficient, load torque, J, b, T_L. Thermal subsystem parameters: $C_{t\text{–rotor}}$, $h_{t\text{–rotor}}$, A_{rotor} [rotor thermal capacitance, film coefficient of heat transfer from rotor surface to air and from air to

stator inner surface, rotor and inner stator surface area (assumed equal)]. $C_{t\text{–stator}}$, $h_{t\text{–stator}}$, A_{stator} (stator thermal capacitance, film coefficient of heat transfer from stator outer surface to air, stator outer surface area).

16.38 We wish to develop a thermal power rating for the (brushless DC) PM synchronous motor of Problem 16.36. To help in developing a rating, you are asked to write a simplified set of dynamic equations linking the electrical dynamics to the thermal dynamics. You may use the following assumptions:

a. All heat is generated in the stator by the stator current. The rotor and stator are at the same temperature, and you may assume specific heat c.

b. The stator is highly thermally conductive, and the dominant heat transfer term is convection. You may assume an overall thermal resistance R_m from stator to air. The motor thermal mass is m.

c. The motor operates at its rated (constant) speed ω_m (the brief acceleration transient to get the motor up to speed takes a short time, so *you do not need to consider the mechanical dynamics*).

Write the differential equation needed to calculate the time it takes the motor to reach its steady-state temperature. Determine, symbolically, the time constant for the temperature rise.

16.39 An electric machine is controlled so that its torque–speed characteristics exhibit a constant-torque region and a constant-power region, as shown in the sketch of Figure P16.35. The average efficiency of the electric drive (combination of machine, plus power electronics, plus control electronics) is 87 percent. The machine torque is constant from 0 to 2,500 rpm, and is equal to 150 N-m. The constant power region is valid from 2,500 to 6,000 rpm. The machine drives a constant torque load requiring 75 N-m.

a. Determine the operating speed of the machine.

b. Determine the electric power needed to operate the machine.

Sketch the solution on the curve of Figure P16.35 with exact numerical values.

16.40 A PM DC motor is rigidly coupled to a fan; the fan load torque is described by the expression $T_L = 5 + 0.05\omega + 0.001\omega^2$ where torque is in N-m and speed in rad/s. The motor has $k_a\phi = k_T\phi = 2.42$. $R_a = 0.2\ \Omega$, and the inductance is negligible. If the motor voltage is 50 V, what will be the speed of rotation of the motor and fan?

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

C H A P T E R

17

POWER ELECTRONICS

Power electronic devices are specially designed diodes and transistors that have the ability to carry large currents and sustain large voltages. These devices are the "muscle" of many electromechanical systems. For example, one finds such devices in many appliances, in industrial machinery, and virtually wherever an electric motor is found since one of the foremost applications of power electronic devices is to supply and control the currents and voltages required to power electric machines.

This chapter surveys the basic properties and types of power electronic devices and illustrates the application of a few of these devices, especially with regard to electric motor drivers. As alluded to above, diodes and transistors are at the heart of power electronic devices and systems. Diodes act as rectifiers and regulators and as simple unidirectional "valves" for current. Transistors act as switches and amplifiers. In particular, power amplifiers often rely on BJTs due to their superior linearity and current handling capability when compared to MOSFETs. MOSFETs are capable of much higher switching speeds than BJTs. When large currents must be switched at high speeds, a hybrid device known as an *insulated gate bipolar*

transistor (**IGBT**) is commonly used because it possesses some of the advantages of both BJTs and MOSFETs. In fact, *switch-mode power supplies*, many of which rely on IGBTs, are popular choices for many power electronic applications.

Page 17-2

Learning Objectives

Students will learn to...

1. Classify power electronic devices and circuits. *Sections 17.1 and 17.2*.
2. Analyze the operation of practical voltage regulators. *Section 17.3*.
3. Utilize transistor power amplifiers. *Section 17.4*.
4. Analyze the operation of single- and three-phase controlled rectifier circuits. *Section 17.5*.

17.1 POWER ELECTRONIC DEVICES AND CIRCUITS

Power semiconductors can be classified into five groups:

1. Power diodes.
2. Thyristors.
3. Power bipolar junction transistors.
4. Insulated-gate bipolar transistors.
5. Static induction transistors.

Figure 17.1 depicts the symbols for common power electronic devices.

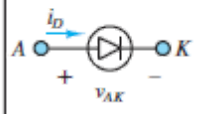
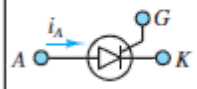
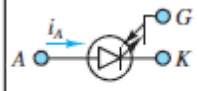
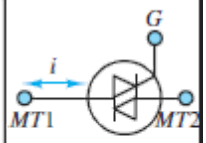
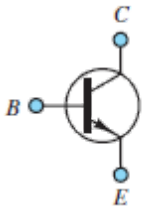
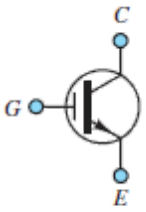
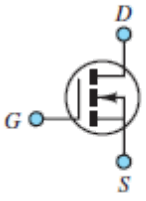
Device	Device symbol
Diode	
Thyristor	
Gate turnoff thyristor (GTO)	
Triac	
npn BJT	
IGBT	
n-channel MOSFET	

Figure 17.1 Classification of power electronic devices

Power diodes are functionally identical to small-signal diodes, except for their ability to manage much larger voltages and currents. For example, a *general-purpose* diode may be rated at 3 kV and 3.5 kA, while a *high-speed* diode may be rated as high as 3 kV and 1 kA. *Schottky* diodes are even faster, with switching intervals in the nanosecond range, but they are typically limited to 0.1 kV and 0.3 kA. The forward voltage drops of small-signal and power diodes are similar, both being on the order of 1 V, which is

usually negligible compared to the voltages found in power applications. As a result, power diodes often approximate ideal switches.

Thyristors are one of the oldest power electronic devices and are similar to power diodes except for an additional gate terminal that determines when conduction begins. A small current injected into the gate will trigger conduction if the diode portion of the thyristor is forward-biased, as indicated in Figure 17.1. Once triggered, the gate current is not needed to maintain conduction, which continues until the diode portion of the thyristor is reverse-biased. Thyristors can be rated at up to 6 kV and 3.5 kA. The **turnoff time** of a thyristor is an important characteristic and can be as short as 10 μs; however, a short turnoff time is usually accompanied by a lower power rating. Over the years many varieties of thyristors have been developed, including the gate turnoff (GTO), reverse-conducting (RCT), static induction (SIT), gate-assisted turnoff (GATT), MOS controlled (MCT), and force- and line-commutated thyristors, as well as the original silicon-controlled rectifier (SCR) and light-activated SCR (LASCR). The operation of each of these devices is a slight modification of the basic Page 17-3thyristor. Of this group, only the GTO can be turned on and off without the aid of a separate commutation circuit. A short negative pulse to the gate turns off a GTO. A **triac** is essentially a pair of oppositely directed thyristors connected in parallel, as shown in Figure 17.1. A single gate enables conduction in either direction. Conduction is maintained until the thyristor is no longer forward-biased in either direction.

Power BJTs can reach ratings up to 1,200 V and 400 A, and they operate in much the same way as a conventional BJT. Power BJTs are used in power converter applications at frequencies up to around 10 kHz. **Power MOSFETs** can operate at somewhat higher frequencies (a few to several tens of kilohertz) but are limited in power (typically up to 1,000 V and 50 A). **Insulated-gate bipolar transistors (IGBTs)** are voltage-controlled (because of their insulated gate, reminiscent of insulated-gate FETs) power transistors that offer superior speed with respect to BJTs but are not quite as fast as power MOSFETs.

Power Electronic Circuits

One possible classification of power electronic circuits is given in Table 17.1. The first four of these circuits may be familiar. For example, **voltage**

regulators and **rectifiers** were introduced in the chapter on diodes. In this chapter, practical considerations and variations of these circuits are discussed. Likewise, principles of transistor **switches** and **amplifiers** were discussed in earlier chapters. Here, high-power versions of these circuits and their characteristics are discussed. For example, it is important to consider carefully power limitations and signal distortion in power amplifiers. Other properties are important when transistors are used as switches.

Table 17.1 Power electronic circuits

Circuit	Function
Voltage regulator	Regulate the output of a DC voltage supply
Diode rectifier	Convert fixed AC voltage to fixed DC voltage (single- or multi-phase)
Power amplifier	Amplify large-signal voltages and currents
Switch	Connect/disconnect one part of a circuit to/from another
Controlled rectifier	Convert fixed AC voltage to variable DC voltage (single- or multi-phase)
AC voltage controller	Convert fixed AC voltage to variable AC voltage (single- or multi-phase)
Chopper	Convert fixed DC voltage to variable DC voltage
Inverter	Convert fixed DC voltage to variable AC voltage (single- or multi-phase)

The last four circuits listed in Table 17.1 were not introduced in any previous chapter. They are all generally known as **converters**, and their principal shared characteristic is the use of a switch to control or vary their output. For example, the DC output of a thyristor-based **controlled rectifier** is variable in the sense that the gate current can be varied to control the initiation of conduction, resulting in Page 17-4a DC output like that illustrated in Figure 17.2. The firing angles α and $\alpha + \pi$ indicate the initiation of conduction through thyristors T_1 and T_2 during each period of the AC supply. This type of AC-DC converter is commonly used in DC motor applications.

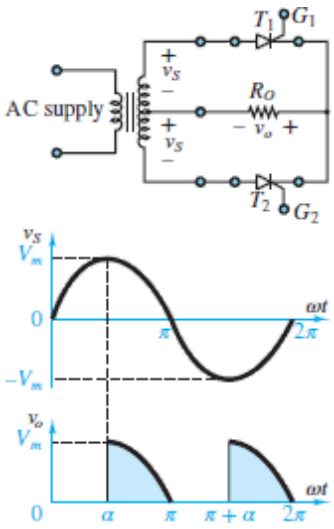


Figure 17.2 AC-DC converter circuit and waveform

As another example, Figure 17.3 shows one type of **AC voltage controller**, where the bidirectional capability of a triac is employed to modify an AC waveform. The period of the output waveform matches that of the input waveform; however, the rms value of the output v_o is smaller than that of the input v_s. This type of circuit is known more generally as an AC-AC converter.

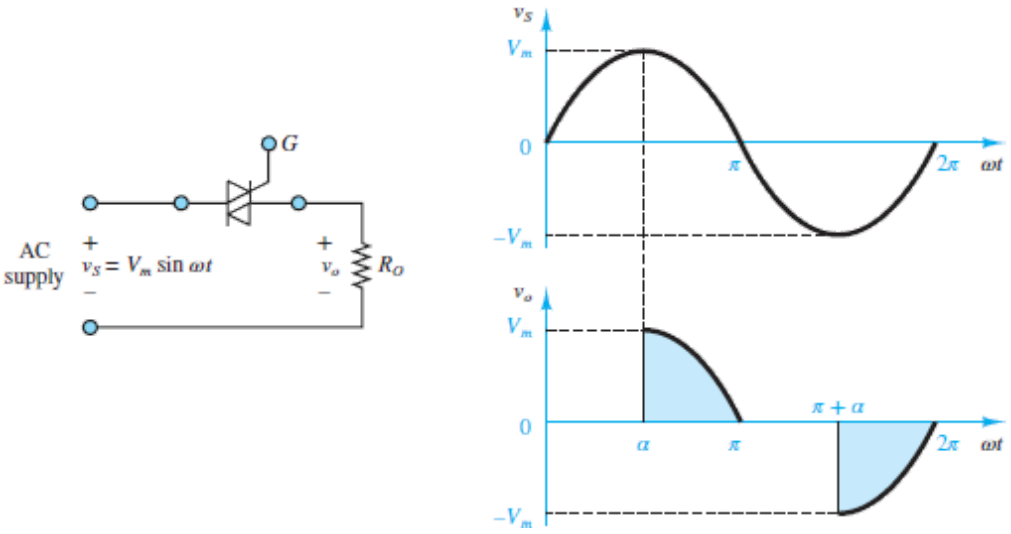


Figure 17.3 AC-AC converter circuit and waveform

Figure 17.4 shows a basic **chopper**, or **switching regulator**, converts a fixed DC source to a variable DC source. In this circuit, a BJT is employed as a switch to enable/disable conduction according to the duty cycle of the base-emitter voltage. The reduced DC average of the output waveform is determined by that duty cycle. Such DC-DC converters are often used as variable-voltage supplies for DC motors.

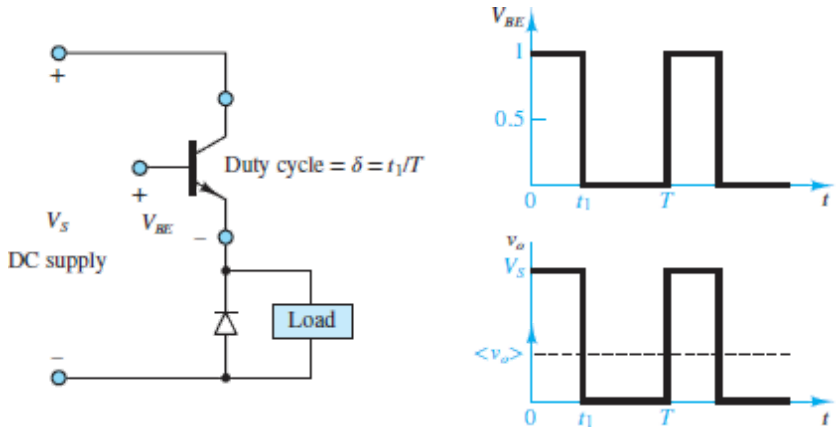


Figure 17.4 DC-DC converter circuit and waveform

Page 17-5

Figure 17.5 shows a basic **inverter**, which converts a fixed DC supply to a variable AC supply by coordinating the switching of pairs of transistors. Such DC-AC converters find application in AC motor control.

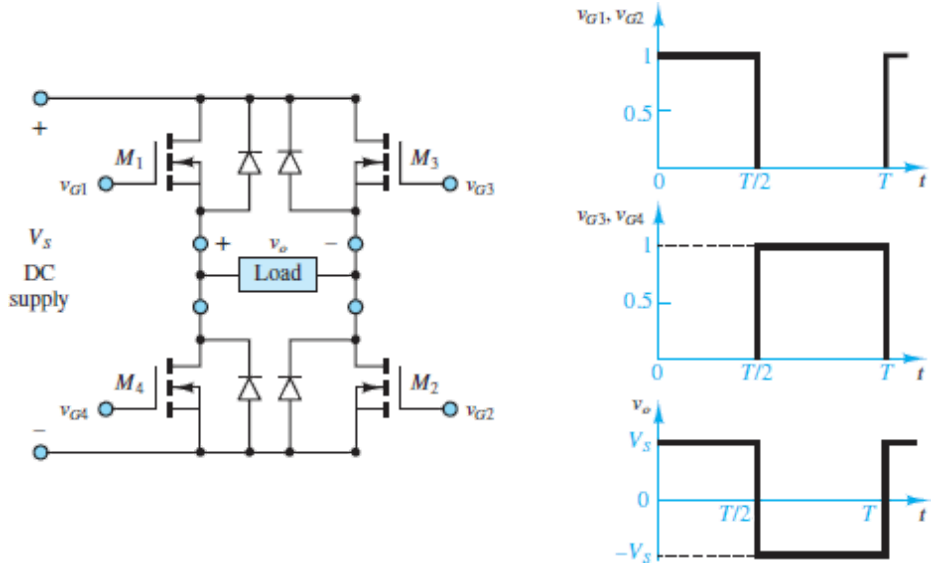


Figure 17.5 DC-AC converter circuit and waveform

17.2 VOLTAGE REGULATORS

Figure 17.6 depicts a Zener diode i-v characteristic and shows a block diagram of a three-terminal voltage regulator. A simple voltage regulator comprised of a resistor in series with a Zener diode was introduced in the chapter on diodes. In practice, a more robust voltage regulator includes a series pass transistor, as shown later in Figure 17.8, where $v_S = V_S \pm \Delta v_S$, $v_o = V_o \pm \Delta v_o$, and $v_Z = V_Z + i_Z r_Z$. When the nominal supply voltage V_S is sufficiently greater than the Zener voltage V_Z, the BJT will be in the active region such that $v_{BE} \approx V_\gamma$, $i_C \gg i_Z$, and $i_C \approx i_o$. The result is that $v_Z \approx V_Z$ and, by KVL, the load voltage is

$$v_o = V_Z - V_\gamma = \text{constant} \tag{17.1}$$

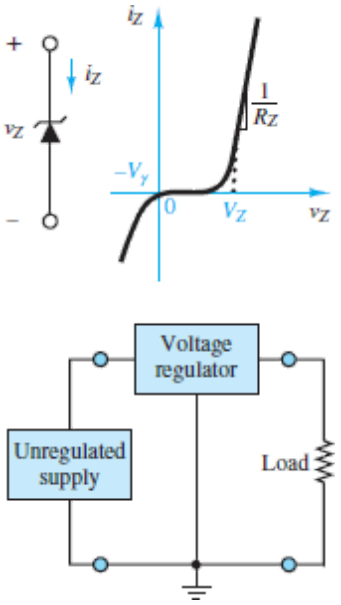


Figure 17.6 Zener diode characteristic and voltage regulator circuit

Note that this result is relatively independent of fluctuations in the unregulated source voltage and of the required load current. Also note that,

by KVL, $v_{CE} = v_S - v_o$. Thus, the required power rating of the BJT may be determined by considering the largest unregulated voltage $v_{S\,\max}$:

$$\begin{aligned} P_{\text{BJT}} &= (v_{S\,\max} - v_o) i_C \\ &\approx (v_{S\,\max} - v_o) i_o \end{aligned} \tag{17.2}$$

Three-terminal voltage regulators are available in a single package that often includes overcurrent protection and is rated in terms of the regulated voltage and maximum power dissipation. The regulated voltage is usually fixed but for some packages may be adjusted through external components. When large power dissipation is required, regulators need to be attached to a **heat sink**, usually a metal plate with fins to enhance heat transfer to the surroundings. Figure 17.7 depicts typical heat sinks.

Page 17-6

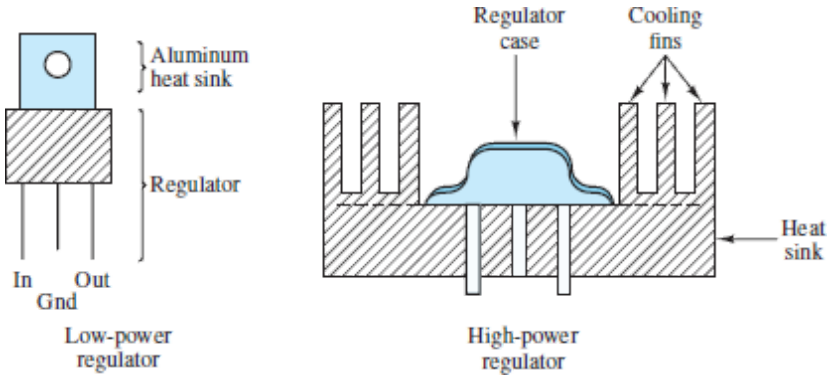


Figure 17.7 Heat sink construction for voltage regulators

EXAMPLE 17.1 Analysis of Voltage Regulator

Problem

Determine the nominal load current i_o and the required Zener diode power rating for the voltage regulator of Figure 17.8.

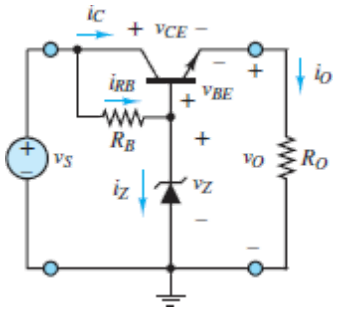


Figure 17.8 Practical voltage regulator

Solution

Known Quantities: Zener voltage; nominal source voltage; BJT base and load resistors; transistor parameters.

Find: $i_{o_{max}}$ and P_Z and P_Z.

Schematics, Diagrams, Circuits, and Given Data: $V_Z = 12.7$ V; $V_S = 20$ V; $R_B = 47\ \Omega$; $R_o = 10\ \Omega$. Transistor data for a TIP31: $V_{BE\,\text{on}} = V_\gamma = 1.8$ V; $\beta = 10$ (see Table 17.2).

Assumptions: None.

Analysis: Notice immediately that since the nominal value V_S of the voltage source is significantly greater than the nominal Zener voltage V_Z the base-collector junction of the BJT is reverse-biased. Also, the base-emitter junction must be forward-biased. (It is a worthwhile exercise to show that the BJT cannot be in cutoff mode.) Consequently, the BJT is in active mode, the Zener diode is on, and the load voltage is being regulated.

To solve for the load current consider Figure 17.9, which depicts the equivalent load circuit. From KVL:

$$V_Z = V_{BE} + R_o I_o$$

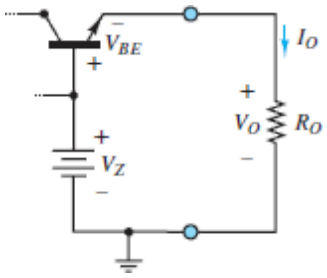


Figure 17.9

Thus, the nominal load current is:

$$I_o = \frac{V_Z - V_\gamma}{R_o} = \frac{12.7 - 1.8}{10} = 1.09 \text{ A}$$

To solve for the required Zener diode power rating it is necessary to determine the nominal power $P_Z = V_Z I_Z$ consumed. Since V_Z is specified, it is only necessary to find I_Z, Page 17-7as shown in Figure 17.10. The first step is to apply KVL to the base circuit, as shown in Figure 17.11, to compute I_{RB}:

$$I_{RB} = \frac{V_S - V_Z}{R_B} = \frac{20 - 12.7}{47} = 155.3 \text{ mA}$$

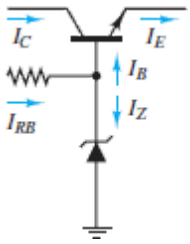


Figure 17.10

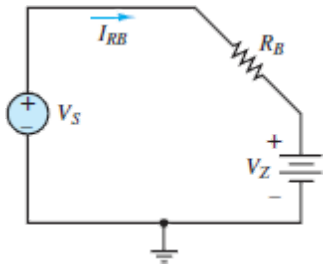


Figure 17.11

The next step is to solve for I_B by recalling that the transistor is in active mode and writing $I_E = (\beta + 1)\, I_B$, where I_E is the same as the load current I_o. Thus:

$$I_B = \frac{I_E}{\beta + 1} = \frac{I_o}{\beta + 1} = \frac{1.09}{10 + 1} = 99.1 \text{ mA}$$

Next, apply KCL at the wire junction shown in Figure 17.10 to find:

$$I_{RB} - I_B - I_Z = 0$$
$$I_Z = I_{RB} - I_B = 155.3 - 99.1 = 56.2 \text{ mA}$$

Finally, the power consumed by the Zener diode is

$$P_Z = I_Z \times V_Z = 0.0562 \times 12.7 = 0.714 \text{ W}$$

The required Zener diode power rating should exceed this value and also allow for small variations in I_Z. A $\frac{3}{4}$-W Zener diode should suffice, while a 1-W Zener diode would be more than adequate.

Comments: Note that the BJT enables regulation while keeping the Zener current relatively small. Also, in this example, the nominal collector-emitter voltage is easily found from KVL to be $V_{CE} = V_S - V_o = V_S - V_Z + V_\gamma = 9.1$ V. Since $V_{CE} \gg V_{CE\,\text{sat}} = 1.3$ (see Table 17.2), the transistor is not close to saturation and modest variations in the source and load voltage will not compromise the regulator.

CHECK YOUR UNDERSTANDING

Repeat Example 17.1 using the MJE3055T transistor (see Table 17.2).

Table 17.2 Typical parameters for representative power BJTs

	MJE3055T	TIP31	MJE170
Type	*npn*	*npn*	*pnp*
Maximum I_C (continuous)	10 A	3 A	–3 A
V_{CEO}	60 V	40 V	–40 A
Power rating	75 W	40 W	12.5 W
β	20 @ $I_C = 4$ A	10 @ $I_C = 3$ A	30 @ $I_C = -0.5$ A
$V_{CE\ sat}$	1.1 V	1.3 V	–1.7 V
$V_{BE\ on}$	8 V	1.8 V	–2 V

17.3 POWER AMPLIFIERS

There are many applications in which substantial power must be delivered to a load, such as a loudspeaker or electric motor. The control of electric power in industry requires electronic devices that can carry currents as high as hundreds of amperes and voltages up to thousands of volts. The aim of this section is to discuss some of the more relevant issues in the design of *power amplifiers*, such as distortion and heat dissipation.

BJTs are still preferred for **audio power amplifiers** due to the relatively linear amplification characteristics of BJTs operating in active mode. Power MOSFETs have advantages in many other applications, particularly those involving high-frequency waveforms. However, an understanding of the issues involved with power BJT circuits also provides insight on the issues present in power MOSFET circuits.

In practice, a BJT is limited in its operation by the maximum collector current, the maximum collector-emitter voltage, and the maximum power dissipation, which is the product of I_C and V_{CE}. Figure 17.12 shows the maximum power limit of a BJT as presented in a plot of i_C versus v_{CE}. As alluded to above, it is also important to keep in mind that linear amplification using a BJT requires operation in the active mode.

1. Exceeding the maximum allowable current $I_{C\ \max}$ on a continuous basis will result in melting the wires that bond the device to the package terminals.
2. Maximum power dissipation is the locus of points for which $V_{CE}I_C = P_{\max}$ at a case temperature of 25°C. The average power dissipation should not exceed $P_{\max}$.
3. The instantaneous value of v_{CE} should not exceed $V_{CE\ \max}$; otherwise, avalanche breakdown of the collector-base junction may occur.

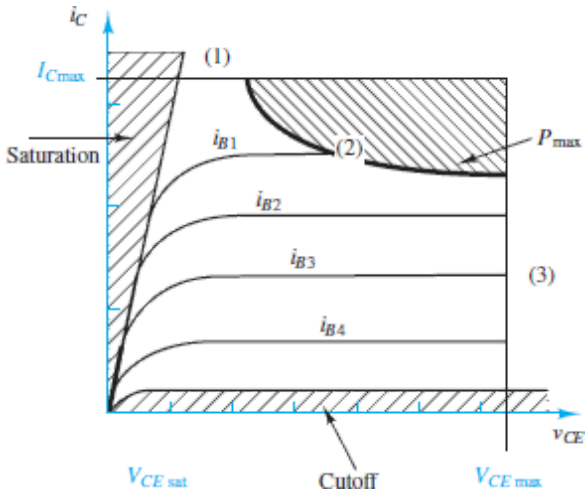


Figure 17.12 Limitations of a BJT amplifier

Consider the effect of driving an amplifier beyond the limits of the linear active region, into saturation or cutoff. The result will be signal distortion. For example, a sinusoid amplified by a transistor amplifier that is forced into Page 17-9saturation, either by a large input or by an excessive gain, will be compressed around the peaks, because of the decreasing device gain in the extreme regions. Thus, to satisfy these limitations—and to fully take advantage of the relatively distortion-free linear active region of operation

for a BJT—the Q point should be placed in the center of the device characteristic to obtain the **maximum symmetric swing**.

Of course, the maximum power dissipation of the device can present a rather severe limitation on the performance of the amplifier, in that the transistor can be irreparably damaged if its power rating is exceeded. Values of the maximum allowable collector current $I_{C\,\max}$, the maximum allowable transistor power dissipation $P_{\max}$, and other relevant power BJT parameters are given in Table 17.2 for a few typical devices. Because of their large geometry and high operating currents, typical power transistor parameters are quite different from those of small-signal transistors.

1. β is low. It can be as low as 5; the typical value is 20 to 80.
2. Typically $I_{C\,\max}$ is in the ampere range; it can be as high as 100 A.
3. V_{CEO} is usually 40 to 100 V, but it can reach 500 V.

EXAMPLE 17.2 Class A and B Amplifiers: Push-Pull Power Amplifier Output Stage (Loudspeaker Driver)

Problem

One of the limitations of transistor power amplifiers, as explained in this section, is their power dissipation, which limits the maximum useful load power a transistor can output. The aim of this example is to compare a typical emitter-follower amplifier with a **push-pull amplifier** when used as an output power stage to drive, say, a loudspeaker. Because of the low impedance of a loudspeaker (8 Ω is typical), the power or output stage of an audio amplifier needs to provide fairly high currents. For example, an 8-Ω, 50-W loudspeaker will require 2.5 A, demanding the use of a power transistor stage. Further, note that the input signal to be amplified is intrinsically an AC signal and that, therefore, the output stage current must swing above and below ground.

Solution

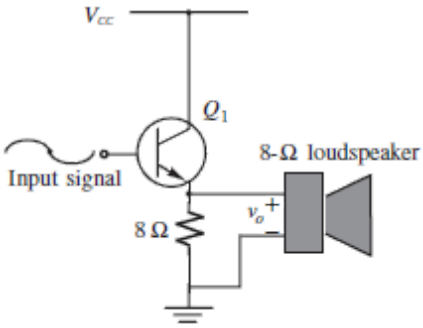
Known Quantities: Transistor parameters for TIP33C-34C complementary power BJT pair.

$$
\begin{aligned}
I_C \text{ (A)} &= 10 \\
V_{CBO} \text{ (V)} &= 100 \\
V_{CEO} \text{ (V)} &= 100 \\
P_D \text{ (W)} &= 80 \\
H_{FE} \text{ (min/max)} &= 20/100 \\
I_C/V_{CE} \text{ (A/V)} &= 3.0/4.0 \\
V_{CE(SAT)} \text{ (V)} &= 1.0 \\
I_C/I_B \text{ (A/mA)} &= 3.0/300
\end{aligned}
$$

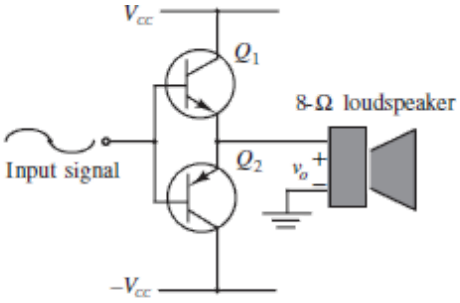
Find: Analyze the two transistor amplifiers in Figure 17.13(a) and (b).

Schematics, Diagrams, Circuits, and Given Data: Figure 17.13.

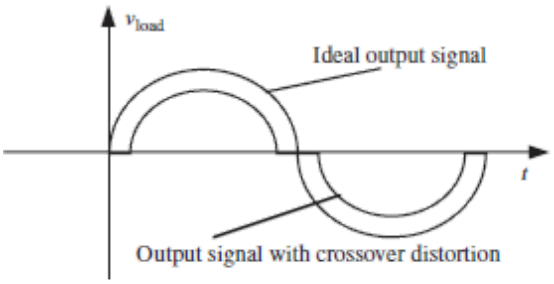
Page 17-10



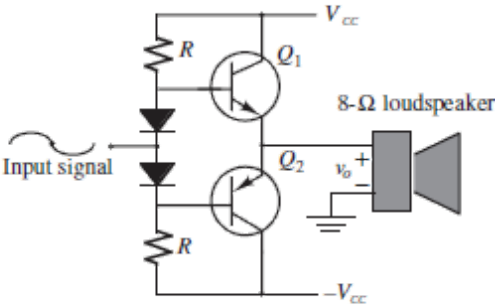
(a) Emitter-follower output stage used as loudspeaker driver



(b) Push-pull output stage used as loudspeaker driver



(c) Crossover distortion



(d) Push-pull output stage correcting for crossover distortion

Figure 17.13 Emitter-follower (class A) and push-pull (class B) amplifier

Assumptions: None.

Analysis: Consider first the emitter follower of Figure 17.13(a). To ensure that the amplifier output can swing above and below ground, the biasing of the amplifier will require that a substantial (quiescent) direct current flow through the amplifier at all times, so that the *Q* point of the amplifier is in the middle of the characteristic curve and the collector current can "swing" over the full range between the cutoff region on the low end and the maximum power dissipation (or saturation) limits shown in Figure 17.12.

This means that if we require a swing of 5 A (±2.5 A around the *Q* point), the amplifier will see a continuous quiescent current of 2.5 A, which will result in substantial continuous power dissipation. We can estimate this quiescent power dissipation to be the sum of the power dissipated by the 8-Ω resistor, approximately $R \times I_{CQ}^2 = 8 \times 2.5^2 = 50$ W plus the power dissipated by the transistor, $I_{CQ} \times V_{CEQ} = 2.5 \times V_{CEQ}$, which could be as much as 10 or 20 additional watts, depending on the collector-emitter bias voltage. So we are looking at a continuous power dissipation that is greater than the (50-W) power actually absorbed by the loudspeaker!

The circuit of Figure 17.13(b) uses two complementary (*npn-pnp*) transistors in a *push-pull* configuration to avoid this problem. Note that this circuit requires a symmetrical power supply [the single-ended collector supply of Figure 17.13(a) is no longer adequate]. When the input signal (assume a sinusoidal signal for simplicity) is in the positive half of the cycle, the top (*npn*) transistor enters the active region (once the input voltage exceeds the junction voltage) and amplifies the base current, while the bottom transistor is in the cutoff region. During the input signal negative half-cycle, the bottom (*pnp*) transistor acts as a linear amplifier, while the top transistor is in cutoff. The advantage of this configuration is that if the input signal is zero, both transistors are off. Thus, there is no quiescent power consumption. Note that no resistors are required to bias this amplifier configuration; Page 17-11thus, there will be no I^2R power consumption either. The primary disadvantage of the push-pull amplifier is illustrated in Figure 17.13(c). The figure clearly shows that the output of the amplifier is zero until one of the transistors has entered the active region. This leads to

the **crossover distortion** shown in the figure. The crossover distortion can be eliminated with the circuit of Figure 17.13(d), in which two diodes are included such that their forward-bias voltages cause the two base-emitter junctions to be always forward-biased. Now the two BJTs are immediately in the active region whenever the appropriate half of the cycle is on.

Comments: The push-pull amplifier is the most common form of output stage in audio amplifier and finds use in many other linear amplifier configurations.

EXAMPLE 17.3 Efficiency of Power Amplifiers

Problem

Example 17.2 illustrated the advantage of a push-pull amplifier output stage in reducing the quiescent power consumption of an amplifier. In this example, the efficiency of three different amplifier configurations are compared: the common-emitter, emitter-follower, and push-pull amplifiers.

Solution

Known Quantities:

Find: The power dissipation $P_o = v_o i_o$ of each of the transistor amplifiers in Figure 17.14(a) to (c). Calculate the efficiency η of each amplifier based on the following definition (ratio of load to input rms power):

$$\eta = \frac{\bar{P}_o}{\bar{P}_{CC}}$$

Schematics, Diagrams, Circuits, and Given Data: Figure 17.14.

Assumptions: In all amplifiers we assume a sinusoidal input, and we also assume that the amplifier is biased to permit maximum symmetrical load voltage swing, from zero to $+V_{CC}$ for the common-emitter and emitter-

follower amplifiers, and $-V_{CC}$ to V_{CC} for the push-pull configuration. This last assumption is not very realistic, but it simplifies the calculations a great deal.

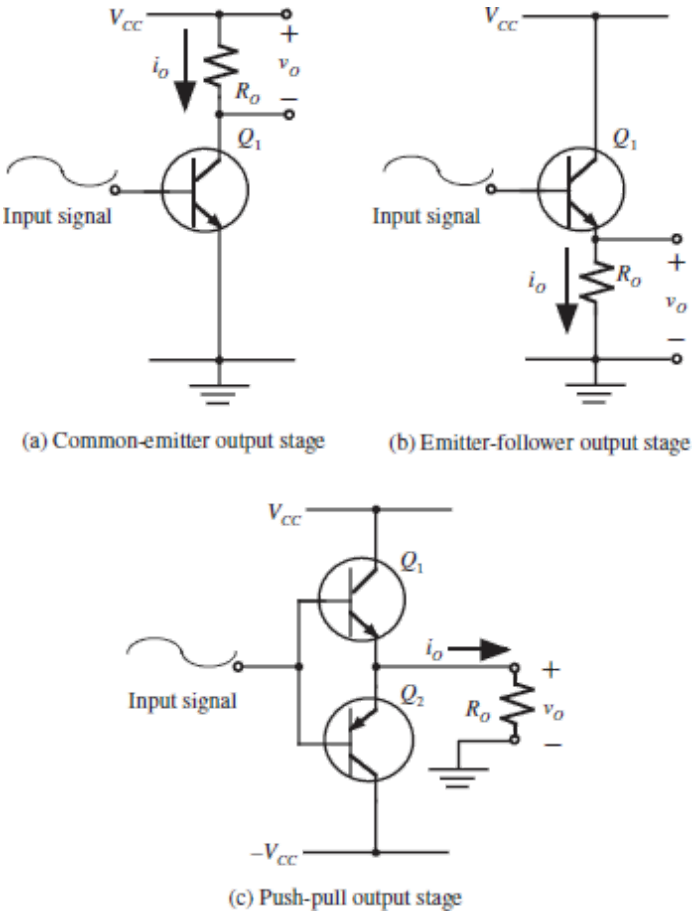


Figure 17.14 Three amplifier output stages

Analysis:

1. *Common-emitter amplifier.* Since we have assumed that the load voltage swing is from 0 to $+V_{CC}$, we can write the following expression for the load voltage:

$$v_o = \frac{V_{CC}}{2} + \frac{V_{CC}}{2}\sin \omega t$$

and the current will be

$$i_o = \frac{V_{CC}}{2R_o} + \frac{V_{CC}}{2R_o}\sin \omega t$$

Now, since the input power is $P_{CC} = V_{CC}I_C$ and $I_C = i_o$, we can write

$$P_{CC} = V_{CC}\left(\frac{V_{CC}}{2R_o} + \frac{V_{CC}}{2R_o}\sin\omega t\right) = \frac{V_{CC}^2}{2R_o}(1 + \sin\omega t)$$

Page 17-12

Conversely, the load power is

$$P_o = v_o i_o = \left(\frac{V_{CC}}{2} + \frac{V_{CC}}{2}\sin\omega t\right)\left(\frac{V_{CC}}{2R_o} + \frac{V_{CC}}{2R_o}\sin\omega t\right) = \frac{V_{CC}^2}{4R_o}(1 + \sin\omega t)^2$$

Thus, the efficiency of the amplifier can be calculated from the expression:

$$\frac{P_o}{P_{CC}} = \frac{(V_{CC}^2/4R_o)(1 + \sin\omega t)^2}{(V_{CC}^2/2R_o)(1 + \sin\omega t)} = \frac{1}{2}(1 + \sin\omega t)$$

The rms value of the function (1 + sin ωt) can be computed to be equal to $\sqrt{1.5} = 1.2247$ and therefore

$$\eta = \frac{\tilde{P}_o}{\tilde{P}_{CC}} = \frac{1.2247}{2} \approx 0.61, \text{ or } 61 \text{ percent.}$$

2. *Emitter-follower amplifier.* Since we have assumed that the amplifier is biased to produce a symmetrical swing, there is very little difference between common-emitter and emitter-follower configurations. In the emitter-follower, the load current is the emitter current (rather than the collector current). A similar derivation can be followed, as shown in part 1, to arrive at the same result.
3. *Push-pull amplifier.* Here, we assume that the load voltage swing is from $-V_{CC}$ to $+V_{CC}$, for maximum amplification. Then, we can write the following expression for the load voltage:

$$v_o = V_{CC}\sin\omega t$$

Page 17-13

and the current will be

$$i_o = \frac{V_{CC}}{R_o}\sin \omega t$$

Now, since the input power is $P_{CC} = V_{CC}I_C$ and $I_C = i_o$ (noting that the load current is the emitter current from the transistor that is on during each half-cycle—see Example 17.2), we can write

$$P_{CC} = V_{CC}\left(\frac{V_{CC}}{R_o}\sin \omega t\right) = \frac{V_{CC}^2}{R_o}\sin \omega t$$

Conversely, the load power is

$$P_o = v_o i_o = (V_{CC}\sin \omega t)\left(\frac{V_{CC}}{R_o}\sin \omega t\right) = \frac{V_{CC}^2}{R_o}(\sin \omega t)^2$$

Thus, the efficiency of the amplifier can be calculated from the expression:

$$\frac{P_o}{P_{CC}} = \frac{(V_{CC}^2/R_o)(\sin \omega t)^2}{(V_{CC}^2/R_o)\sin \omega t} = \sin \omega t$$

The rms value of the above function is 0.707, and therefore

$$\eta = \frac{P_o}{P_{CC}} \approx 0.71, \text{ or } 71 \text{ percent.}$$

Comments: The push-pull amplifier is very useful in power output stages.

17.4 TRANSISTOR SWITCHES

In addition to their application in power amplifiers, power BJTs can serve as controlled power switches, taking advantage of the switching characteristic described in Chapter 9. In addition to the properties already discussed, it is important to understand the phenomena that limit the switching speed of bipolar devices. The parasitic capacitances C_{CB} and C_{BE} that exist at the *CB* and *BE* junctions along with a transistor's internal resistance impose a limiting time constant on the rate at which a transistor can switch from cutoff to saturation. Figure 17.15 illustrates a transistor's response to a step

voltage applied to the base. Whereas the current into the base also experiences a step change, this current must charge the base capacitance such that the *BE* junction voltage requires a finite time to reach V_γ and results in a **delay** t_d. Once the *BE* junction is forward-biased, a collector current ensues and rises steadily over a finite **rise time** t_r. This behavior reflects the steadily rising *BE* junction voltage as the transistor transitions from cutoff, through the active mode, to saturation.

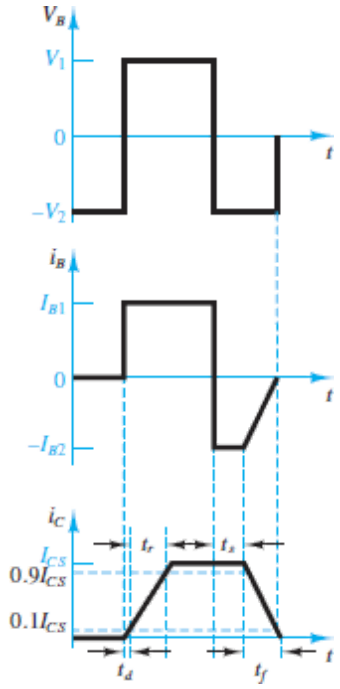


Figure 17.15 BJT switching waveforms

An analogous process (though the physics are different) takes place when the base voltage is reversed to drive the BJT into cutoff. The excess charge accumulated in the base must be discharged before the *BE* junction can be reverse-biased. This discharge takes place over a **storage time** t_s, which can be shortened by driving the base voltage to a negative value ($-V_2$). The reverse-biased *BE* junction capacitance must then be charged to ($-V_2$) before the switching transient is complete; this process takes place during the **fall time** t_f. In the figure, I_{CS} represents Page 17-14the collector saturation current. Thus, the turn-on and turnoff times of the BJT are given by

$$t_{on} = t_d + t_r \quad \text{and} \quad t_{off} = t_s + t_f \tag{17.3}$$

Power MOSFETs

MOSFETs are also used as power switches. The preferred mode of operation of a power MOSFET operated as a switch is the ohmic region, where substantial drain current can flow for relatively low drain voltages. Thus, a MOSFET switch is driven from cutoff to the ohmic state by the gate voltage. In an enhancement MOSFET, positive gate voltages are required to turn the transistor on; in depletion MOSFETs, either positive or negative voltages can be used.

Recall the parasitic capacitances that exist between pairs of terminals: C_{GS}, C_{GD}, and C_{DS}. Once again, these capacitances cause a **turn-on delay** $t_{d(\text{on})}$ corresponding to the time required to charge the equivalent input capacitance to the threshold voltage V_T. As shown in Figure 17.16, a finite rise time t_r is required to charge the gate-to-source voltage v_{GS} beyond first V_T and then V_{GSP} as the MOSFET transitions to the ohmic state.

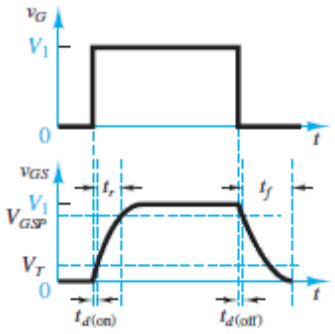


Figure 17.16 MOSFET switching waveforms

The **turnoff delay time** $t_{d(\text{off})}$ is the time required for the input capacitance to discharge so that v_{GS} can fall and v_{DS} can rise. The **fall time** t_f is the time required for v_{GS} to drop below the threshold voltage and turn off the transistor.

Insulated-Gate Bipolar Transistors

The insulated-gate bipolar transistor, or IGBT, is a hybrid device, combining features of both field-effect and bipolar devices. The circuit symbol of the IGBT is shown in Figure 17.1; a simplified equivalent circuit is shown in Figure 17.17. The IGBT is a voltage-controlled device, like a MOSFET, but its performance is closer to that of a BJT. The switching and conduction losses of the IGBT are lower than those of a MOSFET, and the switching speed is greater than that of a BJT (but somewhat lower than that of a MOSFET); the convenience of an insulated gate is an advantage over BJTs.

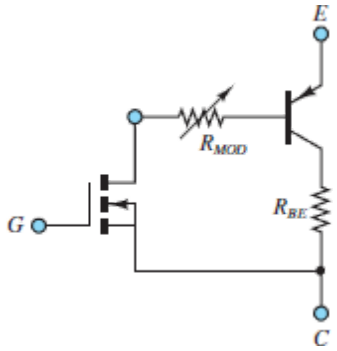


Figure 17.17 IGBT simplified equivalent circuit

IGBTs can be rated up to 400 A and 1,200 V, with switching frequencies as high as 20 kHz.

17.5 AC-DC CONVERTERS

One of the most immediate applications of the semiconductor diode is rectification of AC voltages and currents, to convert AC waveforms to direct current. Rectification can be achieved both with conventional diodes and with controlled diodes, such as thyristors. A simple diode rectifier can provide only a fixed DC voltage level; however, variable DC supplies can be easily obtained with the aid of thyristors.

Single-Phase Rectifiers and Snubbers

The load seen by a DC power supply can impact its performance. In practice, loads are often inductive, such as when a power supply drives a *DC motor*. To operate at a constant speed a DC motor requires a constant voltage. However, variations in the effective load due to variations in the

external work demanded of the motor can, if not accounted for, cause the motor voltage to vary as well.

Page 17-15

The top circuit in Figure 17.18 is a simple half-wave rectifier connected to an inductive load, which has a phase shift $\theta_Z = \tan^{-1}(\omega L/R)$ (i.e., time delay) between the voltage v_0 and the current i_0. For simplicity, assume an ideal diode D_1. Then, when D_1 is forward-biased, KVL requires $v_0 = v_{AC}$. On the other hand, when D_1 is reverse-biased, the current in the loop must be zero and $v_0 = 0$, as shown in the upper plot of Figure 17.19. However, since i_0 is delayed relative to v_0, when v_{AC} transitions from positive to negative, the current i_0 is not yet zero. The consequence is that the current through the inductor is interrupted resulting in a sudden spike in the voltage across the inductor as it attempts to maintain its current. This spike, known as a **flyback voltage**, is proportional to di_o/dt and could seriously damage the load and the diode.

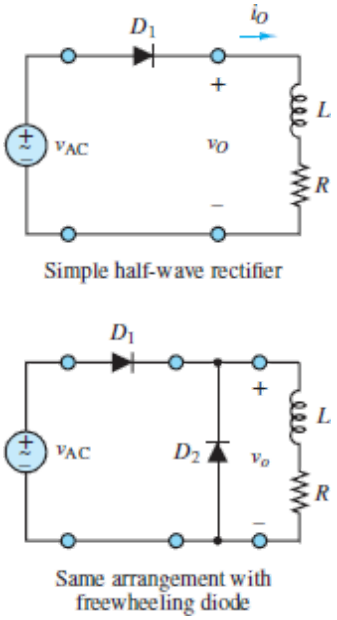


Figure 17.18 Rectifier connected to an inductive load

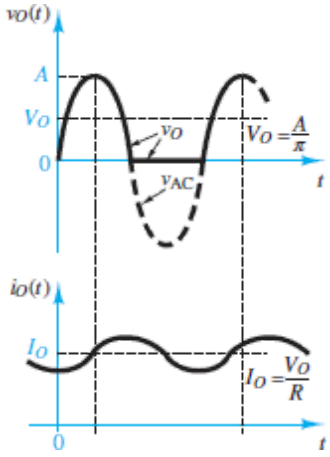


Figure 17.19 Operation of a freewheeling diode

The additional **freewheeling diode** D_2 in the bottom circuit of Figure 17.18 accommodates the current continuity requirement of the inductive load by providing a current pathway when v_{AC} is negative. Such a circuit is known as a **snubber**. Note that diode D_2 is off (reverse-biased) when v_{AC} is positive, but turns on (forward-biased) when v_{AC} is negative and D_1 is reverse-biased and no longer conducts. Consequently, the flyback voltage $L\,di_o/dt$ that developed in the top circuit is relieved because the inductor current can circulate through D_2 when v_{AC} is negative. The lower plot in Figure 17.19 depicts the load current, which is smooth and nearly constant, for the snubber circuit.

An analysis of the snubber circuit is fairly simple. For a sinusoidal voltage source $v_{AC}(t) = A \sin \omega t$ as indicated in Figure 17.19, and assuming ideal diodes, the load voltage is

$$v_0(t) = \begin{cases} A \sin \omega t & t \geq 0 \\ 0 & t \leq 0 \end{cases} \tag{17.4}$$

The time-average (DC) value of $v_0(t)$ is $V_0 = A/\pi$ as marked in Figure 17.19. Since the inductor acts like a short-circuit to this DC voltage, the DC current I_0 is simply:

$$I_0 = \frac{V_0}{R} = \frac{A}{\pi R} \tag{17.5}$$

Fourier analysis is required to compute the exact AC *ripple* current because v_0 is not a pure sinusoid. For the sake of simplicity, assume that most of the energy is at the AC source frequency. Then:

$$i_0(t) \approx I_0 + I_{AC}\cos(\omega t + \theta) \tag{17.6}$$

where I_0 is the average load current, I_{AC} is the peak value of the ripple current, and θ is its phase. The load voltage can also be approximated as shown in Figure 17.20:

$$v_0(t) \approx \frac{A}{2\pi} + \frac{A}{2\pi}\sin\omega t \tag{17.7}$$

This expression can be used to compute I_{AC}.

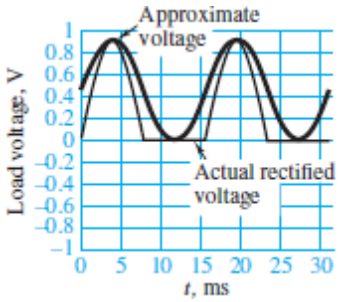


Figure 17.20 Approximation of ripple voltage for a half-wave rectifier

Three-Phase Rectifiers

It is important to realize that the same type of circuit that can be used for single-phase rectifiers can also be employed to design multiphase rectifiers. In many high-power applications, three-phase voltages need to be transformed and rectified to give rise to a single DC supply. Consider the balanced three-phase circuit shown in Figure 17.21. A three-phase wye-connected source is attached to a three-phase transformer, with a Page 17-16delta-connected primary and a wye-connected secondary. The three secondary currents i_a, i_b, and i_c encounter the diode pairs D_1-D_4, D_3-D_6, and D_5-D_2, respectively. The diodes will conduct in pairs depending on the relative line voltages, according to the following sequence: D_1-D_2, D_2-D_3,

D_3-D_4, D_4-D_5, D_5-D_6, and D_6-D_1. Recall that the line-to-line voltage is $\sqrt{3}$ greater than the phase voltage in a three-phase wye-connected source. The instantaneous source voltages and the related diode conduction periods, as well as the load voltage, are shown in Figure 17.22.

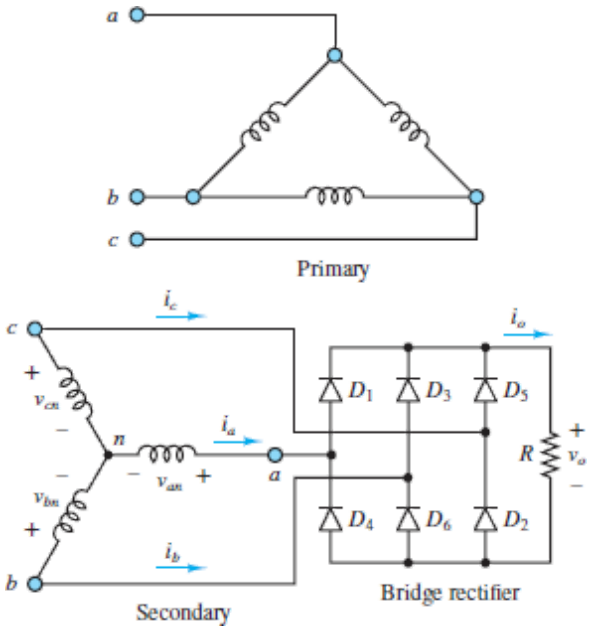


Figure 17.21 Three-phase diode bridge rectifier

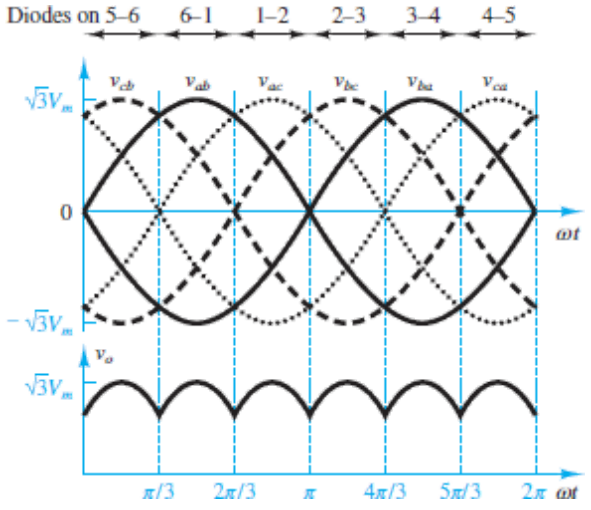


Figure 17.22 Waveforms and conduction times of three-phase bridge rectifier

The average output voltage is given by:

$$V_o = \frac{2}{2\pi/6}\int_0^{\pi/6} \sqrt{3}\,V_m \cos\omega t\, d(\omega t) = \frac{3\sqrt{3}}{\pi} V_m = 1.654 V_m \tag{17.8}$$

Page 17-17

where V_m is the peak phase voltage. The rms output voltage is

$$V_{\text{rms}} = \sqrt{\frac{2}{2\pi/6}\int_0^{\pi/6} 3V_m^2 \cos^2\omega t\, d(\omega t)} = \left[\frac{3}{2} + \frac{9\sqrt{3}}{4\pi}\right]^{1/2} V_m \tag{17.9}$$
$$= 1.6554 V_m$$

Thyristors and Controlled Rectifiers

In a number of applications, it is useful to externally control the current from an AC source to a load. A family of power semiconductor devices that provide this control by means of a third input, the **gate** *G*, are called **controlled rectifiers**. Arguably the oldest member of this family is the **silicon-controlled rectifier (SCR)**, or **thyristor**. Figure 17.23 depicts the internal structure of a thyristor, which consists of four layers of alternating *p*-type and *n*-type material, and its circuit symbol, which suggests that this device acts as a controlled diode.

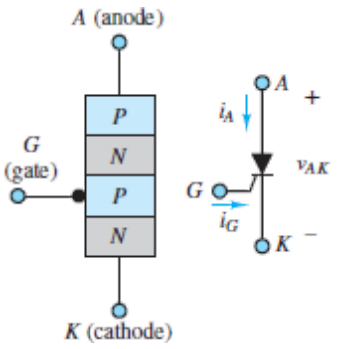


Figure 17.23 Thyristor structure and circuit symbol

When $v_{AK} \leq 0$, the thyristor is reverse-biased and acts as a conventional *pn* junction in the off state. When $v_{AK} > 0$, the thyristor is forward-biased but it will *not* conduct current unless a small current is injected into the gate or the applied forward voltage exceeds the so-called *breakover* voltage. Once the thyristor begins to conduct it will continue to do so, even in the absence of gate current as long as the forward current exceeds the *latching current.* The *holding current* is the minimum forward current required to maintain conduction, regardless of whether a gate current is present. Figure 17.24 shows typical characteristic *i-v* curves for a thyristor.

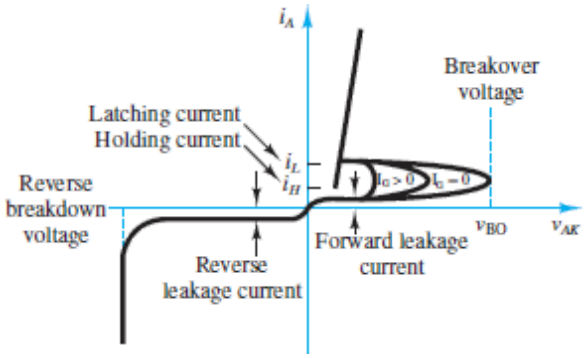


Figure 17.24 Thyristor *i-v* characteristic

The four-layer *pnpn* thyristor may also be modeled as a *pnp* transistor connected to an *npn* transistor, as shown in Figure 17.25, such that $i_{B_p} = i_{C_n}$ and $i_{B_n} = i_{C_p} + i_G$ Suppose, initially, that i_G and i_{B_n} are both zero. Then Q_n is in cutoff and $i_{C_n} = 0$ It then follows that $i_{B_p} = 0$ such that Q_p is also in cutoff and $iCp = 0$, a result which is consistent with the initial assumption. Thus, cutoff is a stable state, in the sense that unless an external condition perturbs the thyristor, it will remain off.

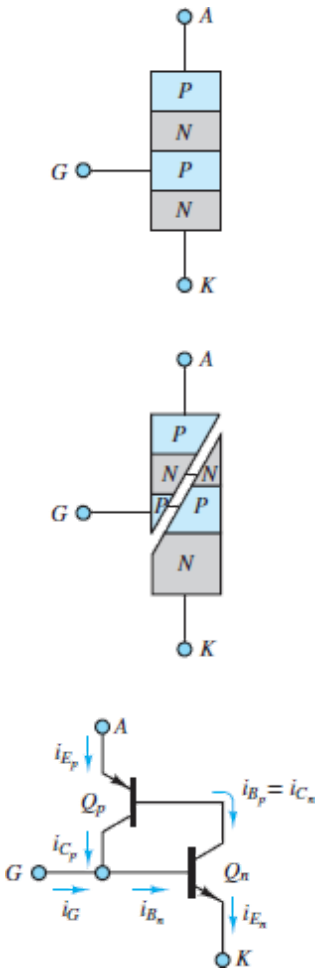


Figure 17.25 Thyristor two-transistor model

Suppose a small pulse of current is injected at the gate when the thyristor is in cutoff. Then $i_{B_n} > 0$ and Q_n will conduct, provided that $v_{AK} > 0$. It then follows that $i_{C_n} = i_{B_p} > 0$ such that Q_p also will conduct and $i_{E_p} > 0$ and $i_{C_p} > 0$. Once this condition is achieved the thyristor will remain on and continue to conduct even if i_G is subsequently set to zero since $i_{B_n} = i_{C_p} + i_G$ remains nonzero. The thyristor reverts to the off state when v_{AK} becomes negative such that both transistors return to cutoff.

Page 17-18

The general waveforms for the circuit of Figure 17.26 can be analyzed as follows. Assume a pulse voltage v_{trigger} is applied to the thyristor gate at $t = \tau$, as depicted in Figure 17.27, such that the thyristor begins to conduct,

and it continues to do so until v_S is negative. The average (DC) component of the load voltage v_0 is zero prior to $t = \tau$.

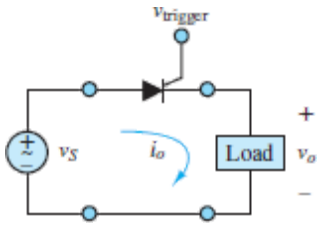


Figure 17.26 Controlled rectifier circuit

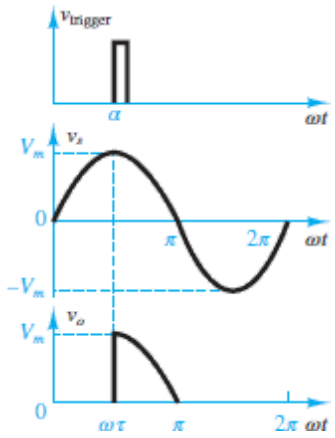


Figure 17.27 Half-wave controlled rectifier waveforms

$$\langle v_0 \rangle = V_0 = \frac{1}{T}\int_{\tau}^{T/2} v_S(t)\,dt \tag{17.10}$$

where T is the period of $v_S(t)$. Let:

$$v_S(t) = A\ \sin \omega t \tag{17.11}$$

to express the average (DC) value of the load voltage as:

$$\langle v_0 \rangle = V_0 = \frac{1}{T}\int_{\tau}^{T/2} A\ \sin\ \omega t\ dt = (1 + \cos \omega\tau)\frac{A}{2\pi} \tag{17.12}$$

where α is the **firing angle**, defined by:

$$\alpha = \omega\tau \tag{17.13}$$

The integral of equation 17.12 can be evaluated to show the dependence of $\langle v_0 \rangle = V_0$ on the firing angle α:

$$\langle v_0 \rangle = V_0 = (1 + \cos\alpha)\frac{A}{2\pi} \tag{17.14}$$

EXAMPLE 17.4 Thyristor-Based Variable Voltage Supply

Problem

Analyze the thyristor-based variable voltage supply shown in Figure 17.28. Determine (1) the rms load voltage as a function of the firing angle and (2) the power supplied to the resistive load at zero firing angle and at firing angles equal to $\pi/2$ and π.

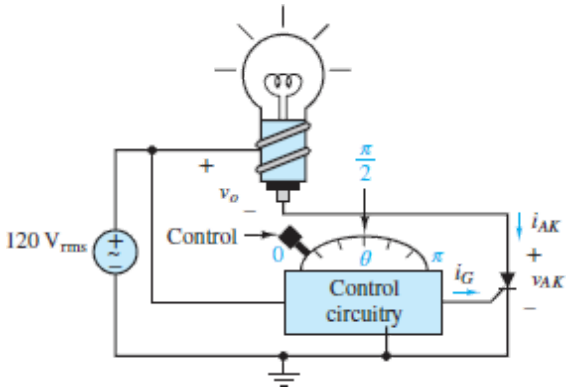


Figure 17.28

Solution

Known Quantities: Load resistance R_0 of the lamp; input voltage source $v_S(\omega t) = V_S \sin(\omega t)$.

Find: $\tilde{V}_0$, $P_0|_{\alpha=0}$, $P_0|_{\alpha=\pi/2}$, $P_0|_{\alpha=\pi}$.

Schematics, Diagrams, Circuits, and Given Data: $V_{AK\ (on)} = 0$ V; $R_0 = 240\ \Omega$; $V_S = 120\sqrt{2}$ V The pulsed gate current $i_G(t)$ is shown in Figure 17.29.

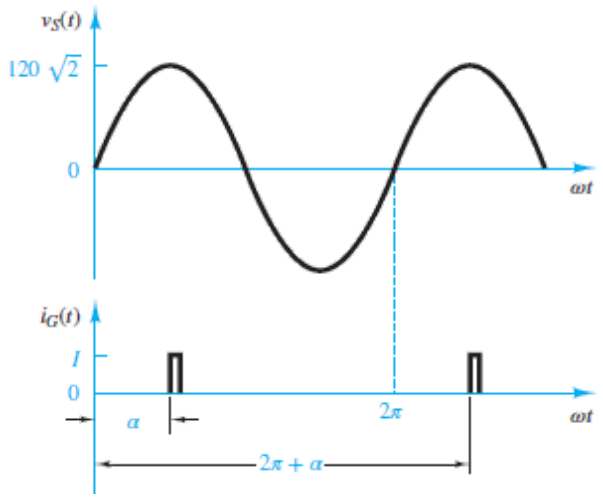


Figure 17.29

Assumptions: The thyristor acts as an ideal diode when on ($V_{AK} > 0$).

Analysis:

1. *Load voltage calculation.* The load voltage is shown in Figure 17.30. The rms value of the load voltage as a function of the firing angle α is computed by:

$$\tilde{V}_0(\alpha) = \sqrt{\frac{1}{T}\int_0^T v_o^2(t)\,dt}$$

$$= \sqrt{\frac{1}{\omega T}\int_0^T v_o^2(t)\,\omega\,dt}$$

$$= \sqrt{\frac{1}{2\pi}\int_0^{2\pi} v_o^2(\omega t)\,d(\omega t)}$$

$$= \sqrt{\frac{1}{2\pi}\left[\int_0^{\alpha} 0\,d(\omega t) + \int_{\alpha}^{\pi} V_S^2 \sin^2(\omega t)\,d(\omega t) + \int_{\pi}^{2\pi} 0\,d(\omega t)\right]}$$

$$= \sqrt{\frac{V_S^2}{2\pi}\int_{\alpha}^{\pi} \sin^2(\theta)\,d\theta}$$

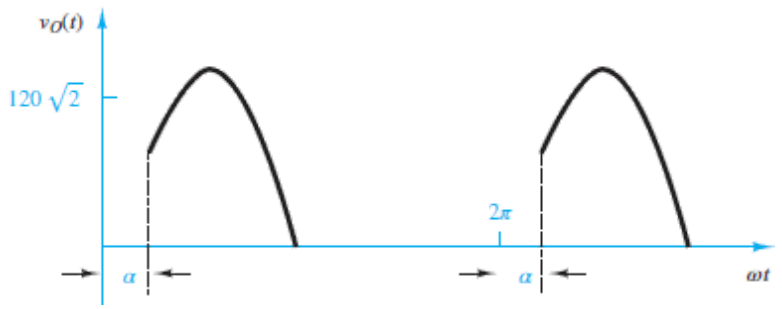


Figure 17.30

Page 17-80

where $\theta = \omega t$. Substitute $V_s = 120\sqrt{2}$ and use the trigonometric identity $\sin^2(\theta) = [1 - \cos(2\theta)]/2$ to evaluate the integral.

$$\tilde{V}_0(\alpha) = \sqrt{\frac{(120\sqrt{2})^2}{2\pi}\int_\alpha^\pi \frac{1-\cos(2\theta)}{2}\,d(\theta)}$$

$$= \frac{120\sqrt{2}}{2}\sqrt{\frac{2}{\pi}\left[\frac{\theta}{2} - \frac{\sin(2\theta)}{4}\right]_\alpha^\pi}$$

$$= \frac{120\sqrt{2}}{2}\sqrt{\left[\frac{\theta}{\pi} - \frac{\sin(2\theta)}{2\pi}\right]_\alpha^\pi}$$

$$= \frac{120\sqrt{2}}{2}\sqrt{\frac{\pi-\alpha}{\pi} + \frac{\sin(2\alpha)}{2\pi}}$$

2. *Load power calculation.* We can now compute the load power for each of the three values of α:

$$P_0 = \frac{\tilde{V}^2}{R_0}$$

For $\alpha = 0$:

$$P_0 = \frac{\tilde{V}^2}{R_0} = \frac{(120\sqrt{2}/2)^2}{240} = 30 \text{ W}$$

For $\alpha = \pi/2$:

$$P_0 = \frac{\tilde{V}^2}{R_0} = \frac{[(120\sqrt{2}/2)\sqrt{1-\frac{1}{2}}]^2}{240} = 15 \text{ W}$$

For $\alpha = \pi$:

$$P_0 = \frac{\tilde{V}^2}{R_0} = \frac{[(120\sqrt{2}/2)\sqrt{1-1}]^2}{240} = 0 \text{ W}$$

Comments: Note that no power is wasted when the firing angle is set for zero load voltage. This would not be the case if a resistive voltage divider were used to adjust the load voltage.

EXAMPLE 17.5 Automotive Battery Charger

Problem

Qualitatively explain the operation of the automotive battery charger shown in Figure 17.31.

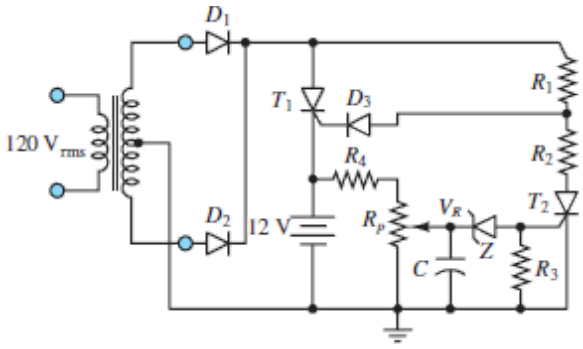


Figure 17.31 Automotive battery charger

Solution

Analysis: The charging circuit is connected to a standard 110-V single-phase supply. Diodes D_1 and D_2 form a full-wave rectifier; resistors R_1 and R_2 and

thyristor T_2 form a variable-voltage divider.

Assume that thyristor T_2 is not in the conducting state and that the anode voltage of D_3 is such that D_3 conducts. Then T_1 will be fired near the beginning of the positive half-cycle of the AC source voltage, and its period of conduction will be long, providing a substantial current to the battery (resistors R_4 and R_p are sufficiently large that most of the current flowing through T_1 will go to the battery).

The potentiometer R_p is set so that when the battery voltage is low, the voltage V_R is not sufficient to turn on the Zener diode Z. Thus, Z is effectively an open-circuit, and T_2 remains off (recall that we had initially assumed T_2 to be off—this confirms the correctness of the assumption). As the battery charges to a progressively higher value, Z will eventually conduct; when Z conducts, a gate current is injected into T_2, which is then turned on.

When T_2 conducts, the voltage across the R_2-T_2 series connection becomes significantly lower, because T_2 is now nearly a short-circuit. Resistors R_1 and R_2 are selected so that when T_2 conducts, D_3 becomes reverse-biased. Once this condition occurs, T_1 is turned off and charging stops. Note that the circuit has built-in overcharging protection.

EXAMPLE 17.6 Thyristor Circuit

Problem

Determine the value of R in the circuit of Figure 17.32 such that the average current through the thyristor is 1 A.

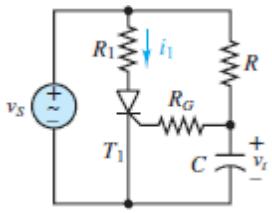


Figure 17.32

Solution

Known Quantities: Resistances and source voltage.

Find: Resistor R such that $\langle i_1 \rangle = 1$ A.

Schematics, Diagrams, Circuits, and Given Data: $v_S = 200$ V rms, 250 Hz; $V_{AK\,(\text{on})} = 0$ V; $R_1 = 75\ \Omega$; $R_G = 1$ kΩ; $C = 1\ \mu$F.

Assumptions: The thyristor acts as an ideal diode when on ($V_{AK} > 0$).

Analysis: Figure 17.33 depicts the relative timing of the source voltage $v_S(t)$, thyristor current $i_1(t)$, and triggering voltage $v_t(t)$. The source voltage is

$$v_S(t) = \sqrt{2} \times 200 \sin(2\pi \times 250t) \qquad \text{(recall that } \omega = 2\pi f\text{)}$$

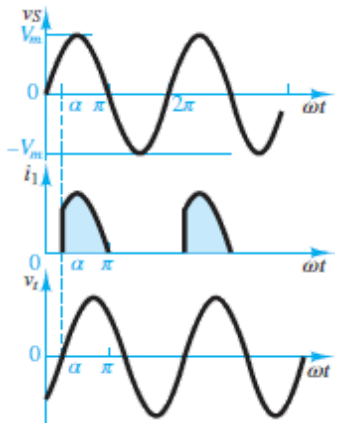


Figure 17.33

The current through $R_1 = 75\ \Omega$ is

$$i_1(t) = \begin{cases} 0 & 0 \le \omega t < \alpha \\ \dfrac{\sqrt{2} \times 200}{75} \sin(2\pi \times 250t) & \alpha < \omega t \le \pi \\ 0 & \pi \le \omega t \le 2\pi \end{cases}$$

and the triggering voltage is

$$v_t(t) = V_t \sin(2\pi \times 250t - \alpha)$$

Page 17-22

The triggering voltage will go positive at the desired firing angle α, thus injecting a current into the gate of the thyristor and turning it on. To set $\langle i_1 \rangle = 1$ A requires:

$$\langle i_1(t) \rangle = \frac{1}{\pi} \int_{\alpha}^{\pi} \frac{\sqrt{2} \times 200}{75} \sin(\omega t')\, d(\omega t') = 1 \text{ A}$$

Upon evaluating the integral, the requirement is

$$\frac{\sqrt{2} \times 200}{\pi \times 75}(1 + \cos\alpha) = 1$$

for which $\alpha = 99.6°$.

Observe that the AC source voltage appears across RC. If the gate current is assumed to be negligibly small, v_t can be found (in phasor notation) by voltage division:

$$\begin{aligned} \mathbf{V}_t(j\omega) &= \frac{1/j\omega C}{R + 1/j\omega C} V_S(j\omega) \\ &= \frac{V_S}{\sqrt{1 + \omega^2 R^2 C^2}} \angle[-\arctan(\omega RC)] \end{aligned}$$

To solve for R observe that the phase of $V_t(j\omega)$ is the firing angle α. Thus:

$$-\arctan(\omega RC) = \alpha = 99.6°$$

$$R = -\frac{\tan(\alpha)}{\omega C} = \frac{\tan(99.6)}{2\pi \times 250 \times 10^{-6}} = 3759.5\ \Omega$$

CHECK YOUR UNDERSTANDING

Using the approximations given in equations 17.6 and 17.7, find the DC and AC loads for the circuit of Figure 17.18 if $R = 10\ \Omega$, $L = 0.3$ H, $A = 170$ V, and $\omega = 377$ rad/s.

Answer: $I_0 = 5.4$ A, $I_{AC} = 0.75$ A, $\theta = -84.95°$

CHECK YOUR UNDERSTANDING

Calculate the rms load voltage in the circuit of Figure 17.26 for $A = 100$ V and $\alpha = \pi/3$ rad. Let the input AC rms voltage be 240 V in the circuit of Example 17.4. Find the rms value of the load voltage and the power dissipated in the load if the firing angle $\alpha = \pi/4$ rad.

Answer: $V_{0\,\text{rms}} = 44.85$ V; $V_{0\,\text{rms}} = 161.81$ V, $P_0 = 109$ W

CHECK YOUR UNDERSTANDING

Compute the rms value of the load current in the circuit of Figure 17.32.

Answer: $\frac{V_m}{2R_1}\sqrt{\frac{\pi-\alpha}{\pi}+\frac{\sin(2\alpha)}{2\pi}}$

Conclusion

The following learning objectives should have been mastered upon completion of this chapter:

1. *Learn the classification of power electronic devices and circuits.* Power electronic devices can handle up to a few thousand volts and up to several hundred amperes and are used in many industrial applications. The various families of power electronics devices and systems are introduced in this section.
2. *Analyze the operation of practical voltage regulators.* Voltage regulators are a basic element of DC power supplies and are used to provide a stable voltage output. The principal element of a voltage regulator is the Zener diode, used as a voltage reference.
3. *Understand the principal limitations of transistor power amplifiers.* Transistors, both BJT and MOSFET, as well as IGBTs are commonly used as power amplifiers and switches. The principal limitations in these applications are the allowable power dissipation and the switching time.
4. *Analyze the operation of single- and three-phase controlled rectifier circuits.* Rectifiers, like voltage regulators, are essential elements in DC power supplies. Controlled rectifiers can provide a variable DC voltage, and they are therefore very useful in the design of power converters for DC motors, and in other industrial applications that require variable DC voltage supplies.

HOMEWORK PROBLEMS

Section 17.2: Voltage Regulators

17.1 Repeat Example 17.1 for a 7-V Zener diode.

17.2 For the current regulator circuit shown in Figure P17.2, derive an expression for R_S in terms of V_Z and I.

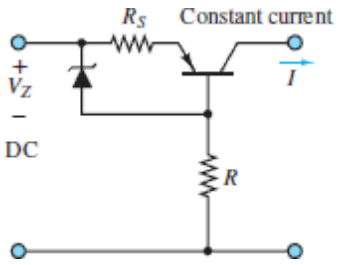


Figure P17.2

17.3 For the shunt-type voltage regulator shown in Figure P17.3, find an expression for the output voltage V_o in terms of R_S, V_Z, and R.

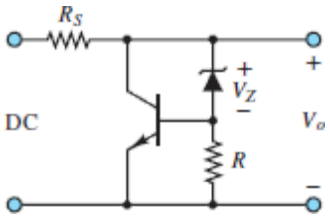


Figure P17.3

Section 17.3: Power Amplifiers

17.4 The circuit of Figure P17.4 is a very effective battery charger. Its operation is simple, and the TIP-33C *npn* power transistor can sink 10 A if a big enough heat sink is used. Assuming that the transistor remains in the active operating region, determine the power delivered to the 1.2-V rechargeable battery in the circuit.

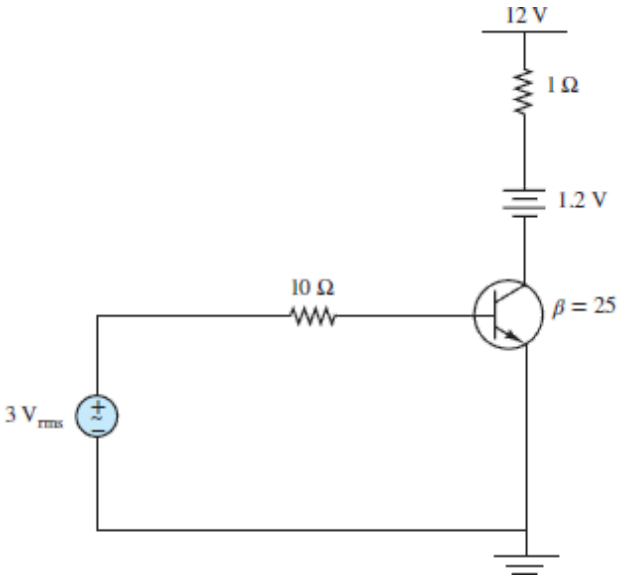


Figure P17.4

Page 17-24

17.5 An IGBT can be modeled as shown in the circuit of Figure P17.5. With $V_t = 4$ and $K = 0.01$ A/V^2 for the MOSFET, and $\beta = 200$ for the BJT, determine the current through R_o and the voltage across it. Let $V_G = 8$ V and $R_o = 2\ \Omega$.

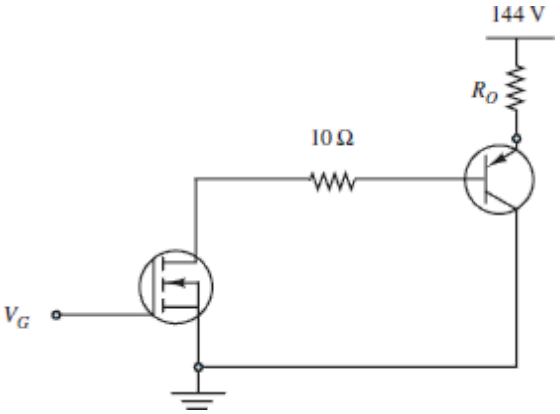


Figure P17.5

Section 17.5: AC-DC Converters

17.6 Consider the half-wave rectifier shown in Figure 17.18. Assume $v_{AC}(t) = V_M \sin(\omega t)$. Determine the current $i_o(t)$ for one period of the source waveform with and without the freewheeling diode. Assume ideal diodes D_1 and D_2.

17.7 Consider the half-wave rectifier shown in Figure 17.18. Determine $v_o(t)$ with and without the freewheeling diode. Assume each diode has a forward resistance of 50 Ω and a forward bias voltage of 0.7 V. Also assume $R = 10$ Ω and $L = 2$H.

17.8 For the circuit shown in Figure P17.8, v_{AC} is a sinusoid with a 12-V peak amplitude, $R = 3$ kΩ, and the forward-conducting voltage of D is 0.65 V.

a. Sketch the waveform of $v_o(t)$.

b. Find the average value of $v_o(t)$.

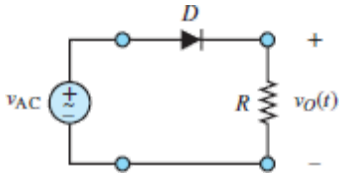


Figure P17.8

17.9 A vehicle battery charge circuit is shown in Figure P17.9. Describe the circuit, and draw the output waveform (L_1 and L_2 represent the inductances of the windings of the alternator).

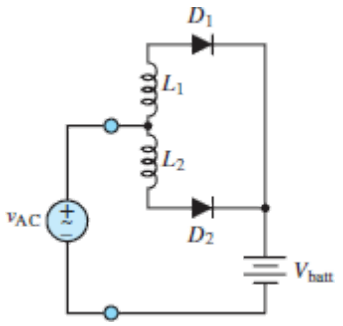


Figure P17.9

17.10 Repeat Example 17.4 for $\alpha = \pi/4$ and $\alpha = \pi/8$.

17.11 The circuit shown in Figure P17.11 is a speed control system for a DC motor. Assume $v_{AC}(t) = 115\sqrt{2}\sin(\omega t)$ the thyristors S_1 and S_2 are fired at α = 60°, and that the motor direct current is 20 A and is ripple-free.

a. Sketch the output voltage waveform v_o.

b. Compute the average power (in watts) absorbed by the motor.

c. Compute the apparent power (in volt-amperes) supplied by the source.

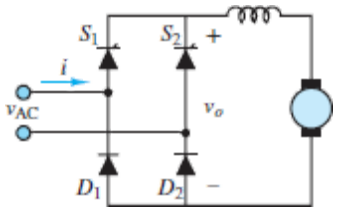


Figure P17.11

17.12 Consider a full-wave, single-phase controlled rectifier, similar to that shown in Figure 17.2, that is used to drive a DC motor instead of the resistive load R_o. Assume that the motor is rated at 4 kW and 110 V_{rms} and that the AC supply is 80 Vrms 60 Hz. Also assume that the motor inductance is very large such that the motor current is ripple-free and that the motor constant is 0.055 V/rpm. If the motor runs at 1,000 rpm at rated current:

a. Determine the firing angle of the converter.

b. Determine the rms value of the supply current.

17.13 For the light dimmer circuit of Example 17.4, determine the load power at firing angles α = 0°, 30°, 60°, 90°, 120°, 150°, and 180°, and plot the load power as a function of α.

17.14 In the circuit shown in Figure P17.14:

$V_o = 10\,\text{V}$ $\quad V_{\text{ripple}} = 10\% = 1\,\text{V}$

$I_o = 650\,\text{mA}$ $\quad v_{\text{line}} = 170\cos\omega t$

$\omega = 2{,}513\,\text{rad/s}$

Assume $v_o = V_o + V_{\text{ripple}}(t)$ and the diodes are fabricated from silicon. Determine the conduction angle of the diodes.

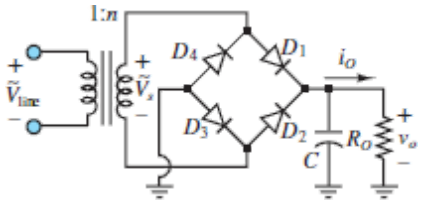


Figure P17.14

17.15 Assume that the conduction angle of the silicon diodes shown in the circuit of Figure P17.15 is

$\phi = 23°$

$v_{s1}(t) = v_{s2}(t) = 8\cos\omega t \quad \text{V}$

$\omega = 377\,\text{rad/s}$ $\quad R_o = 20\,\text{k}\Omega$

$C = 0.5\,\mu\text{F}$

Determine the rms value of the ripple voltage.

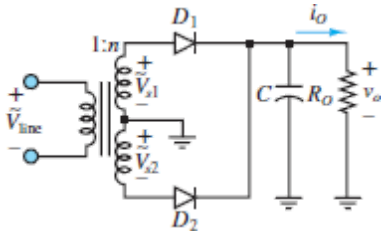


Figure P17.15

17.16 The diodes in the full-wave DC power supply shown in Figure P17.14 are silicon. Assume:

$I_o = 85$ mA $\quad V_o = 5.3$ V

$V_{ripple} = 0.6$ V $\quad \omega = 377$ rad/s

$v_{line} = 156 \cos \omega t$ V

$C = 1{,}023\ \mu$F

ϕ = conduction angle = 23.90°

Determine the value of the average and peak current through each diode.

17.17 The diodes in the full-wave DC power supply shown in Figure P17.15 are silicon. Assume:

$I_o = 600$ mA $\quad V_o = 50$ V

$V_{ripple} = 8\% = 4$ V $\quad v_{line} = 170 \cos \omega t$

Determine the value of the conduction angle for the diodes and the average and peak current through the diodes. The load voltage waveform is shown in Figure P17.17.

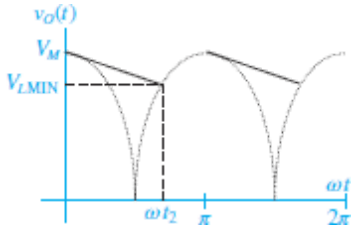


Figure P17.17

17.18 A power supply is shown in Figure P17.18. Sketch the signals V_{ab}, V_{cd}, V_{ef}, and V_{gh}. Assume $V_z = 8$ V and silicon rectifier diodes.

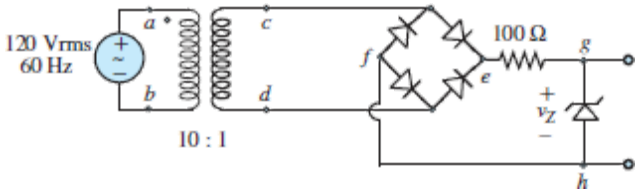


Figure P17.18

17.19 A power supply is shown in Figure P17.19. Sketch the signals V_{ab}, V_{cd}, V_{ef}, and I_Z. Assume $V_Z = 6$ V and a silicon rectifier diode.

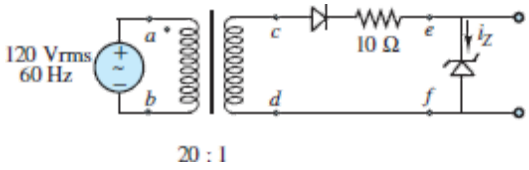


Figure P17.19

17.20 Figure P17.20 depicts a low-cost full-wave bridge rectifier with a Zener diode voltage regulator. Sketch the voltages across terminals *a*-*b*, *c*-*d*, and *e*-*f*. The transformer turns ratio is 10:1 (step down). Assume $V_Z = 12$ V and silicon rectifier diodes.

Page 17-26

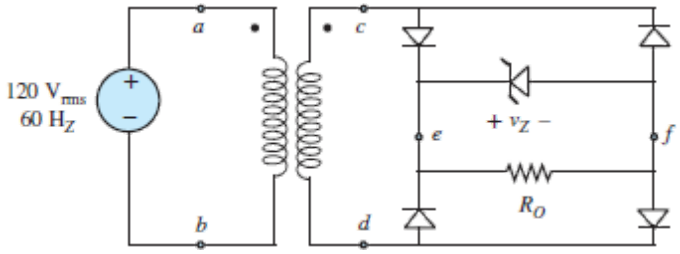


Figure P17.20

17.21 In the DC power supply shown in Figure P17.21, sketch the voltage across *a*-*b*, *c*-*a*, and *d*-*e*, assuming that *R* is so large as to make any ripple negligible. Assume silicon rectifier diodes.

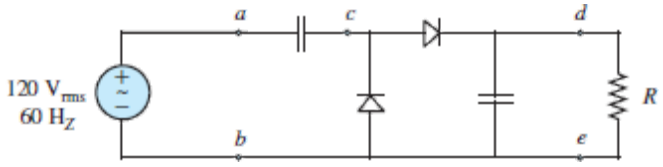


Figure P17.21

17.22 A DC power supply known as a *voltage doubler* is shown in Figure P17.22. It is assumed that the capacitors are large enough that the ripple is not significant in the output voltage. Sketch the signals v_{ab} and v_{cd}. Assume silicon rectifier diodes.

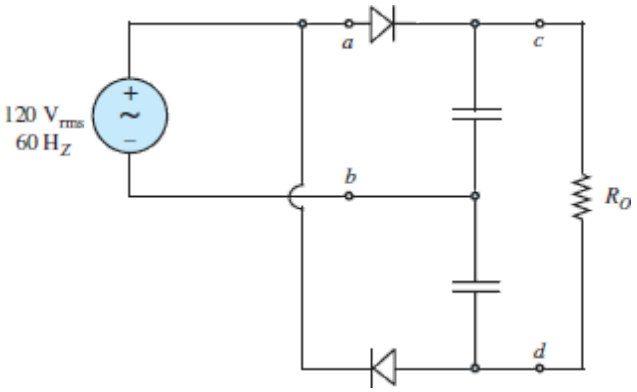


Figure P17.22

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

C H A P T E R

18

ANALOG COMMUNICATION SYSTEMS

This chapter introduces the foundations of electrical communication systems, emphasizing basic analog communications ideas.

The subject of electrical communications is one that touches everyone's life: telephones, TV, and radio have been a part of our lives for many decades. Today, new means of communications are as pervasive as the traditional ones. Computer networks, satellite weather systems, and personal communication systems (pagers, cellular phones, etc.) are essential parts of our everyday lives. The aim of this chapter is to present the basic mathematics of spectrum analysis, which are the foundations of all communication systems, and the basic operation of amplitude- and frequency-modulation systems. The explanation of these concepts is supplemented by the use of computer-aided tools. In addition, the chapter also includes an overview of different types of commonly used communication systems.

Learning Objectives

Students will learn to...

1. Be familiar with the most common types of communication systems in block diagram form. *Sections 18.1 and 18.5*.
2. Be capable of performing spectral analysis of simple signals using analytical and computer-aided tools. *Section 18.2*.
3. Understand the principles of amplitude modulation (AM) and demodulation and perform basic calculations and numerical computations on AM signals *Section 18.3*.
4. Understand the principles of frequency modulation (FM) and demodulation and perform basic calculations and numerical computations on FM signals *Section 18.4*.

18.1 INTRODUCTION TO COMMUNICATION SYSTEMS

The modern era of communications began with the **telegraph** and the **Morse code** and rapidly moved toward radio and television. Table 18.1 summarizes some of the major dates in the history of communication systems.

Table 18.1 A brief history of communications

Date	Event
1838	Samuel F. B. Morse demonstrates telegraph
1876	Alexander Graham Bell patents the telephone
1897	Marconi patents a complete wireless telegraph system
1906	Lee DeForest invents the triode amplifier
1915	Bell System completes a transcontinental telephone line
1918	B.H. Armstrong perfects the superheterodyne receiver
1920	Commercial radio broadcasting
1934–45	Radar and microwave systems are developed
1937	Alec Reeves conceives pulse code modulation
1938	Television broadcasting begins
1948	The transistor is invented; Claude Shannon publishes Mathematical Theory of Communications
1956	First transoceanic telephone cable
1960	First communications satellite, Telstar I, is launched
1962–66	High-speed digital communications
1965	Mariner IV transmits pictures from Mars to Earth
1969	ARPANET
1970	Color TV
1970	Commercial relay satellite telecommunications
1974	TCP/IP protocol published
1975	Intercontinental computer communication networks
1983	First commercial mobile phone
1990	World Wide Web and first commercial ISP introduced
1991	First digital cellular network is introduced

Information, Modulation, and Carriers

The purpose of communication systems is to communicate *information;* the four most common sources of information are *speech* (or *sound*), *video*, and *data*. Regardless of Page 18-3the source, the information that is transmitted and received in a communication system consists of a signal, which has the information encoded in some appropriate fashion. Figure 18.1 depicts the general layout of a communication system: an **input transducer** (e.g., a microphone) converts the **input message** into a **message signal** (e.g., a time-varying voltage) that is transmitted over a **channel** and converted by a **receiver** into an **output signal**. An **output transducer** (e.g., a loudspeaker) converts the received signal into an **output message** (e.g., sound). The transmitter performs a very important function on communication signals by encoding the signals in some fashion making use of a **carrier signal**. The information is contained in a so-called **modulating signal** that *modulates* a carrier signal. For example, in FM radio the modulating signal consists of speech and music, and the carrier is a sinusoidal wave of predetermined frequency, much higher than the modulating signal frequency. Table 18.2 summarizes the **frequency band allocation** and typical applications in each frequency band.

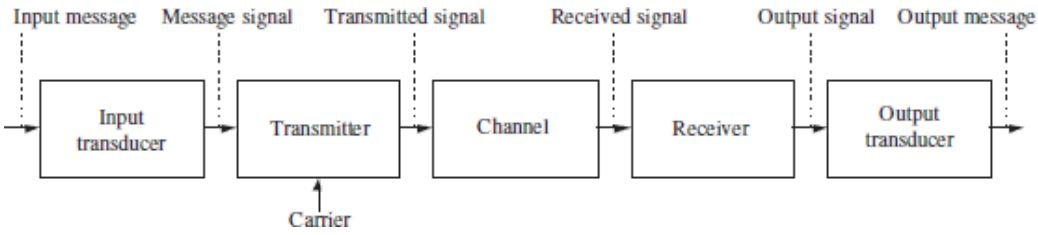


Figure 18.1 Block diagram of a communication system

Table 18.2 Frequency bands

Frequency band	Name	Medium	Applications
3–30 kHz	Very low frequency (VLF)	Wire pairs	Long-range navigation, sonar
30–300 kHz	Low frequency (LF)	Wire pairs	Navigational aids, radio beacons
300–3,000 kHz	Medium frequency (MF)	Coaxial cable	Maritime radio, direction finding, Coast Guard, commercial AM radio
3–30 MHz	High frequency (HF)	Coaxial cable	Search and rescue, aircraft communications with ships, telegraph, telephone, and facsimile
30–300 MHz	Very high frequency (VHF)	Coaxial cable	VHF television channels, FM radio, private aircraft, air traffic control, taxi cabs, police
0.3–3 GHz	Ultra high frequency (UHF)	Coaxial cable, waveguide	UHF television channels, surveillance radar, satellite communications
3–30 GHz	Super high frequency (SHF)	Waveguide	Satellite communications, airborne radar, approach radar, weather radar, land mobile
30–300 GHz	Extremely high frequency (EHF)	Waveguide	Railroad service, radar landing systems, experimental
>300 GHz	Optical frequencies	Optical fiber	Wideband data, experimental

There are two principal reasons for the use of a very broad *spectrum* of carrier frequencies. The first is that allowing for a broad spectrum permits many simultaneous users to broadcast information at different frequencies without interference among different transmissions; the second is that depending on the frequency of the carrier, the electromagnetic waves that are transmitted have different propagation characteristics. Thus, different carrier frequencies are better suited for propagating over long distances than others.

Classification of Communication Systems

Communication systems can be classified into two basic families, based on the nature of the message signal: **analog communication systems** and **digital communication systems**. In this chapter we focus on analog communications although it should be noted that digital communications are taking an increasingly

prominent role even in the most common applications.[1] Digital communications are covered in Chapter 19. Another classification may be made based on the type of transmission: *light wave* versus *radio frequency*, or RF transmission, as is explained in the next section. A third classification is that of *carrier* versus *direct baseband* transmission system. This latter classification is based on whether the signal of interest is directly transmitted (e.g., as in the case of the telegraph), or whether the signal *modulates* a *carrier wave*, as in the case of AM and FM radio transmission.

Communication Channels

The modulated transmitted signal can reach the receiver in a number of ways. In some cases, communication systems are hard wired. Examples of this configuration are local area computer networks, local telephone systems, and local cable TV networks. Depending on the frequency range, the transmitted signal can be carried by **twisted wire pair**, **coaxial cable**, **waveguides**, or **optical fiber**. However, in most communications systems, after the signal has been carried over a wire or cable, it is eventually broadcast over air by an antenna, to be received by a similar antenna elsewhere. Figure 18.2 depicts some typical communication system components.

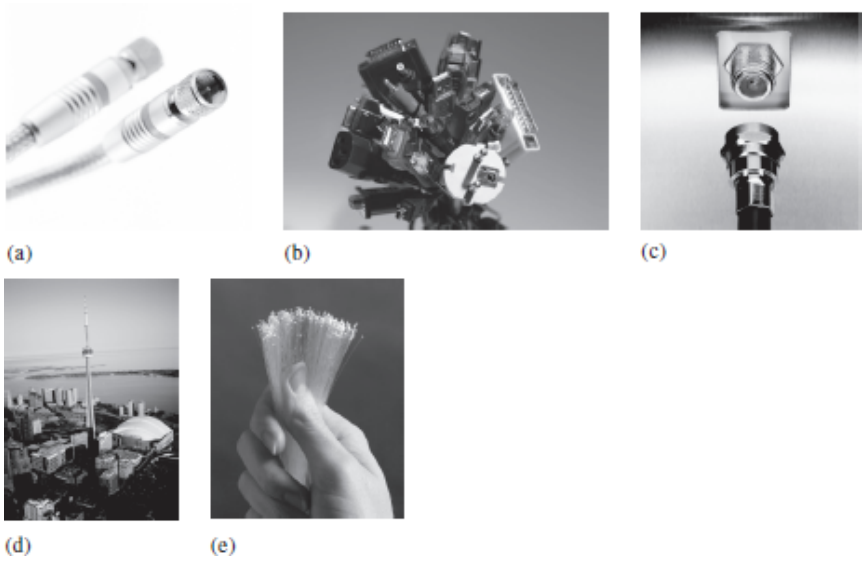


Figure 18.2 Communication system components; clockwise from top left: (a) coaxial cables; (b) RF cabling components; (c) detail of coaxial cable; (d) monopole antenna; (e) optical fiber bundle (*Photos:* (a) *Kevin Jordan/Getty Images* (b) *Jochen Tack/imageBROKER/age fotostock* (c)

Don Farrall/Getty Images (d) *Sylvain Grandadam/Getty Images* (e) *Glow Images*)

Page 18-5

The range of transmission can be significant—consider that signals can be received from the far reaches of the solar system via **radio astronomy**. The most common means of transmission of communication signals is via the broadcast of **radio frequency waves** over the air. To understand the different types of wave propagation, we need to briefly explain the geometry of the earth's atmosphere. With reference to Figure 18.3, the atmosphere is composed of layers, of which the **troposphere** and the **ionosphere** are the most important for radio wave transmission. The troposphere (up to about 20 km above sea level) is where the earth's air is contained; air density, temperature, and humidity decrease with increasing altitude. The propagation of radio waves in air depends on various properties of the medium. The speed of propagation of electromagnetic waves and the refractive index of the medium (causing the deflection of the wave) increase with altitude. As a consequence, radio waves tend to bend back toward the earth as they propagate through the troposphere.

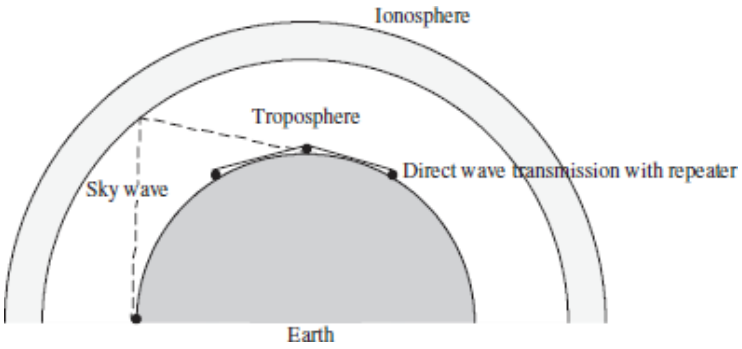


Figure 18.3 Propagation of radio frequency waves

The ionosphere is so called because of the ionization of the small amounts of air present at these altitudes (50 to 600 km). Electromagnetic waves reaching the ionosphere may propagate through it with some losses (attenuation) or may be reflected down to earth, depending on the frequency of the transmissions. In general, frequencies above 30 MHz will propagate through the ionosphere and are therefore suitable for **space communications**.

To achieve long-range communications over the earth, **sky waves** are used. These waves are reflected by the ionosphere and can reach points beyond the horizon. These waves must have frequencies below 30 MHz to allow them to be

reflected by the ionosphere. Short-wave radio makes use of sky waves. **Tropospheric waves** can also propagate beyond the horizon, but instead of being reflected, as in the case of sky waves, they bend around the earth due to diffraction. **Direct waves** are used in *line-of-sight* transmission, where the transmitter and receiver are in the line of "sight" of one another. The earth's curvature is the primary limitation to the distance of such transmissions; however, due to reflections from the ground, and to ground and surface waves, this transmission can achieve greater distances than one would calculate simply based on the earth's curvature and the height of the antennas.

Coaxial cables are very commonly used for the transmission of radio-frequency waves over short to medium distances, typically in the frequency range between a fraction of a megahertz to hundreds of megahertz. Coaxial cable consists of a copper core, surrounded by an insulating layer, in turn surrounded by a conductive (ground) layer and by an external protective sheath. Today, the most common Page 18-6example of the use of coaxial cable is the distribution of cable television signals from the receiving station to individual homes.

An increasingly common type of communication system is based on **light wave transmission**. Light is also electromagnetic radiation, but at much higher frequencies than radio waves. The main drawback in the use of light as a carrier is that it needs to be enclosed in a guide to travel over significant distances; **optical fibers** are used to achieve such transmission. An optical fiber consists of a hair-thin strand of glass, the *core*, surrounded by a protective layer, the *cladding*. Snell's law of optics ensures that if light enters the fiber at a sufficiently low angle of incidence, the transmission benefits from *total internal reflection*, confining the light signal to the core with minimal losses. High-speed computer communications networks are increasingly making use of optical fibers.

18.2 SPECTRAL ANALYSIS

Signal Spectra

Signals can be represented in time-domain and frequency-domain forms. Phasor notation is the starting point of the frequency-domain representation, or **spectral representation**, of signals: a phasor describes a sinusoidal signal's amplitude and phase as a function of frequency. The **spectrum** of a signal is its frequency-domain representation. For example, the signal $x(t) = A_1\cos(\omega_1 t + \phi_1)$ only contains a single sinusoidal frequency, ω_1, and its spectrum therefore consists of a pair of **spectral lines** at the frequency $\pm\omega_1$.[2] Figure 18.4(a) to (c) depict the

representation of a sinusoidal signal in the *time domain* and in the *frequency domain*. A complete representation of the frequency domain requires both *magnitude* and *phase*.

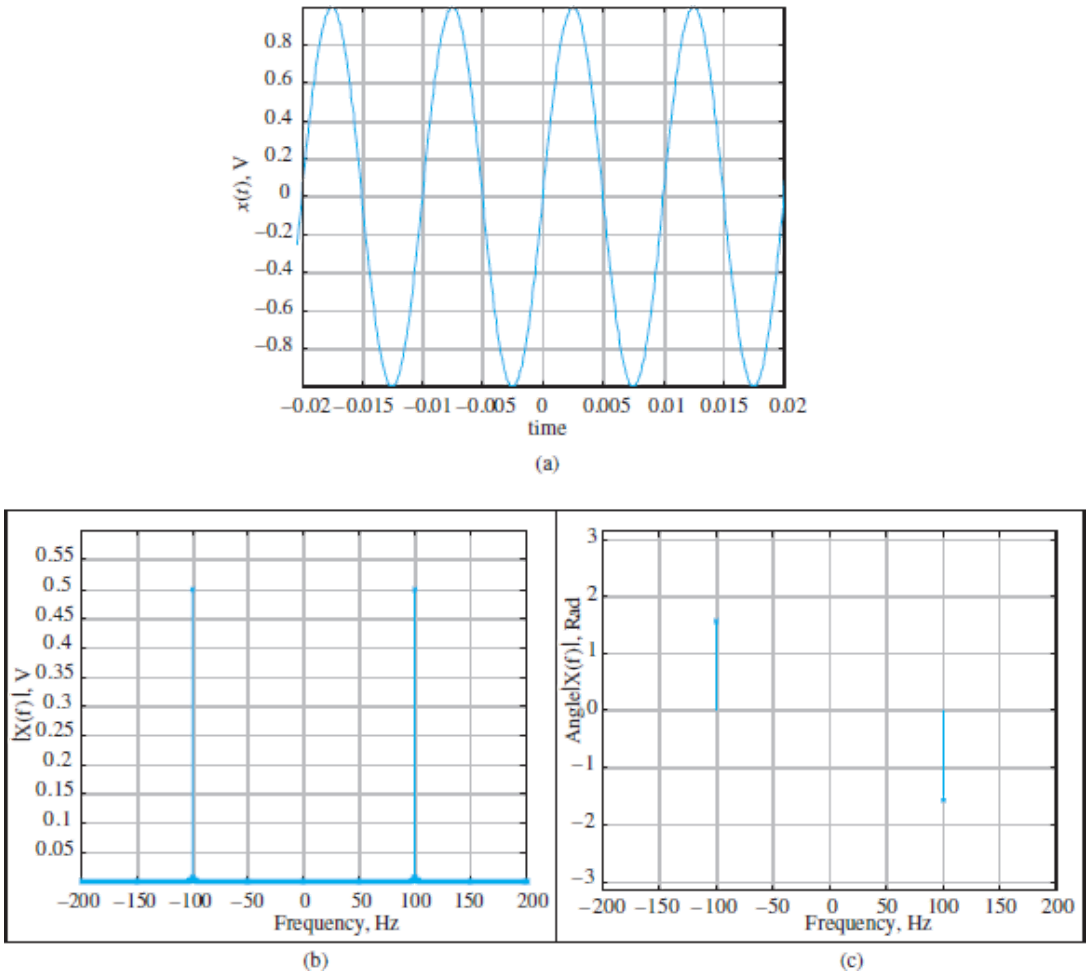


Figure 18.4 (a) Time-domain and (b, c) frequency-domain [(b) magnitude, (c) phase] representation of a sinusoidal voltage (amplitude: 1 V-peak; phase: 0 rad)

Periodic Signals: Fourier Series

Periodic signals can be represented as an infinite summation of sinusoids, as explained in Section 5.2. Each sinusoid in the summation oscillates at its own unique frequency, each of which is determined by the characteristics of the original signal. These sinusoidal components of the original signal are known as its *harmonics*.

Nonperiodic Signals: Fourier Transform

Practical communication signals have both periodic and nonperiodic components. Typically, the carrier waveform is periodic (usually a sine wave), while the modulating signal, consisting of speech, music, video, or data, is nonperiodic. The analysis of nonperiodic signals uses a mathematical tool different from (but related to) the Fourier series: the **Fourier transform**. The Fourier transform, also named after the French mathematician Jean-Baptiste Joseph Fourier, is an integral transform, so called because it is mathematically represented by an integral and because it performs a *transformation* between two domains: the **time domain** and the **frequency**, or **spectral**, **domain**.

Page 18-7

The Fourier transform of a function $x(t)$ is the function $X(\omega)$ defined by the integral

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt \qquad \text{Fourier transform} \tag{18.1}$$

Conversely, if the function $X(\omega)$ is known, the **inverse Fourier transform** is defined by:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega \qquad \text{Inverse Fourier transform} \tag{18.2}$$

Page 18-8

The pair $x(t)$ and $X(\omega)$ represent a Fourier transform pair, a relationship usually denoted by $x(t) \leftrightarrow X(\omega)$. Table 18.3 provides a summary of Fourier transform pairs that are often useful in computations.

Table 18.3 Properties of Fourier transforms

Property	Signal	Fourier transform
Time shift	$x(t - t_0)$	$e^{-j\omega t_0}X(\omega)$
Frequency shift	$e^{j\omega_0 t}x(t)$	$X(\omega - \omega_0)$
Complex conjugate	$x^*(t)$	$X^*(-\omega)$
Time reflection	$x(-t)$	$X(-\omega)$
Scaling	$x(at)$	$\frac{1}{a}X\left(\frac{\omega}{a}\right)$
Convolution	$x(t)^* y(t)$	$X(\omega) \cdot Y(\omega)$
Multiplication	$x(t) \cdot y(t)$	$\frac{1}{2\pi}X(\omega) * Y(\omega)$
Differentiation	$\frac{d}{dt}x(t)$	$j\omega X(\omega)$
Integration	$\int_{-\infty}^{t} x(t)\,dt$	$\frac{1}{j\omega}X(\omega) + \pi X(0)\,\delta(\omega)$

The **unit impulse** or **delta function**, $\delta(t)$, which is the derivative of the unit step function, plays an important role in Fourier transform analysis. (SeeA ppendix B, equation B.13.)

$$\delta(t) = \frac{du(t)}{dt} \quad \text{or} \quad u(t) = \int_{-\infty}^{t} \delta(t')dt' \quad \text{Delta or unit impulse function} \tag{18.3}$$

The unit impulse function has three important properties: (1) its area is equal to one; (2) it has infinite amplitude; and (3) it has zero duration, that is, its occurrence is concentrated at one instant in time. Clearly, such a signal is a mathematical abstraction since it is impossible to physically generate a signal that has zero duration and infinite amplitude. Figure 18.5 shows how one can think of the delta function as the limit of a sequence of rectangular pulses that are increasingly narrow and tall, such that the product of height ($1/\varepsilon$) and width (ε) is always equal to 1: the delta function can be thought of as the limit of this sequence as ε approaches zero.

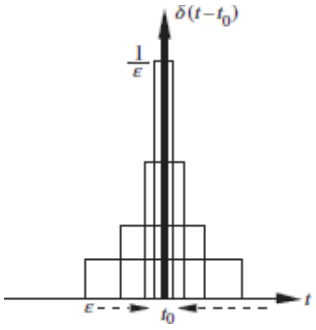


Figure 18.5 Delta function as the limit of a sequence of rectangular pulses of unit area

The delta function has one further property that is of interest in signal analysis:

$$\int_{-\infty}^{\infty} x(t)\,\delta(t-t_0)\,dt = x(t_0) \tag{18.4}$$

that is, the delta function "samples" the function $x(t)$ at the time of the occurrence of the impulse. Table 18.4 lists the Fourier transform of $\delta(t)$ along with that of other common functions.

Table 18.4 Fourier transform pairs

	$x(t)$	$X(\omega)$
1	$\delta(t)$	1
2	1	$2\pi\delta(\omega)$
3	$u(t)$ (unit step)	$2\pi\delta(\omega) + \frac{1}{j\omega}$
4	$tu(t)$ (unit ramp)	$2\pi\frac{d\delta(\omega)}{d\omega} - \frac{1}{\omega^2}$
5	$e^{-\alpha t}u(t) \quad \alpha > 0$	$\frac{1}{\alpha + j\omega}$
6	$te^{-\alpha t}u(t) \quad \alpha > 0$	$\frac{1}{(\alpha + j\omega)^2}$
7	$e^{j\omega_0 t}$	$2\pi\delta(\omega - \omega_0)$
8	$\cos(\omega_0 t)$	$\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$
9	$\sin(\omega_0 t)$	$-j\pi[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]$
10	$\cos(\omega_0 t)\, u(t)$	$\frac{\pi}{2}[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] + \frac{j\omega}{\omega_0^2 - \omega^2}$
11	$\sin(\omega_0 t)\, u(t)$	$-j\frac{\pi}{2}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] + \frac{\omega_0}{\omega_0^2 - \omega^2}$

The importance of the Fourier transform is that it allows us to view signals in the frequency or spectral domain. The spectral representation of signals is much Page 18-9more convenient and effective in representing communication signals, among other reasons because it allows important concepts, such as *bandwidth* and *spectrum allocation*, to be defined. A sinusoidal signal is represented by a single frequency. In the following examples, the Fourier transforms of a sinusoid, a single rectangular pulse, and a sine wave burst, or RF pulse, are computed. These three signals are frequently present in communication systems.

Bandwidth

The **bandwidth** of a signal is the range of frequencies comprising the spectrum of the signal. Bandwidth is a very important concept in communication systems, as the allocation of the radio frequency spectrum for different communication systems permits the transmission of a signal within a certain specified bandwidth. For example, standard FM radio allows a bandwidth of 200 kHz for each radio station. The most common definition of bandwidth is that of **3-dB bandwidth**, also called **half-power bandwidth**. The 3-dB bandwidth of a signal is defined as the frequency range between points where the signal level is 3 dB below its maximum passband value. This informal definition is illustrated in Figure 18.6, where an arbitrary voltage signal is shown to have a spectrum $V(f)$, with *center frequency* f_0 and 3-dB bandwidth $2B$.

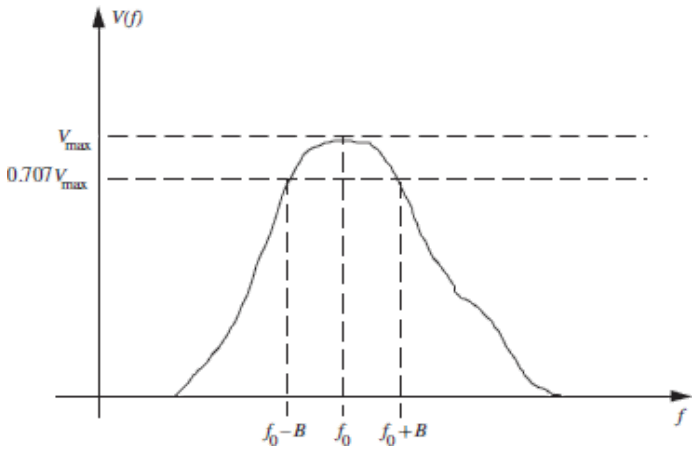


Figure 18.6 Definition of 3-dB (half-power) bandwidth

You will recall from the definitions given in Chapter 5 that the 3-dB point in a frequency plot is the frequency where the amplitude has dropped to a value equal to $1/\sqrt{2}$, or 0.707, times the maximum value. Since signal power is proportional to the square of the voltage, the 3-dB bandwidth is also called the half-power bandwidth. Thus, half of the signal power is contained in the frequency band $f_0 - B$ to $f_0 + B$; we call $2B$ the bandwidth of the signal. Please observe that this informal definition assumes that the signal spectrum has a bandpass shape. This is usually the case for most, if not all, communication signals.

Page 18-10

How much bandwidth does a signal require? This depends on two factors: (1) the bandwidth of the signal itself and (2) the type of modulation. We shall revisit the concept of bandwidth when we explore amplitude- and frequency-modulation systems.

EXAMPLE 18.1 Sinusoidal Signal Spectrum

Problem

Generate the spectrum of a signal consisting of the addition of two unity-amplitude sine waves for different frequencies and phases. Plot the time-domain

sum and the spectrum of the sum.

Solution

Known Quantities: Sine wave amplitude, frequency, and phase.

Find: Plot the time-domain sum of the signals and the frequency-domain spectrum.

Schematics, Diagrams, Circuits, and Given Data: $\omega_1 = 300$ rad/s; $\omega_2 = 500$ rad/s; $\phi_1 = 0$ rad; $\phi_2 = \pi/4$ rad/s.

Assumptions: None.

Analysis: The time-domain signals $x_1(t)$ and $x_2(t)$ and their sum are shown in Figure 18.7(a). Figure 18.7(b) depicts the frequency spectrum of the sum signal.

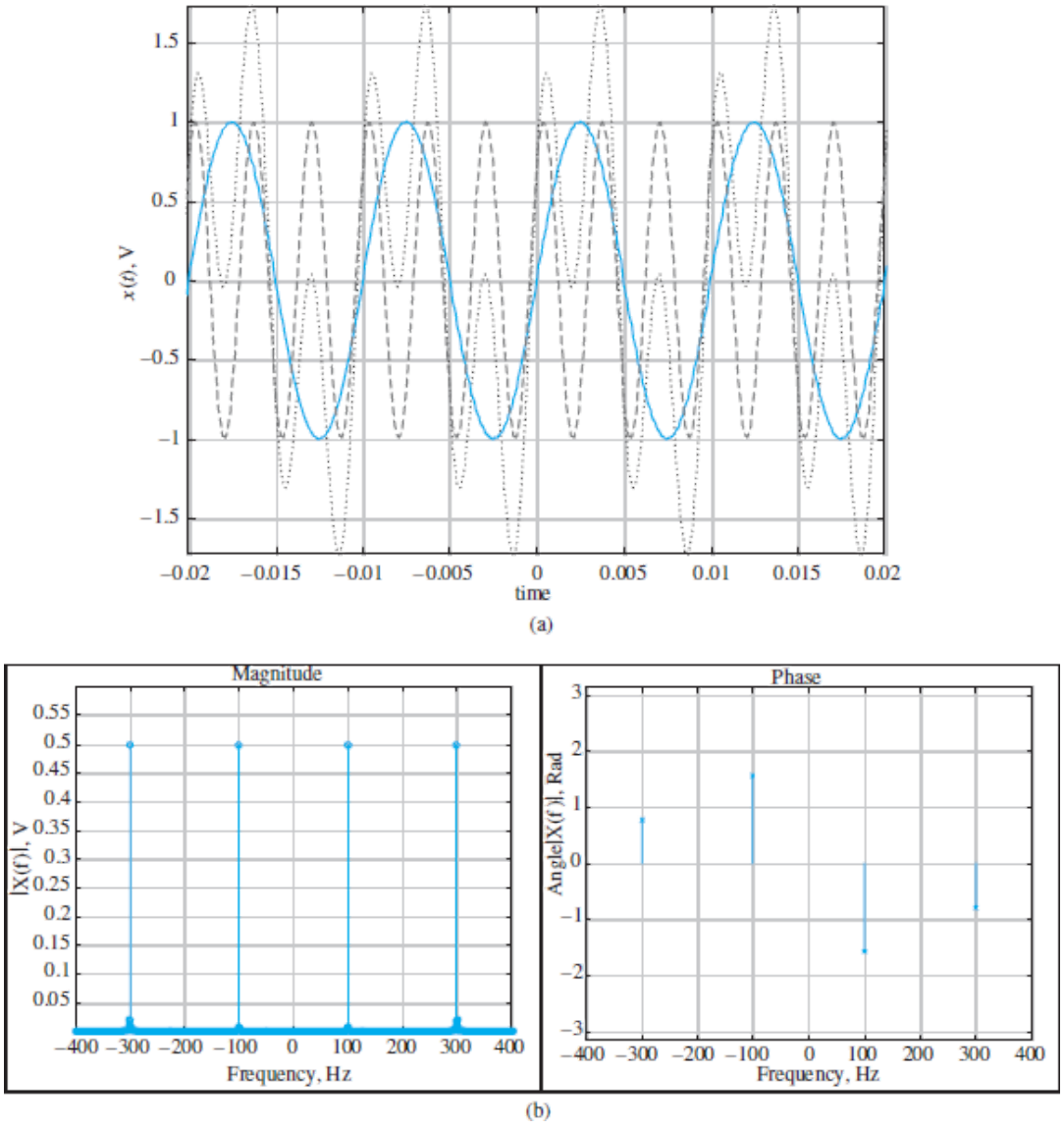


Figure 18.7 (a) Time-domain and (b) frequency-domain representation of the sum of two sinusoidal voltages

Comments: Note that the signal amplitude in the time domain is divided between two spectral lines at each signal component frequency and at the corresponding negative frequency; thus, signal power is preserved. The phase angle (at the positive frequencies) is shown to be $-\pi/2$ for the signal $x_1(t)$ because the signal is a sine wave (in Chapter 3 we defined the cosine as the reference function, with zero phase angle); thus, $x_2(t)$ has phase angle $-\pi/2 + \pi/4 = -\pi/4$.

EXAMPLE 18.2 Fourier Series of Pulse Train

Problem

Compute the complete Fourier series of the periodic pulse train shown in Figure 18.8.

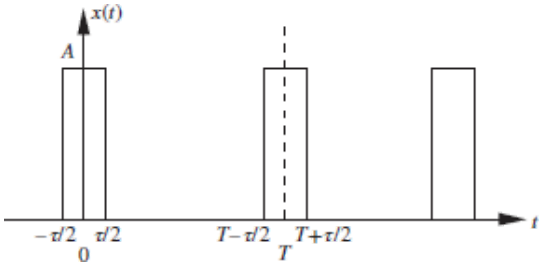


Figure 18.8 Periodic pulse train

Solution

Known Quantities: Amplitude, period, and functional form of the signal.

Find: b_n and c_n coefficients as a function of n.

Schematics, Diagrams, Circuits, and Given Data: $\delta = \frac{\tau}{T} = 0.2.$

Page 18-12

Assumptions: The function repeats periodically.

Analysis: The expression for $x(t)$ is simple:

$$x(t) = \begin{cases} A & -\frac{\tau}{2} < t < \frac{\tau}{2} \\ 0 & \frac{\tau}{2} < t < T - \frac{\tau}{2} \end{cases}$$

Evaluate the integral below to determine the complex coefficients of the Fourier series (see Chapter 5 for more on Fourier coefficients):

$$\gamma_n = \frac{1}{T}\int_{(-T/2)}^{(T/2)} x(t)\, e^{(-jn(2\pi/T)t}\, dt) = \frac{1}{T}\int_{(-\tau/2)}^{(\tau/2)} A\, e^{(-jn(2\pi/T)t)}\, dt$$

$$= \frac{A}{T}\frac{1}{(-jn(2\pi/T))}[e^{(-jn(2\pi/T)t)}]_{(-\tau/2)}^{(\tau/2)} = \frac{A}{\pi n}\left[\frac{e^{-jn(\pi\tau/T)} - e^{jn(\pi\tau/T)}}{-2j}\right] = \frac{A}{\pi n}\left[\frac{e^{jn\pi\delta} - e^{-jn\pi\delta}}{2j}\right]$$

$$= \frac{A}{\pi n}\sin(n\pi\,\delta) \qquad n = 0, \pm 1, \pm 2, \ldots$$

We may simplify the notation by using the *sinc function*, defined as:

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

Thus we may rewrite the coefficients as:

$$\gamma_n = \frac{A}{\pi n}\sin(n\pi\,\delta) = \delta\,\text{sinc}\,(n\delta)$$

Figure 18.9(a) depicts the pulse train corresponding to the numerical values given above, and Figure 18.9(b) its Fourier series coefficients up to n = 1,000. The *envelope* of the discrete-frequency coefficients is the sinc function defined above.

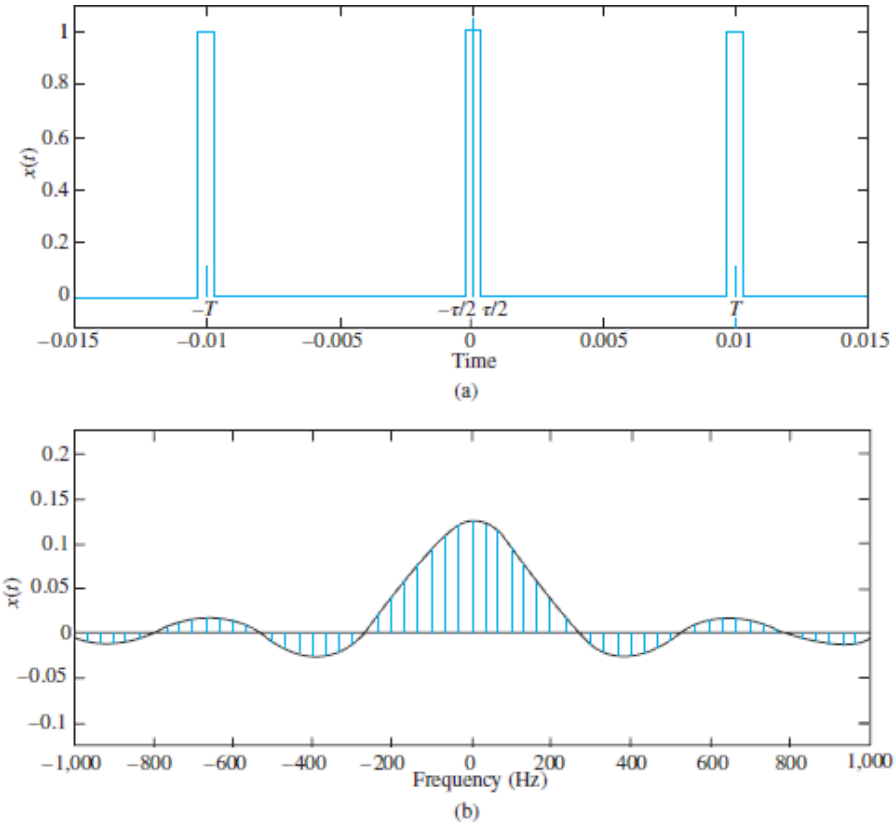


Figure 18.9 Time-domain representation and Fourier series spectrum of periodic pulse train with $\delta = 0.2$

Comments: In this example, the integral representation of the Fourier coefficients is easily evaluated because $x(t)$ is a simple constant. Fourier integrals with polynomial forms of $x(t)$ can be evaluated by repeated application of the method of integration by parts. More generally, Fourier integrals with periodic forms of $x(t)$ can be evaluated using Euler's theorem and expressions found in Chapter 5.

Note that in the complex form of the Fourier series, the coefficients range from negative infinity to positive infinity.

EXAMPLE 18.3 Fourier Transform of Sine Wave

Problem

Compute the Fourier transform of an arbitrary sinusoidal signal.

Solution

Known Quantities: Functional form of the signal $x(t)$.

Find: $X(\omega)$.

Schematics, Diagrams, Circuits, and Given Data: Figure 18.10.

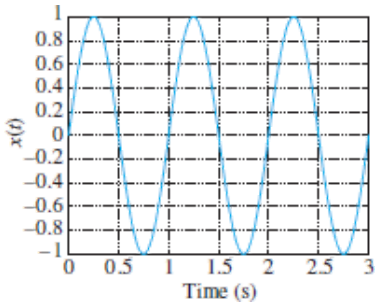


Figure 18.10 Sinusoidal signal of frequency $\omega_o = 2\pi$ Rad/s.

Assumptions: None

Analysis: The signal shown in Figure 18.10 is defined as

$$x(t) = \sin(\omega_0 t)$$

Page 18-14

The Fourier transform for the signal is calculated using the complex representation

$$\begin{aligned} \gamma_n &= \frac{1}{T}\int_{(-T/2)}^{(T/2)} x(t)\, e^{-jn(2\pi/T)t}\, dt \\ &= \frac{1}{T}\int_{(-T/2)}^{(T/2)} \sin(\omega_0 t)\, e^{-jn(2\pi/T)t}\, dt \\ &= \frac{1}{T2j}\int_{(-T/2)}^{(T/2)} \left[e^{(j2\pi/T)t} - e^{(-j2\pi/T)t}\right] e^{-jn(2\pi/T)t}\, dt \\ &= \frac{1}{2j}\frac{1}{T}\int_{(-T/2)}^{(T/2)} \left[e^{j(2\pi/T)(1-n)t} - e^{j(2\pi/T)(-n-1)t}\right] dt \end{aligned}$$

Using the relation from Table 18.4 for the defined Fourier transform pairs, the above integral can be evaluated as:

$$\begin{aligned} \gamma_n &= \frac{1}{2j}[2\pi\delta(\omega - \omega_0) - 2\pi\delta(\omega + \omega_0)] \\ &= \frac{\pi}{j}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] \end{aligned}$$

Thus, the Fourier transform of an arbitrary sinusoid consists of a pair of delta functions in the frequency domain, as shown in Figure 18.11.

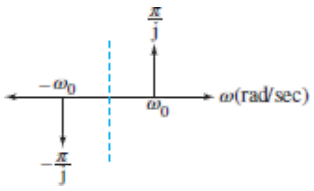


Figure 18.11 Fourier transform of sinusoid

Comments: We can extend the result of this example to an arbitrary periodic signal since we know that any periodic signal can be represented by the sum of an infinite number of sinusoidal functions. The Fourier transform $X(\omega)$ of a periodic

signal $x(t)$ is a train of impulses occurring at the harmonically related frequencies and for which the area of the impulse at the nth harmonic frequency $n\omega_0$ is 2π times the nth Fourier series coefficient a_n. The Fourier series coefficients for this sinusoid signal are

$$a_1 = \frac{1}{2j}$$
$$a_{-1} = -\frac{1}{2j}$$
$$a_k = 0 \qquad |k| \neq 1$$

Hence, the Fourier transform coefficients are given by

$$a_1 = \frac{\pi}{j}$$
$$a_{-1} = -\frac{\pi}{j}$$
$$a_k = 0 \qquad |k| \neq 1$$

EXAMPLE 18.4 Fourier Transform of Rectangular Pulse Signal

Problem

Compute the Fourier transform of a square pulse signal.

Page 18-15

Solution

Known Quantities: Functional form of the signal $x(t)$.

Find: $X(\omega)$.

Schematics, Diagrams, Circuits, and Given Data: Figure 18.12.

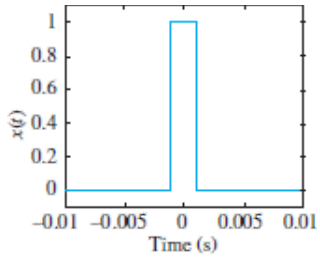


Figure 18.12 Rectangular pulse

Assumptions: None.

Analysis: Consider the rectangular pulse $x(t)$ of duration T and unity amplitude shown in Figure 18.12. We define this pulse mathematically as follows:

$$x(t) = \begin{cases} 1 & |t| \le \frac{T}{2} \\ 0 & |t| > \frac{T}{2} \end{cases}$$

The Fourier transform of $x(t)$ is

$$X(\omega) = \int_{(-T/2)}^{(T/2)} e^{-j\omega t}\,dt = \left[\frac{e^{-j\omega t}}{-j\omega}\right]_{-(T/2)}^{(T/2)} = \frac{2}{\omega}\left[\frac{e^{j\omega(T/2)} - e^{-j\omega(T/2)}}{2j}\right] = \frac{2}{\omega}\sin\left(\omega\frac{T}{2}\right)$$

$$= T\,\text{sinc}\left(\omega\frac{T}{2}\right) \quad \text{where} \quad \text{sinc}(x) = \frac{\sin(x)}{x}$$

$$X(f) = T\,\text{sinc}(\pi f T)$$

A plot of the spectrum is shown in Figure 18.13. The figure illustrates the characteristics of the sinc function, with zero crossings at integer multiples of $2\pi/T$ rad/s or $1/T$ Hz, and peak amplitude of T.

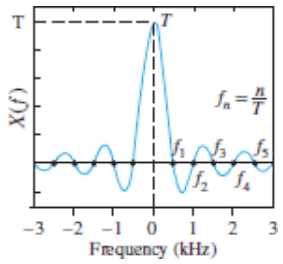


Figure 18.13 Fourier transform of square pulse (magnitude only)

Comments: Single and repetitive square bursts occur commonly in communication systems. The analysis completed in this example will be useful in the following sections.

EXAMPLE 18.5 Fourier Transform of Sine Burst (RF Pulse)

Problem

Compute the Fourier transform of the sine wave burst shown in Figure 18.14.

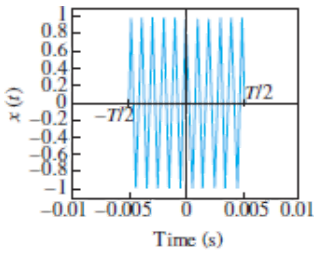


Figure 18.14 Radio-frequency (RF) burst

Solution

Known Quantities: Functional form of the signal $x(t)$.

Find: $X(\omega)$.

Schematics, Diagrams, Circuits, and Given Data: Figure 18.14.

Assumptions: None.

Analysis: The pulse signal $x(t)$ shown in Figure 18.14 consists of a sinusoidal wave of unit amplitude and frequency f_c, for a duration $t = -T/2$ to $t = T/2$. The signal $x(t)$ can be defined mathematically as follows:

$$x(t) = \begin{cases} A\cos(2\pi f_c t) & |t| \le \dfrac{T}{2} \\ 0 & |t| > \dfrac{T}{2} \end{cases}$$

The Fourier transform of $x(t)$ is

$$X(\omega) = \int_{(-T/2)}^{(T/2)} x(t)\, e^{-j\omega t}\, dt = \int_{(-T/2)}^{(T/2)} \frac{A}{2}\left(e^{j\omega_c t} + e^{-j\omega_c t}\right) e^{-j\omega t} dt$$

$$= \frac{A}{2}\left[\frac{e^{-j(\omega-\omega_c)t}}{-j(\omega-\omega_c)} + \frac{e^{-j(\omega+\omega_c)t}}{-j(\omega+\omega_c)}\right]_{(-T/2)}^{(T/2)}$$

$$= A\left[-\frac{e^{-j(\omega-\omega_c)(T/2)}}{2j(\omega-\omega_c)} - \frac{e^{-j(\omega+\omega_c)(T/2)}}{2j(\omega+\omega_c)} + \frac{e^{j(\omega-\omega_c)(T/2)}}{2j(\omega-\omega_c)} + \frac{e^{j(\omega+\omega_c)(T/2)}}{2j(\omega+\omega_c)}\right]$$

$$= A\left[\frac{\sin\pi(f+f_c)T}{2\pi(f+f_c)} + \frac{\sin\pi(f-f_c)T}{2\pi(f-f_c)}\right]$$

$$= \frac{AT}{2}[\text{sinc}(f+f_c)T + \text{sinc}(f-f_c)T]$$

Note that we could have used the frequency-shifting property of the Fourier transform to obtain the above result without carrying out the integration explicitly. The magnitude spectrum of the RF pulse is shown in Figure 18.15. It clearly illustrates the *frequency-shifting property* of the Fourier transform.

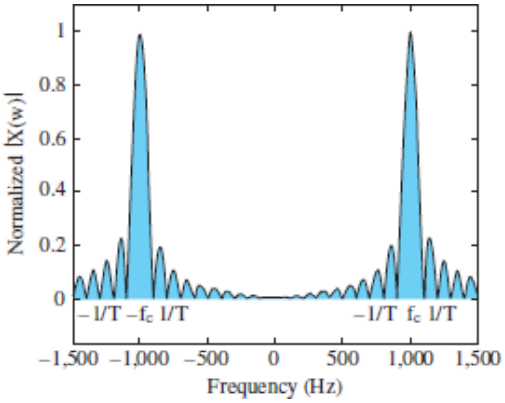


Figure 18.15 Magnitude spectrum of RF burst

Comments: This signal is specifically referred to as a *RF pulse* when the frequency f_c of the sinusoid wave falls in the radio frequency range.

EXAMPLE 18.6 Bandwidth of Commercial AM (or TV, or FM) Signals

Problem

Analyze the bandwidth of the signal from a commercial AM station and determine how many stations can be assigned frequencies over the frequency band assigned to commercial AM.

Solution

Schematics, Diagrams, Circuits, and Given Data: See www.ntia.doc.gov/osmhome/allochrt.pdf.

Page 18-17

Analysis: As can be seen in the diagram displayed in the given website, the AM band frequency allocation goes from 535 to 1,605 kHz. Each channel is allocated a bandwidth of approximately 10 kHz (we shall see in the next section what this calculation includes). Thus, the total number of stations that can operate in the same region is approximately equal to:

$$N = \frac{1{,}605 - 535}{10} = 107$$

Each AM station can operate with a total bandwidth of 10 kHz. As we shall see in the next section, this actually corresponds to a signal bandwidth of only 5 kHz.

Comments: You are probably aware of the fact that the FCC licenses many more than 107 AM stations in the United States. This is possible because AM broadcast has a limited range, and two stations can be assigned the same frequency if they are located sufficiently far apart. You may also have noticed that at night it is possible to receive AM radio signals from much greater distances (for example, in Ohio one can tune in to stations from as far as New York City and New Orleans late at night). This is a consequence of the change in ionization density in the ionosphere during the night, permitting reflection of radio waves over a longer range. The FCC regulates not only the frequency allocated, but also the power allocated to a given station; a station may be required to switch to a lower-power transmitter during certain times of the day.

CHECK YOUR UNDERSTANDING

a. Compute the coefficients of the complete Fourier series expansion for the signal $x(t) = 1.5\cos(100t)$.

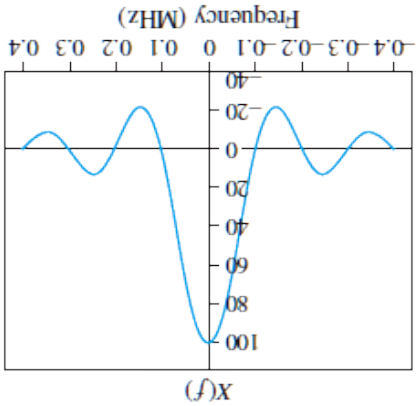
b. Sketch the Fourier transform of a square pulse with unity amplitude and with a duration of 10 μs.

c. The spectrum of a signal can be described by the function $X(\omega) = (\alpha\omega/\omega^2 + \alpha^2)$. Let $\alpha = 10^3$, and calculate the 3-dB bandwidth of the signal. What is the center frequency of the signal?

Answers:

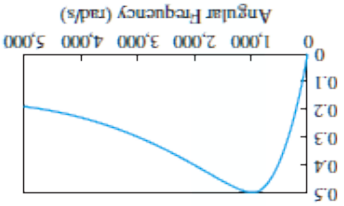
a. Integration by parts solves the coefficient integral to yield: $\gamma_n = \frac{0.75n}{j\pi}\frac{1 - e^{-j2\pi n}}{n^2 - 1}$

b. Square-pulse Fourier transform:



Answers:

c. Center frequency: 1000 rad/s, 3-dB bandwidth: 2,000 rad/s



18.3 AMPLITUDE MODULATION AND DEMODULATION

The concept of **amplitude modulation** (AM) was introduced in Chapter 3 (Focus on Measurements box, "Capacitive Displacement Transducer and Microphone"), where it was shown that the signal produced by a capacitive microphone (displacement transducer) inserted in a Wheatstone bridge circuit modulated the amplitude of a sinusoidal excitation signal. In that example, the pressure changes sensed by the microphone constituted the *modulation*, while the sinusoidal excitation provided a *carrier*. In Chapter 8 (Focus on Measurements box, "Peak Detector for Capacitive Displacement Transducer"), it was shown that a diode circuit was capable of demodulating the AM signal and of recovering the desired information (pressure changes corresponding to acoustic waves, that is, sound). In this section, the same basic principles introduced in the above mentioned examples will be discussed more formally, as they apply to AM communication systems.

The most common manifestation of amplitude modulation in communication systems is commercial AM radio, or *standard AM*. The *Federal Communications Commission* (FCC), a body that regulates the usage and allocation of the radio frequency spectrum in the United States has assigned the frequency band between 540 and 1,600 kHz to commercial AM radio transmission. Each station can occupy a bandwidth of 10 kHz centered around its carrier. As we shall see, this corresponds to an effective signal bandwidth of 5 kHz—sufficient for good reproduction of speech and acceptable reproduction of music.

Basic Principle of Amplitude Modulation

AM signals are generated by modulating the amplitude of a carrier signal. Let the carrier signal be a sinusoid at frequency ω_c:

$$c(t) = A_c \cos(\omega_c t) \qquad \textbf{Carrier signal} \tag{18.5}$$

and, for illustration purposes, let the modulation also be a single tone (sinusoid), at a frequency $\omega_m \ll \omega_c$:

$$m(t) = A_m \cos(\omega_m t) \qquad \textbf{Modulating signal} \tag{18.6}$$

With these definitions, we can define the basic AM signal as follows:

$$s(t) = [A_c + m(t)]\cos(\omega_c t) \tag{18.7}$$

Page 18-19

or

$$s(t) = A_c\left[1 + \frac{A_m}{A_c}\cos(\omega_m t)\right]\cos(\omega_c t) \qquad \textbf{AM signal} \tag{18.8}$$

The **modulation index** μ is defined to be the ratio of the modulation to carrier signal amplitudes:

$$\mu = \frac{A_m}{A_c} \qquad \textbf{Modulation index} \tag{18.9}$$

and for proper amplitude modulation should be less than 1. If equation 18.8 is expanded, we see that an AM signal is composed of two terms: a sinusoidal carrier wave, plus a wave that is the product of two sinusoidal terms. Using trigonometric identities, we can write the following expression:

$$\begin{aligned} s(t) &= A_c\cos(\omega_c t) + \mu A_c\cos(\omega_c t)\cos(\omega_m t) \\ &= A_c\cos(\omega_c t) + \mu\frac{A_c}{2}\cos(\omega_c - \omega_m)t + \mu\frac{A_c}{2}\cos(\omega_c + \omega_m)t \end{aligned} \tag{18.10}$$

Equation 18.10 shows that the AM signal is really composed of three sinusoidal waveforms: (1) a carrier wave; (2) a **lower sideband** signal, at frequency ($\omega_c - \omega_m$), containing the modulating signal; and (3) an **upper sideband** signal, at frequency ($\omega_c + \omega_m$), also containing information (the modulation). Example 18.7 illustrates some important properties of an AM signal with pure sinusoidal modulation. An understanding of single-tone modulation, as in Example 18.7, is essential to understanding double-tone and nonperiodic modulation found in more realistic AM signals.

AM Demodulation; Integrated Circuit Receivers

Demodulation is the process of recovering the modulating signal from a received modulated signal. With reference to Figure 18.1, one can think of the transmitter in an AM signal as the device that imposes the modulation on a carrier, while the receiver extracts the modulating signal from a received AM signal. To understand the basic principle of modulation and demodulation, we observe that amplitude modulation consists in effect of *multiplying* the carrier signal times the

modulating signal. This process is often called *mixing*, and a **mixer** is the device that implements this function, that is, multiplication.

Consider the AM signal of equation 18.8. Multiply it by a second signal at the same frequency as the carrier signal:

$$2 \cdot s(t) \cdot c(t) = 2A_c[A_c + \mu \cos(\omega_m t)]\cos(\omega_c t) \cdot A_c \cos(\omega_c t) \tag{18.11}$$

Multiply the two $\cos(\omega_c t)$ terms and use the trigonometric identity $\cos^2(\omega_c t) = [1 + \cos(2\omega_c t)]/2$ to yield:

$$\begin{aligned} 2 \cdot s(t) \cdot c(t) &= A_c^2[1 + \mu \cos(\omega_m t)]\,[1 + \cos(2\omega_c t)] \\ &= A_c[A_c + m(t)] + A_c[A_c + m(t)]\cos(2\omega_c t) \end{aligned} \tag{18.12}$$

In the case that $A_c = 1$:

$$2 \cdot s(t) \cdot c(t) = [1 + m(t)] + [1 + m(t)]\cos(2\omega_c t)$$

The result of the mixing operation consists of two terms: a constant plus the original modulation signal $m(t)$—what we desire to recover—and an amplitude-modulated term at twice the carrier frequency. Since the baseband of the modulation signal is typically (for example, 0 to 5 kHz for speech and music) much lower than the carrier band, the modulating signal can be recovered by low-pass filtering the output of the mixer. Figure 18.16 depicts a block diagram of a conceptual AM demodulator as well as the spectra of the AM signal before and after mixing.

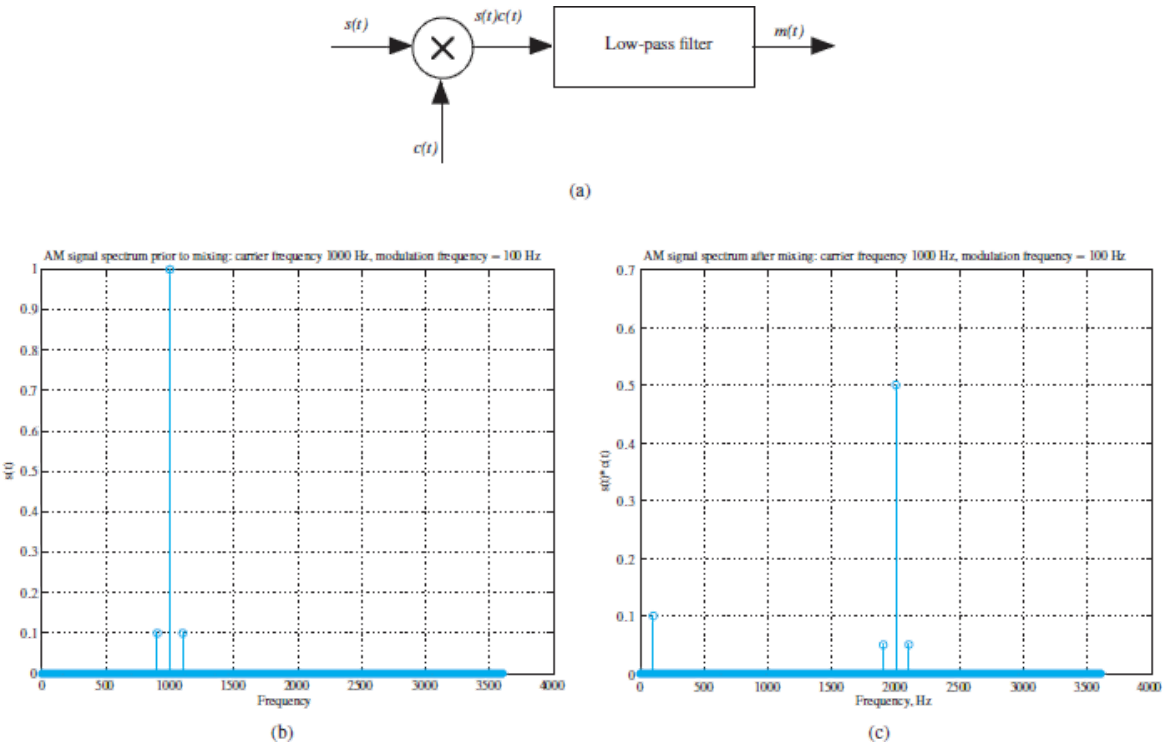


Figure 18.16 (a) Block diagram of AM demodulator; (b) spectrum of AM signal; (c) spectrum of signal following the mixer (multiplier)—note that the modulation signal (at 100 Hz) can now be recovered via low-pass filtering.

The process of demodulating AM signals is carried out today by means of **integrated circuit receivers**.

A sample MatLab® code for calculating AM signals and carrying out the mixing operation is listed below.

```
%AM sample code
t = 0.05/360:0.05/360:1;
Fs = 7200; %360/0.05 data points
Ac = 1; % Carrier amplitude
fc = 1000; % Carrier signal cyclical frequency
wc = 2.0*pi*fc; % Carrier angular frequency
```

```
c = Ac*cos(wc*t); % Carrier signal
figure(1)
plot(c)
Am = 0.2; % Modulating signal amplitude
fm = 100; % Modulating signal cyclical frequency
wm = 2.0*pi*fm; % Modulating signal angular frequency
m = Am*cos(wm*t); % Modulating signal
figure(2)
plot(m)
u = Am/Ac; % Modulation index
s = Ac*(1+u*cos(wm*t)).*cos(wc*t); % Modulated signal (equation 18.8)
figure(3)
plot(s)
mm = 2.0*s.*c; % Mixed signal (equation 18.11 or 18.12)
figure(4)
plot(mm)
Ss = abs(fft(s-mean(s)))/3600;
% Discrete Fourier transform of normalized AM signal
Smm = abs(fft(mm-mean(mm)))/3600;
figure(5)
stem(Ss(1:3600))
xlabel('Frequency, Hz')
ylabel('s(t)')
title('AM signal spectrum prior to mixing: carrier frequency 1,000 Hz,
  modulation frequency = 100 Hz')
grid
figure(6)
stem(Smm(1:3600))
xlabel('Frequency, Hz')
ylabel('s(t)*c(t)')
title('AM signal spectrum after mixing: carrier frequency 1,000 Hz,
  modulation frequency = 100 Hz')
grid
```

EXAMPLE 18.7 Single-Tone Amplitude Modulation

Problem

Analyze the spectrum of a single-tone modulation signal based on the WOSU AM 820 radio station. Use both analytical and computational tools.

Solution

Known Quantities: Carrier frequency; modulation index.

Find: Derive expressions for and plot time- and frequency-domain waveforms of the AM signal.

Schematics, Diagrams, Circuits, and Given Data: $f_c = 0.82$ MHz; $\mu = 0.5$.

Page 18-22

Assumptions: Assume unity amplitude for the carrier, $A_c = 1$ and a modulating frequency $f_m = 10$ kHz.

Analysis: Define the modulating signal $m(t)$ and the carrier signal $c(t)$ as

$$m(t) = A_m \cos(2\pi f_m t)$$
$$c(t) = A_c \cos(2\pi f_c t)$$

These waveforms are plotted in Figure 18.17(a) and (b), respectively. The spectra of these signals are plotted in Figure 18.17(d) and (e).

The AM wave $s(t)$ is given by

$$s(t) = A_c[1 + \mu \cos(2\pi f_m t)]\cos(2\pi f_c t)$$

and is plotted in Figure 18.17(c). Using the Fourier transform pairs given in Table 18.4 (pair 8), the Fourier transform of $s(t)$ can be expressed as the sum of three delta functions, centered at the carrier and at the sum (upper sideband) and difference (lower sideband) frequencies. Note also that the spectrum is repeated for negative frequencies, as explained earlier.

$$S(f) = \frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] + \mu\frac{A_c}{4}[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)]$$
$$+\mu\frac{A_c}{4}[\delta(f - f_c + f_m) + \delta(f + f_c - f_m)]$$

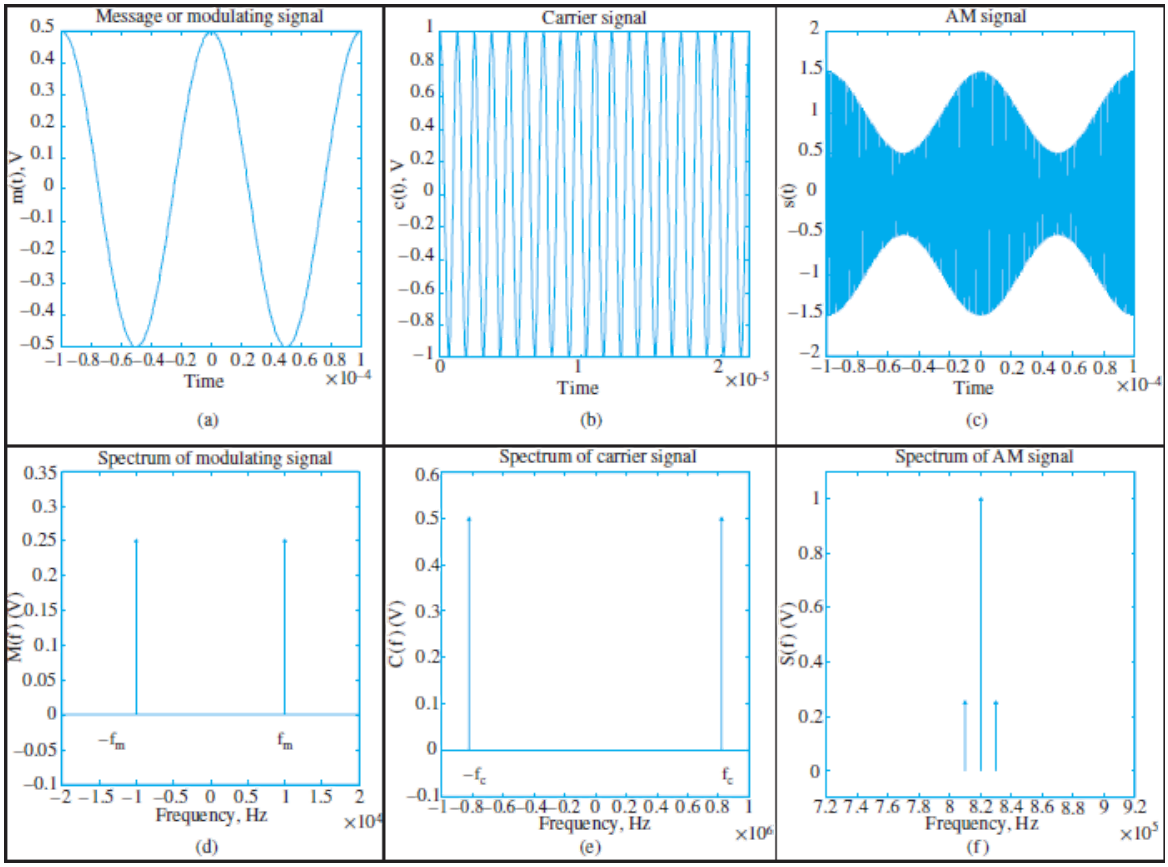


Figure 18.17 Time- and frequency-domain waveforms of single-tone AM signal

Page 18-23

Thus, the spectrum of an AM wave for the special case of sinusoidal modulation consists of the delta function at $\pm f_c$, $f_c \pm f_m$, and $-f_c \pm f_m$, as shown in Figure 18.16(f).

Comments: If you like, you may experiment with the value of the modulation index and see its effect on the AM wave. It is recommended that the modulation index μ be nearly equal to 1, but not greater. If the modulation index is greater than 1 for any t, the carrier wave becomes over-modulated, resulting in carrier phase reversals whenever the function $[1 + \mu m(t)]$ crosses zero.

EXAMPLE 18.8 Double-Tone Modulation

Problem

Plot the frequency spectrum of a carrier signal with unity amplitude and frequency f_c = 1 MHz, which is amplitude modulated with a modulating signal $m(t)$ consisting of two sinusoidal frequencies.

Page 18-24

Solution

Known Quantities: Carrier frequency and amplitude; modulating signal.

Find: Modulation index and frequency spectrum of the AM wave with the defined carrier and modulating signal.

Schematics, Diagrams, Circuits, and Given Data: f_c = 1 MHz; A_c = 1 V; $m(t)$ = 0.5 cos(2π 10,000t) + 0.4 cos(2π80,000t).

Assumptions: None.

Analysis: The modulation index for the signal is defined as

$$\mu = \frac{A_m}{A_c} = \frac{\max[m(t)] - \min[m(t)]}{2V_c} = \frac{0.9 + 0.8626}{2} = 0.8813$$

The spectrum of the AM wave in this case consists of delta functions at $\pm f_c, f_c \pm f_{m1}$, $f_c \pm f_{m2}$, $-f_c \pm f_{m1}$, and $-f_c \pm f_{m2}$, where f_{m1} and f_{m2} are the frequencies contained in the modulating signal. This is seen in Figure 18.18, where all time- and frequency-domain waveforms are plotted.

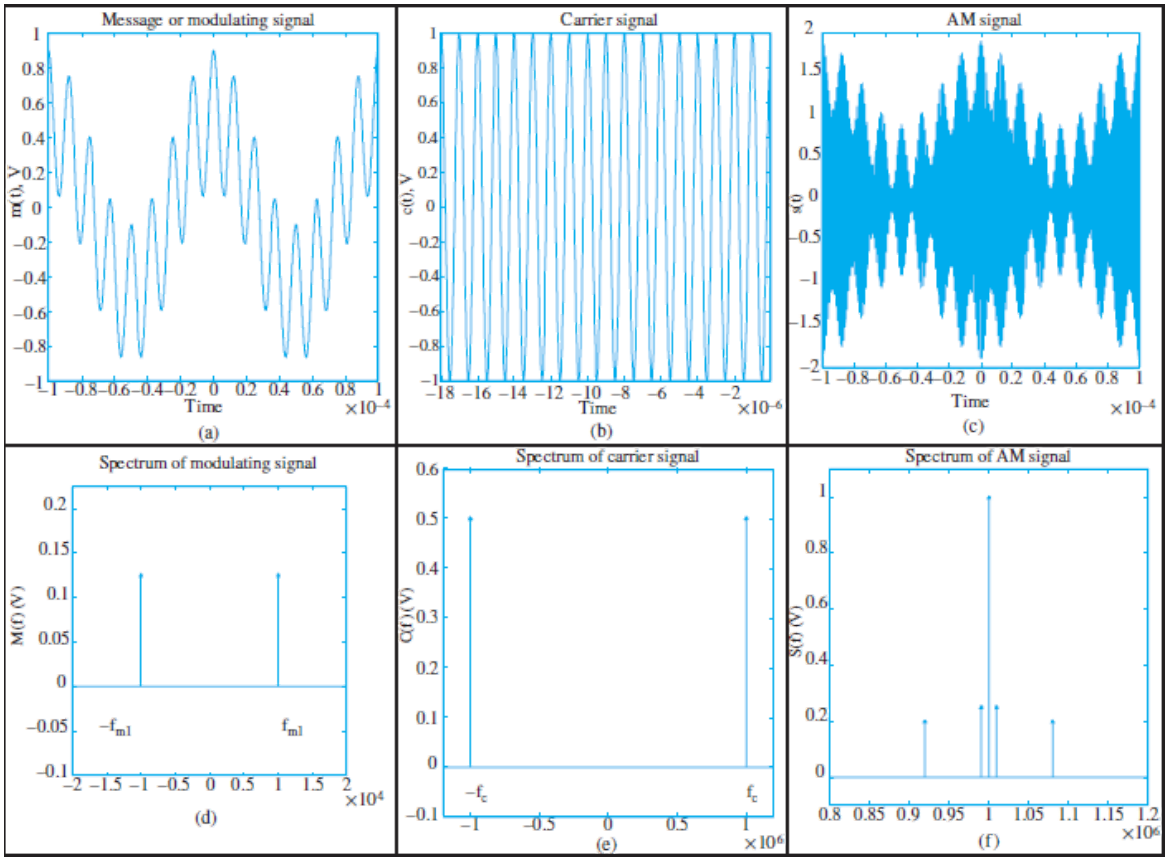


Figure 18.18 Time- and frequency-domain waveforms of double-tone AM signal

Comments: The frequency spectrum of the AM wave is just a shifted version of the original modulating signal with the shift in frequency equal to the carrier frequency. The portion of the spectrum of an AM wave lying above the carrier frequency f_c is the *upper sideband* whereas the symmetric portion below f_c is called the *lower sideband*.

EXAMPLE 18.9 Nonperiodic Amplitude Modulation

Problem

Plot the frequency spectrum of a carrier signal with unity amplitude and frequency f_c = 0.1 MHz, which is amplitude modulated with a non-periodic modulating signal $m(t)$ having a defined shape.

Solution

Known Quantities: Carrier frequency and amplitude; modulating wave $m(t)$ defined for a certain interval of time.

Find: Frequency spectrum of the AM wave.

Schematics, Diagrams, Circuits, and Given Data:

$$m(t - T/2) = \begin{cases} \sin(2\pi f_1 t) + \sin(2\pi f_2 t) + \sin(2\pi f_3 t) + \sin(2\pi f_4 t) + u(t) & 0 \le t \le T \\ 0 & \text{Otherwise} \end{cases}$$

$$u(t) = 1$$

$$c(t) = A_c \cos(2\pi f_c t)$$

$f_c = 0.1$ MHz; $A_c = 1$; $f_1 = 1$ kHz, $f_2 = 2$ kHz, $f_3 = 3$ kHz, $f_4 = 4$ kHz, $T = 0.5$ ms.

Assumptions: None.

Analysis: The signal waveform and the frequency spectrum of the modulating signal and of the AM wave are shown in Figure 18.19(a) to (d). The spectrum of the AM wave is a shifted version of the modulating signal spectrum around the carrier frequency.

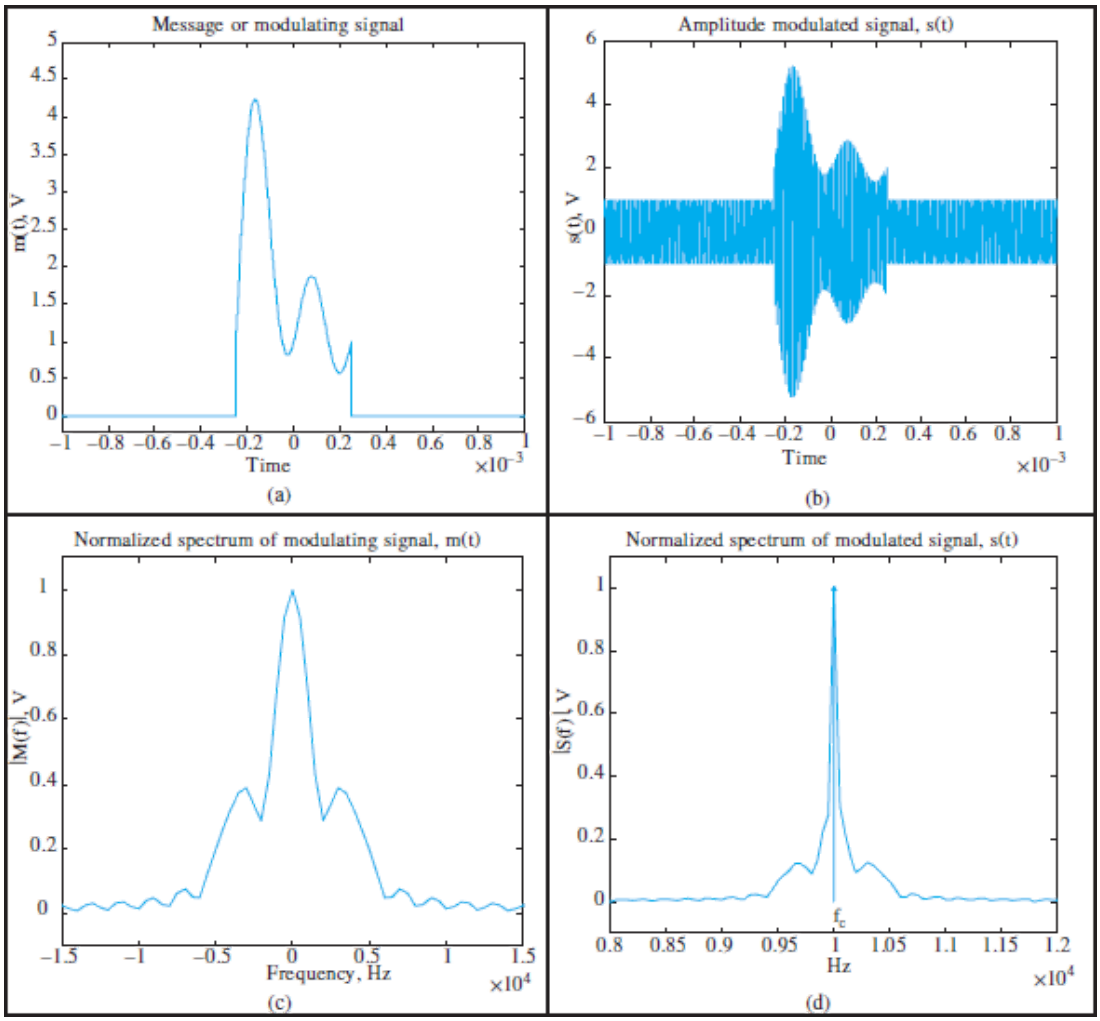


Figure 18.19 Time- and frequency-domain waveforms of nonperiodic AM signal

Comments: If B_{max} is the bandwidth of the modulating signal (the highest frequency in the modulating signal), the bandwidth of the AM wave is defined as *twice the highest frequency in the modulating signal*, i.e., $B = 2B_{max}$.

CHECK YOUR UNDERSTANDING

a. Use the Matlab® files that accompany Examples 18.7 and 18.8 to plot only the positive spectrum of the single- and double-tone AM signals. Determine the bandwidth of the AM signal in each case.

b. Determine the bandwidth of the modulating signal in Figure 18.19(c); what is the bandwidth of the AM signal? Is this consistent with commercial AM practice? Would the FCC allow a commercial station to broadcast such a signal?

Page 18-26

18.4 FREQUENCY MODULATION AND DEMODULATION

You are certainly familiar with the term *frequency modulation* (FM) because of the great diffusion of FM radio. As its name implies, frequency modulation refers to encoding the information contained in a modulating signal in the frequency of a carrier signal. Figure 18.20 depicts a sinusoidally modulated FM waveform and its corresponding magnitude spectrum. FM transmission permits significant improvements over AM, but at the cost of an increased requirement for bandwidth. In this section you will be introduced to the basic signal models for FM. Two different cases are discussed: **narrowband FM** and **wideband FM**. The plots of Figure 18.20 correspond to a wideband FM signal. Note the significant spread of signal frequencies relative to the carrier frequency!

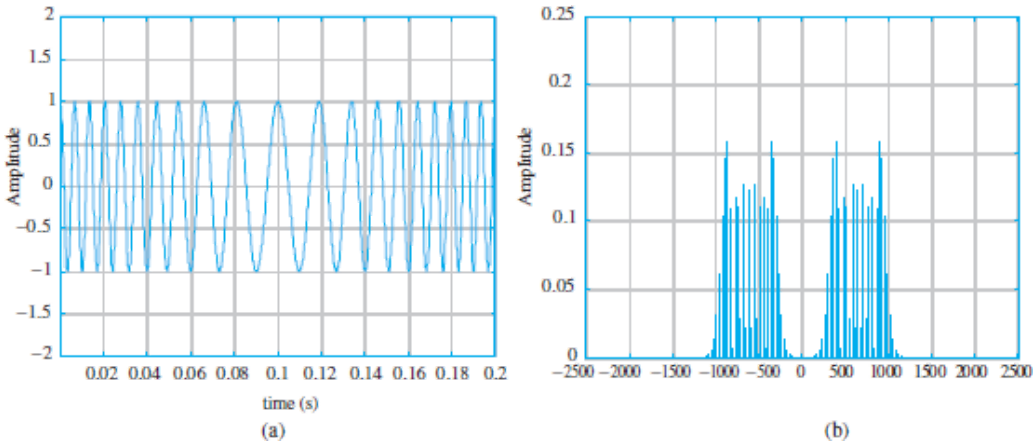


Figure 18.20 (a) FM signal time waveform; (b) FM signal magnitude spectrum

Basic Principle of FM

The basic principle underlying FM is that the **instantaneous frequency** of the carrier is modulated by the information-carrying signal. If we assume a sinusoidal

carrier, as is usually the case, say $c(t) = \cos(2\pi f_c t)$, the modulation will cause the frequency f_c to be a function of time. In the signal of Figure 18.20, the carrier frequency varies sinusoidally as well. Before proceeding with the analysis of FM signals, it will be useful to examine the relationship between the instantaneous phase and frequency of a sinusoidal signal. We first define the relationship between the instantaneous phase and frequency of a sine wave as follows:

$$\theta_i = 2\pi \int_0^t f_i(t)\,dt; \qquad f_i(t) = \frac{d\theta_i(t)}{dt} \tag{18.13}$$

To illustrate this definition, consider the case of a simple sinusoidal signal, $v(t) = A\cos(\omega t)$; in this signal we recognize that the phase angle (the argument of the cosine function) is $\theta_i(t) = \omega t$, and therefore the instantaneous signal frequency is given by $f_i(t) = (d\theta_i(t)/dt) = \omega$. This result should not surprise you: all sinusoidal signals must have a phase angle that increases linearly with time so that their instantaneous frequency is constant. In the case of an FM signal, we might Page 18-27have a phase angle that varies sinusoidally with time, thus causing the instantaneous frequency to also vary in a sinusoidal fashion; this is the simplest case of an FM signal and is treated next.

Single-Tone Modulation

Consider the case of single-tone modulation, where the modulating signal is:

$$m(t) = A_m\cos(\omega_m t) = A_m\cos(2\pi f_m t) \tag{18.14}$$

The instantaneous frequency of the FM signal varies linearly with the modulation; that is, the carrier frequency increases and decreases with the modulating signal, as shown in equation 18.15:

$$f(t) = f_c + k_f A_m\cos(2\pi f_m t) = f_c + \Delta f\cos(2\pi f_m t) \tag{18.15}$$

In the above expression we have implicitly defined the **frequency deviation** Δf:

$$\Delta f = k_f A_m \qquad \text{Frequency deviation} \tag{18.16}$$

The frequency deviation is a very important characteristic of an FM signal; Δf represents the maximum instantaneous deviation of the FM signal frequency from the carrier frequency. It is dependent on the *amplitude* of the modulating signal

and is *independent of the modulating signal frequency.* The value of the constant k_f depends on the technique used for generating the modulated signal.

The instantaneous phase of the FM signal is calculated using equation 18.13:

$$\theta_i = 2\pi \int_0^t f_i(t)\,dt = 2\pi f_c t + \frac{\Delta f}{f_m}\sin(2\pi f_m t) \tag{18.17}$$

Note that equation 18.17 introduces another important constant: the ratio of the frequency deviation to the modulating frequency is called the **modulation index** β:

$$\beta = \frac{\Delta f}{f_m} \tag{18.18}$$

The parameter β represents the maximum instantaneous deviation of the phase angle of the FM signal from the angle of the carrier. Now we can write the instantaneous angle of the FM signal as follows.

$$\theta_i = 2\pi f_c t + \beta \sin(2\pi f_m t) \tag{18.19}$$

Finally, the sinusoidally modulated FM signal is defined in equation 18.20:

$$s_{\text{FM}}(t) = A_c \cos\left[2\pi f_c t + \beta \sin(2\pi f_m t)\right] \tag{18.20}$$

It is now possible to make a formal distinction between the two cases mentioned earlier, *narrowband FM* and *wideband FM,* on the basis of the modulation index.

Narrowband FM

Narrowband FM corresponds to FM signals with a small modulation index (i.e., $\beta \ll 1$). We can use a trigonometric identity to expand equation 18.20:

$$s_{\text{FM}}(t) = A_c \cos(2\pi f_c t)\cos\left[\beta \sin(2\pi f_m t)\right] - A_c \sin(2\pi f_c t)\sin\left[\beta \sin(2\pi f_m t)\right] \tag{18.21}$$

and if $\beta \ll 1$, we can use the small-angle approximations $\cos[\beta \sin(2\pi f_m t)] \approx 1$ and $\sin[\beta \sin(2\pi f_m t)] \approx \beta \sin(2\pi f_m t)$ to write

$$\begin{aligned} s_{\text{FM}}(t) &\approx A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t)\sin(2\pi f_m t) \\ &\approx A_c \cos(2\pi f_c t) - \frac{1}{2}\beta A_c \left\{\cos\left[2\pi(f_c + f_m)t\right] - \cos\left[2\pi(f_c - f_m)t\right]\right\} \end{aligned} \tag{18.22}$$

Note the similarity between the expression for narrowband FM and that we derived earlier for the AM signal, repeated below for convenience:

$$s_{\text{AM}}(t) = A_c \cos(\omega_c t) + \mu \frac{A_c}{2} \cos(\omega_c - \omega_m)t + \mu \frac{A_c}{2} \cos(\omega_c + \omega_m)t \tag{18.10}$$

A reasonable rule of thumb is that the approximation given in equation 18.22 holds for $\beta < 0.3$. Figure 18.21(a) to (d) depicts the spectra of FM signals with various values of β. Note how the bandwidth increases with the value of the modulation index. Only in (a) and (b) is the signal narrowband FM.

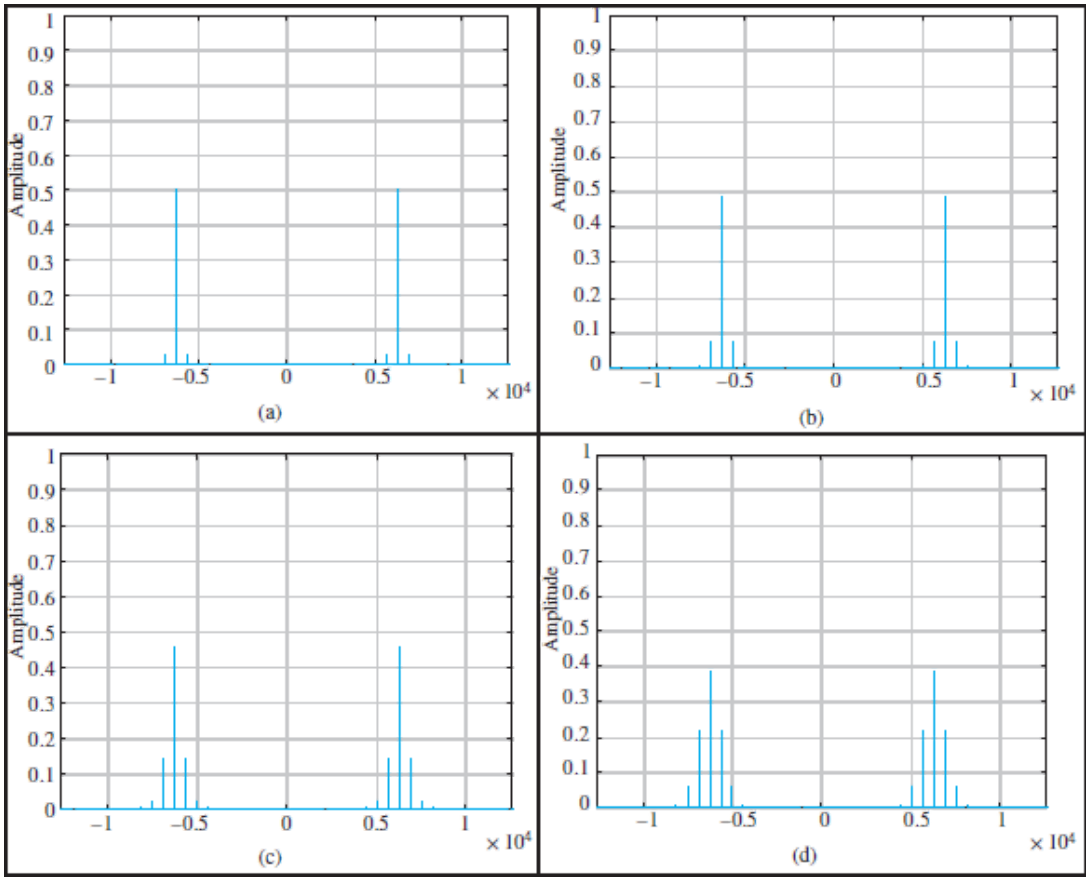


Figure 18.21 Bandwidth increases with modulation index in FM signals: (a) $\beta = 0.1$; (b) $\beta = 0.3$; (c) $\beta = 0.6$; (d) $\beta = 1$

Wideband FM

The mathematical representation of wideband FM signals is far more complex than the approximation of equation 18.22. The nonlinearity of the wideband FM signal is described using Bessel functions. This analysis is beyond the scope of

this chapter, Page 18-29and the interested reader is referred to any one of a number of excellent textbooks in electrical communications.

Transmission Bandwidth of FM Signals

The transmission bandwidth of a frequency-modulated signal is theoretically infinite; however, practical approximations are possible. In the case of narrowband FM, we have already seen that we have the same transmissions bandwidth as in an AM signal: $B = 2f_m$. For large values of the modulation index β, the bandwidth of the FM signal can be experimentally observed to be close to the total frequency excursion $\pm\Delta f$ or $2\Delta f$. These observations lead to the well-known Carson's rule, relating the approximate transmission bandwidth to the frequency deviation and to the modulation index:

$$B = 2\Delta f + 2f_m = 2\Delta f\left(1 + \frac{1}{\beta}\right) \quad \text{Carson's rule} \tag{18.23}$$

Carson's rule suggests that as β becomes larger, the bandwidth approaches $2\Delta f$, while as β decreases, the bandwidth becomes closer to $2f_m$.

FM Demodulation

Demodulation of an FM signal is accomplished by performing a **frequency-to-voltage conversion**, that is, by converting the frequency modulation into a voltage signal. This is the reverse of the modulation process and can be realized in a number of ways. We describe two basic approaches in the following subsections.

Frequency-to-Voltage Conversion

If a pulse of fixed amplitude A and fixed duration τ is generated at each zero crossing of the sensor waveform, it can be readily shown that a voltage proportional to the instantaneous signal frequency may be obtained. Figure 18.22 depicts the functional form of a frequency-to-voltage converter.

Ideally frequency-to-voltage (F-V) conversion could be obtained by computing the following integral:

$$\int_{t_{i-1}}^{t_i} v_{os}(t)\,dt = \frac{At}{\Delta T_i} = 2A\tau f_i \tag{18.24}$$

yielding a voltage proportional to the frequency of $v_S(t)$ during the ith cycle of the carrier waveform. In practice it is quite difficult to reset the integrator of Figure 18.22 at each zero-crossing, so practical F-V converters employ a low-pass filter in place of the ideal integrator.

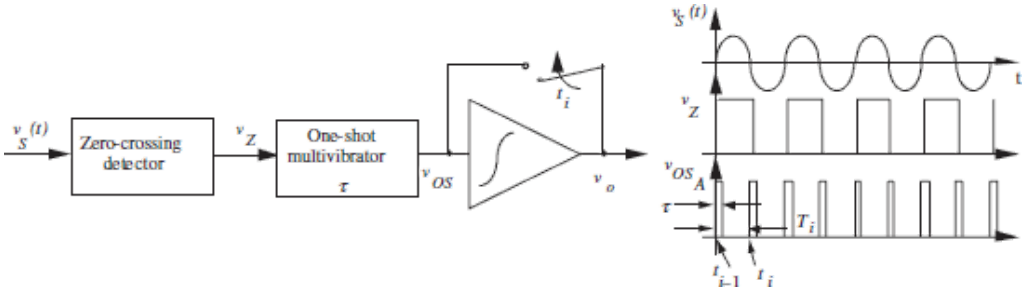


Figure 18.22 Frequency-to-voltage conversion

Page 18-30

Phase-Locked Demodulation

Another method for implementing F-V conversion utilizes a **phase-locked loop** (PLL) as a frequency-to-voltage converter, or FM demodulator. The PLL can act as an FM demodulator once it is phase-locked to an input signal waveform. When the PLL is in lock, any change in the input signal frequency generates an error voltage at the output of the phase detector, which can be either analog (a mixer) or digital, consisting of a pair of zero-crossing detectors.

The output voltage of the PLL $v_o(t)$ is the voltage that is required to maintain a voltage-controlled oscillator (VCO) running at the same frequency as the input signal and changes in the input signal frequency are matched by changes in $v_o(t)$. In this sense, the PLL acts as an F-V converter, with the input-output characteristic shown in Figure 18.23. Note that the PLL can offer only a finite *lock range*.

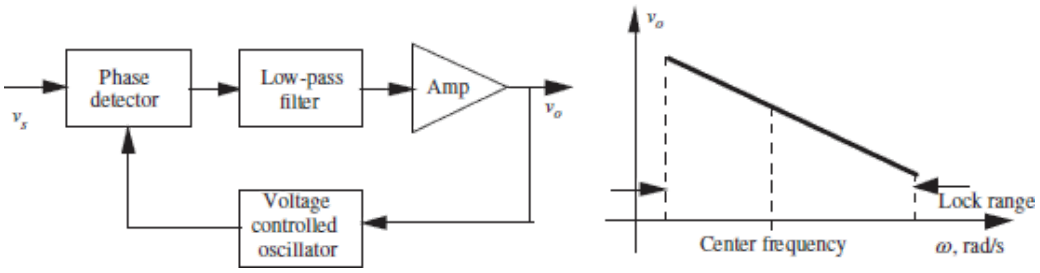


Figure 18.23 Phase-locked loop F-V conversion

EXAMPLE 18.10 Narrowband FM

Problem

Compare the spectrum of a narrowband FM waveform with that of an AM waveform with the same modulating and carrier frequencies.

Solution

Known Quantities: Carrier frequency, modulation frequency, modulation index.

Find: Plot the frequency-domain waveforms of the narrowband FM and AM signals.

Schematics, Diagrams, Circuits, and Given Data: f_c = 1,000 Hz; f_m = 100 Hz; A_c = 1 V; A_m = 0.2 V; μ = 0.2; β = 0.2.

Assumptions: Assume sinusoidal modulation and unity amplitude for the carrier, A_c = 1.

Analysis: Figure 18.24 depicts two signals: the first signal is an FM waveform with β = 0.2; the second is an AM signal with μ = 0.2. The resulting spectra are plotted in Figure 18.24(a) and (b), respectively.

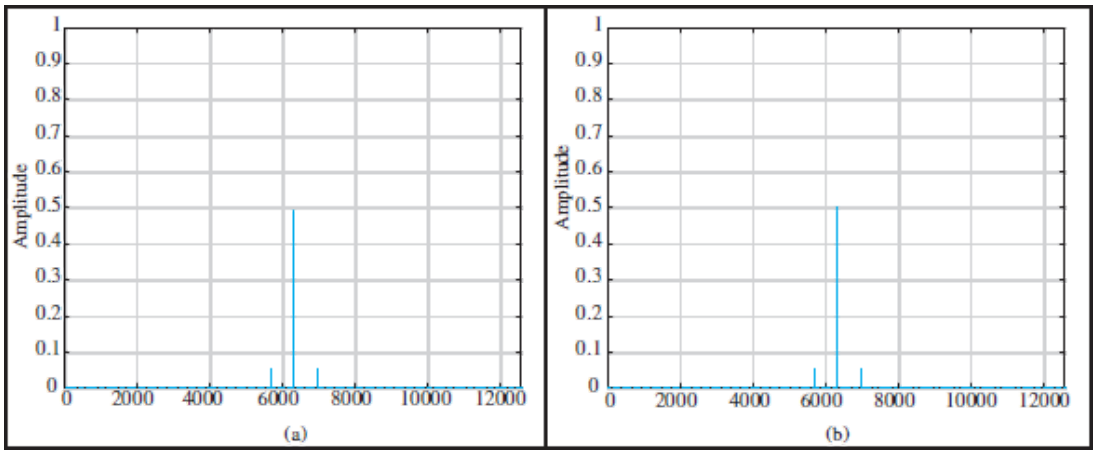


Figure 18.24 Comparison of narrowband FM and AM signal spectra: (a) FM, $\beta = 0.2$; (b) AM, $\mu = 0.2$

Comments: Note how the two amplitude spectra are virtually identical. The only difference between the two spectra is in the phase angle of the signals.

Page 18-31

LO

EXAMPLE 18.11 Commercial FM Broadcast

Problem

Use Carson's rule to analyze the bandwidth of a commercial FM station.

Solution

Known Quantities: Carrier frequency, modulation index.

Find: Approximate signal bandwidth.

Schematics, Diagrams, Circuits, and Given Data: $f_c = 90.5$ MHz; $A_c = 1$; $A_m = 1$; $f_m = 10$ kHz; $k_f = 6{,}000$; $\beta = 0.2$.

Assumptions: Assume sinusoidal modulation.

Analysis: For sinusoidal modulation, the frequency deviation is $\Delta f = k_f A_m = 6$ kHz. Then, Carson's rule predicts a bandwidth of

$$B = 2\Delta f + 2f_m = 2\Delta f\left(1 + \frac{1}{\beta}\right) = 1.2 \cdot 10^4\left(1 + \frac{1}{0.2}\right) = 72\,\text{kHz}$$

EXAMPLE 18.12 Bandwidth

Problem

Given an FM message signal, find:

a. The bandwidth of the message signal B_m.

b. The bandwidth of the modulated carrier signal B_c.

c. The band of frequencies occupied by the signal B.

Page 18-32

Solution

Known Quantities: Modulation frequency, carrier frequency, modulation constant.

Find: B_m, B_c.

Schematics, Diagrams, Circuits, and Given Data: $v_1 = 5\cos(200\pi t)$; $f_m = 100\,f_m$; $k_f = (1000/2\pi$ Hz/V).

Assumptions: Assume sinusoidal modulation.

Analysis:

a. The message signal is $v_m = 5\cos(200\pi t)$. Hence,

$$f_m = 100\,\text{Hz}$$
$$A_m = 5$$
$$f_c = 100 f_m = 10\,\text{kHz}$$

The bandwidth of the message signal is

$$B_m = 2f_m = 200 \text{ Hz}$$

b. The maximum frequency deviation is given by

$$\Delta f = k_f A_m = \frac{5{,}000}{2\pi} \text{Hz} \approx 795 \text{ Hz}$$

Thus, the bandwidth of the modulated carrier signal is approximately given by

$$B_c = 2(\Delta f + f_m) = 2(795 + 100) = 1{,}790 \text{ Hz}$$

Page 18-33

c. The carrier frequency is f_c = 10 kHz. The frequency band is centered about the carrier frequency. Therefore, the band of frequencies occupied spans from

$$\left(f_c - \frac{B_c}{2}\right) \quad \text{to} \quad \left(f_c + \frac{B_c}{2}\right)$$

Hence, the frequency band is

$$\left(10{,}000 - \frac{1{,}790}{2}\right) \quad \text{to} \quad \left(10{,}000 + \frac{1{,}790}{2}\right)$$

which equals a band from 9.105 to 10.895 kHz.

Figure 18.25 depicts the modulating signal and the FM spectrum of the FM signal examined in this example.

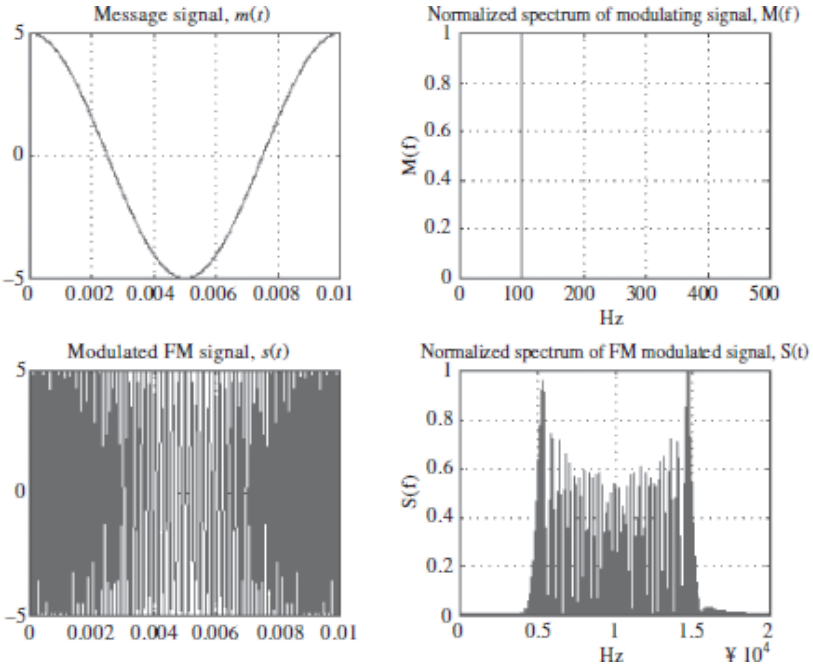


Figure 18.25 Time-domain and spectral plots for Example 18.12

EXAMPLE 18.13 Bandwidth

Problem

In the United States the assigned band for FM commercial broadcast is from 88.0 to 108.0 MHz with 100 possible channels. Find the bandwidth for each channel.

Solution

Known Quantities: Band for FM commercial broadcast, number of channels.

Find: The bandwidth of each channel.

Schematics, Diagrams, Circuits, and Given Data: See problem statement.

Assumptions: None.

Analysis: The bandwidth for each channel is defined as:

$$B = \frac{\text{Total band}}{\text{Number of channels}} = \frac{108.0 - 88}{100} = 200 \text{ kHz}$$

Comments: Each commercial FM radio station has a bandwidth allocation of 200 kHz.

CHECK YOUR UNDERSTANDING

a. Use the Matlab® files that accompany Example 18.10 to compute and plot the phase angle of the two spectra. Are the phase spectra identical?

b. What is the effect of changing the amplitude of the modulating signal on the bandwidth of the FM signal?

c. Investigate the effect of changing β on the FM signal bandwidth and analyze the spectrum of the FM signal of Example 18.10 with $\beta = 0.5$.

d. Find the carrier frequency for Channel 11 used by WCBE Radio, Columbus, OH, per the FCC regulations for the commercial broadcast in the United States (use the data given in Example 18.13).

Page 18-34

18.5 EXAMPLES OF COMMUNICATION SYSTEMS

The objective of this chapter is to summarize some important applications of modern communication systems. The overview given in this chapter is certainly not exhaustive, but it will give you a good summary of the technology underlying some of today's communication systems and of their capabilities.

Global Positioning System (GPS)

The **Global Positioning System (GPS)**, illustrated in Figure 18.26, is rapidly supplanting older navigation technologies based on radio communications used by the aircraft and marine industries, such as Loran-C. GPS is based on 24 satellites linked to ground stations and effectively replaces—with much greater accuracy—the century-old system of navigation based on star position. You can think of the satellites as "man-made stars." Differential GPS is capable of position measurements accurate to within a few centimeters. In recent years, GPS receivers have been miniaturized, and today amateur sailors, private pilots, and

other private users have the ability to purchase hand held GPS units. Among the uses of GPS we list navigation systems for cars, boats, planes, and guiding agricultural machining for "precision agriculture." The operation of GPS is explained in five basic steps, as shown in Table 18.5.

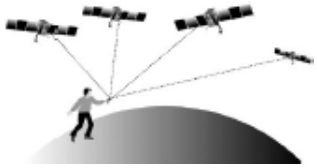


Figure 18.26 GPS principle (*Courtesy of Trimble*)

Table 18.5 Basic operation of GPS

1. GPS operates based on triangulation of signals received from satellites.
2. Triangulation is performed by the GPS receiver by measuring distance based on the travel time of radio signals.
3. Travel time is computed with the aid of accurate timing references.
4. The exact position of the satellites in space is used to calculate distance.
5. Delays experienced by the signals in traveling from the satellite to the ground stations are corrected.

Triangulation

The technique of triangulation is based on distance (range) measurements from the receiver location to the satellite. Four satellites are needed to determine the exact position of any receiver location on earth.

Measuring Distance

To measure distance, GPS computes the time required to receive each satellite signal. The receiver and satellite both generate a synchronized signal; comparison of the signal received from the satellite with that in the receiver is used to calculate the time of travel. Since the speed of travel of electromagnetic waves is known, distance can be calculated.

Timing Accuracy

Satellites carry atomic clocks on board to provide accurate timing. The clock in a low-cost GPS receiver need not be as accurate, as an additional range (distance) measurement can be used to remove the timing error in the receiver.

Satellite Positions

The positions of the satellites are essential in calculating distance. The orbits of GPS satellites are known, and any deviations are measured by the Department of Page 18-35Defense; any errors are transmitted to the satellites, which in turn transmit this error information to the receivers along with their timing signals.

Correcting Errors

Signals traveling through the ionosphere experience additional delays, which turn into transmission errors. There are many methods for correcting errors; the technique called **differential GPS** can eliminate almost all errors. Additional information on these topics may be found on the Web.

Radar

The acronym **RADAR** stands for RAdio Direction And Ranging. Radar systems operate by radiating electromagnetic waves (typically at microwave frequencies) and by detecting the echo returned from reflecting objects (targets). Radar technology was developed during World War II and played a significant part in the success of the Allied Forces. While military applications are obvious, today radar finds widespread civilian application in air traffic control and in tracking weather conditions, as well as in marine navigation (see Figure 18.27 for some examples of radar technology used in the latter application, and Figure 18.28 for an example of weather radar). In addition to detecting the position of a stationary target, radar is capable of determining the trajectory of a moving target, thus predicting its future location. This function is, for example, very useful in weather radar where one wishes to predict weather conditions. The principle on which radar operation is based is that of the **doppler shift**, a concept with which you are probably already intuitively familiar (think of the sound of a train whistle as the train moves by a stationary observer). The doppler shift permits a moving target to be distinguished from a stationary one, thus allowing the radar system to discern the echo of a stationary target from that of a moving target.

Figure 18.27 Radar antenna (top) and displays (©*Copyright ©2014 Japan Radio Co., Ltd.*)

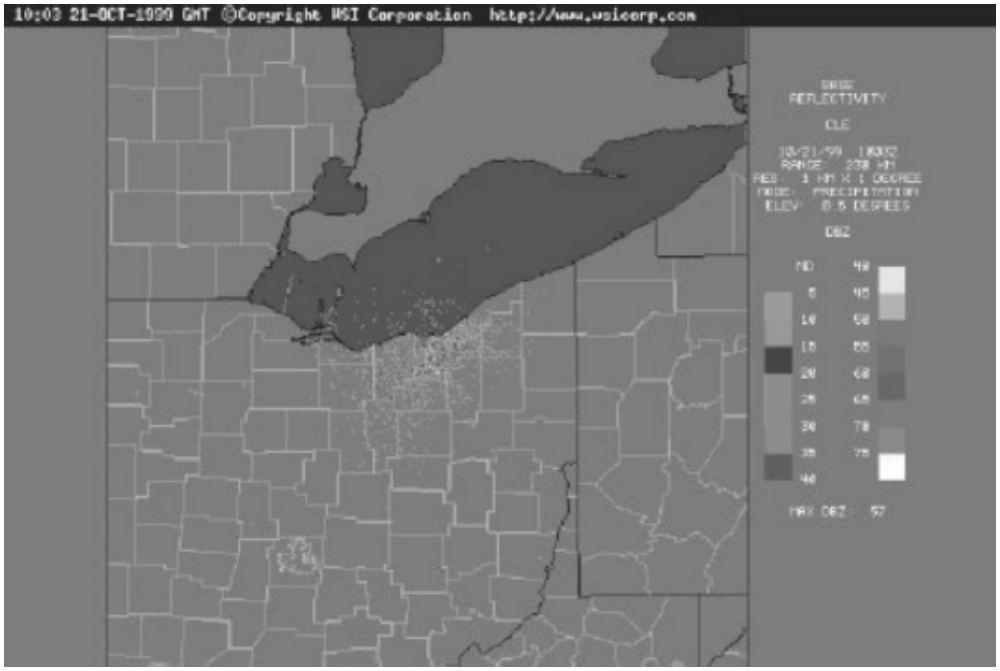


Figure 18.28 Weather radar (*Courtesy of WSI Corporation. www.wsicorp.com*)

Page 18-36

Sonar

The term **SONAR** stands for SQund NAvigation and Ranging. Sonar is conceptually similar to radar, in that it uses information about the reflection and transmission of waves to determine the behavior of targets or the properties of the environment. The principal difference between sonar and radar is that the former is based on the reflection and propagation of acoustic (sound) waves, while the latter is based on radio-frequency electromagnetic waves.

The applications of sonar are numerous, ranging from inexpensive depth finders and imaging systems used to aid the navigation of small and large sea vessels (see, for example, Figure 18.29), to underwater navigation, to ocean thermal mapping. You will find a number of interesting resources related to sonar on the Web.

Figure 18.29 Sonar displays for marine navigation (*©Copyright ©2014 Japan Radio Co., Ltd.*)

Conclusion

This chapter introduces the basic principles of analog communication systems. Upon completing the chapter you should have mastered the following learning objectives.

1. Be familiar with the most common communication systems.
2. Perform spectral analysis of simple signals using analytical and computer-aided tools.
3. Understand the principles of AM modulation and demodulation, and perform basic calculations and numerical computations on AM signals.

4. Understand the principles of FM modulation and demodulation, and perform basic calculations and numerical computations on FM signals.

HOMEWORK PROBLEMS

Section 18.2: Fourier Series and Transform

18.1 Compute the Fourier series coefficients for a periodic square wave of unit amplitude, time period τ, and duty cycle η as shown in Figure P18.1 and defined mathematically as:

$$x(t) = \begin{cases} 1 & |t| \le \eta\tau \\ -1 & |t| < \tau \end{cases}$$

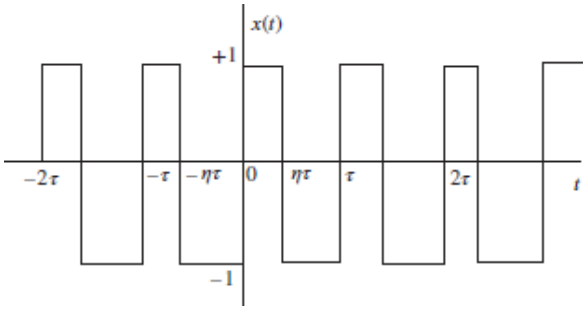


Figure P18.1 Square wave signal of periiod τ and duty cycle η

Use Matlab® to plot the frequency spectrum of this signal with time period $\tau = (1/300)$ s and duty cycle

a. $\eta = 50$ percent

b. $\eta = 30$ percent

18.2 A full wave-rectified sinusoidal signal of natural frequency ω_0 rad/s is shown in Figure P18.2.

a. Find the Fourier series coefficients for the full wave-rectified sinusoid.

b. Generate the frequency spectrum for a full wave-rectified sinusoid of natural frequency $\omega_0 = 200\pi$ rad/s.

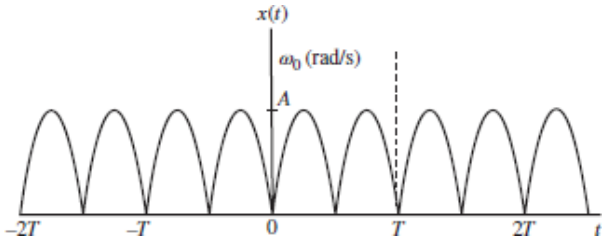


Figure P18.2 Full wave-rectified sine wave of natural frequency ω_0 rad/s

18.3 A full wave-rectified cosine signal of natural frequency ω_0 rad/s is shown in Figure P18.3.

a. Find the Fourier series coefficients for the full wave-rectified cosine.

b. Generate the frequency spectrum for a full wave-rectified cosine with natural frequency $\omega_0 = 150\pi$ rad/s.

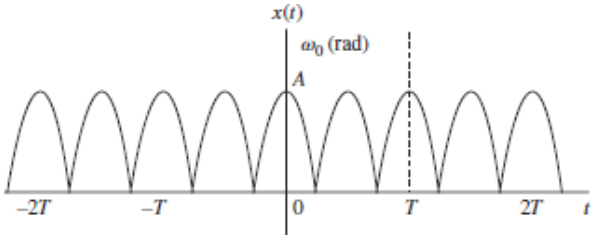


Figure P18.3 Full wave-rectified cosine wave of natural frequency ω_0 rad/s

18.4 Compute the Fourier series coefficients for the cosine-burst signal shown in Figure P18.4.

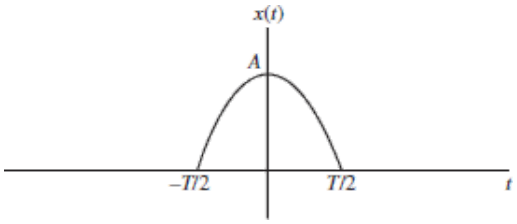


Figure P18.4 Cosine-burst signal

18.5 The triangular pulse signal shown in Figure P18.5 is mathematically defined as:

$$x(t) = A\left[1 - \frac{|t|}{T}\right][u(t+T) - u(t-T)]$$

a. Find the Fourier transform of the triangular pulse.

b. Plot the frequency spectrum of a triangular pulse of period $T = 0.01$ s and amplitude $A = 0.5$.

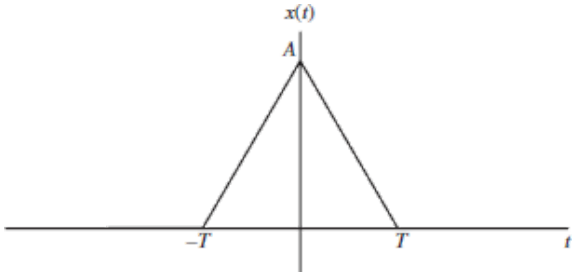


Figure P18.5 Triangular pulse x(t) of duration $2T$

18.6 A double exponential signal shown in Figure P18.6 is defined mathematically by:

$$x(t) = \begin{cases} e^{-at} & t > 0 \\ 0 & t = 0 \\ -e^{at} & t < 0 \end{cases}$$

a. Compute the Fourier transform of the signal. (*Hint:* Use the linearity property of the Fourier transform.)

b. Plot the frequency spectrum of the signal using Matlab® for $a = 8$.

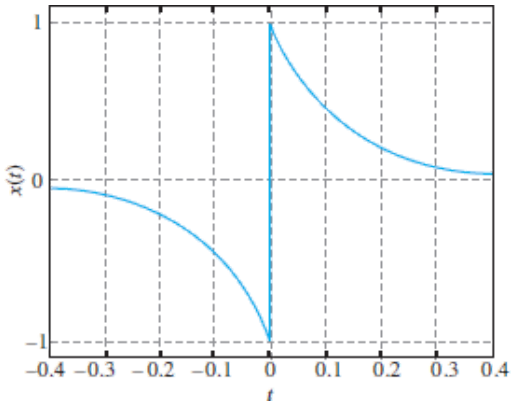


Figure P18.6 Double exponential signal

18.7 Evaluate the Fourier transform of the damped sinusoidal wave shown in Figure P18.7 and having the functional form:

$$x(t) = \exp(-at)\cos(2\pi f_c t)u(t)$$

where $u(t)$ is the unit step function.

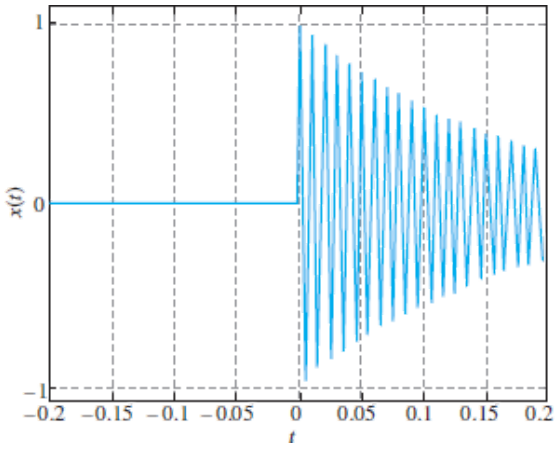


Figure P18.7 Damped sinusoidal signal

18.8 An ideal sampling function consists of an infinite sequence of uniformly spaced delta functions and is mathematically defined as:

$$\delta_{T_0}(t) = \sum_{m=-\infty}^{\infty} \delta(t - mT_0)$$

Page 18-38

a. Compute the Fourier transform of the ideal sampling function shown in Figure P18.8.

b. Also, use Matlab® to generate the time-domain signal and its amplitude spectrum for $T_o = 0.01$ s.

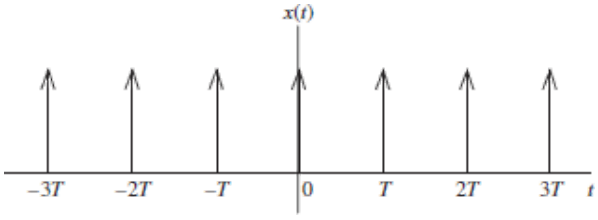


Figure P18.18 Dirac delta function with period T_0

18.9 Download the utterance signal utter.au and use the Matlab® command **auread('utter.au')** to load it to the workspace. Use the FFT tools in Matlab® to identify the frequency components of the signal.

18.10 A **bat echolocation chirp** signal is provided to you on the book website. Do a frequency analysis of the signal and explain what you observe. (*Courtesy Digital Signal Processing Group, Rice University*.)

Section 18.3: Amplitude Modulation and Demodulation

18.11 Find the modulation index μ for an AM signal having a carrier of amplitude $A_c = 1.0$, and where the amplitude of the carrier at the maximum is $A_{\max} = 3.0$ and at the minimum is $A_{\min} = 0.6$.

18.12 Plot the anticipated frequency spectrum of a carrier signal with an amplitude of unity and frequency $f_c = 1.3$ MHz that is amplitude modulated ($\mu = 1$) with a signal $m(t)$, where

$$m(t) = 0.8\sin(2\pi 5{,}000t) + 0.4\sin(2\pi 10{,}000t) + 0.2\sin(2\pi 20{,}000t)$$

18.13 Plot the anticipated time-domain response of a carrier signal with an amplitude of unity and frequency $f_c = 10$ MHz that is amplitude modulated ($\mu = 1$) with a signal $m(t)$, where

$$m(t) = A\left[1 - \frac{|t|}{T}\right] T = 0.01 \text{ s}$$

Hint: The message signal is a triangular wave of time period T.

18.14 An AM radio station uses a carrier signal of unity amplitude and frequency $f_c = 1.6$ MHz. The message signal is a voice signal having certain frequency components and defined as:

$$m(t) = 0.4\sin(2\pi 340t) + 0.35\sin(2\pi 960t) + 0.3\sin(2\pi 1{,}345t) + 0.2\sin(2\pi 2{,}230t) + 0.1\sin(2\pi 2{,}890t)$$

Plot the time-domain and the frequency-domain AM signal of modulation index $\mu = 1$.

18.15 A non-periodic message signal $m(t)$ is amplitude modulated ($\mu = 1$) by a carrier signal of unity amplitude and frequency $f_c = 0.5$ MHz. Plot the time- and frequency-domain signal.

$$\begin{aligned} m(t) &= 0.4\sin(2\pi 340t) + 0.35\sin(2\pi 960t) \\ &\quad + 0.3\sin(2\pi 1{,}345t) + 0.2\sin(2\pi 2{,}230t) \\ &\quad + 0.1\sin(2\pi 2{,}890t) + u(t) \end{aligned}$$

$$u(t) = \begin{cases} 1 & t \le 0.01 \\ 0 & \text{Otherwise} \end{cases}$$

18.16 Consider a modulating wave $m(t)$ that consists of a single frequency component and is defined as:

$$m(t) = A_m \cos(2\pi f_m t)$$

where A_m is the amplitude of the modulating wave and f_m is the frequency. The sinusoidal carrier wave has amplitude A_c and frequency f_c. The signal is amplitude modulated to produce the signal $s(t)$. Find the average power delivered to a 1-Ω resistor by $s(t)$.

18.17 The carrier frequency of the W-OSU channel is 0.82 MHz. If the upper sideband of the AM signal has frequency components of amplitude 0.4 at 0.825 MHz, 0.2 at 0.83 MHz, and 0.25 at 0.84 MHz,

a. Find an expression for the modulating signal.

b. Plot the spectrum of the modulating signal.

c. Plot the spectrum of the AM signal including the lower sideband.

18.18 The AM commercial radio band in the United States is authorized to operate from 525 kHz to 1.7 MHz. A carrier frequency is assigned to each station, and regulations require them to be separated by 10 kHz. Find:

a. The number of channels that can be accommodated in the given frequency range.

b. The maximum modulating frequency that can be transmitted without overlap.

18.19 The speech signal utter.au is to be amplitude modulated for transmission on an AM commercial radio band in the United States. Plot the frequency spectrum of the AM signal. Use any channel according to its separations and a modulation index $\mu = 1$ in your design.

Section 18.4: Frequency Modulation and Demodulation

18.20 The message signal given by $m(t) = 5\cos(750\pi t)$ is frequency modulated by a carrier frequency 105 times the message frequency, and the modulation constant is $k_f = 1{,}005$. Find the bandwidth of the message signal.

18.21 If the message signal given by $m(t) = 2\cos(360\pi t)$ is frequency modulated by a carrier frequency 100 times the message frequency and the modulation constant $k_f = 1{,}000$, find the bandwidth of the modulated carrier signal.

18.22 Find the band of frequencies occupied by the FM signal of Problem 18.21.

18.23 A message signal $m(t)$ is frequency modulated by a carrier of unity amplitude and frequency $f_c = 10.0$ MHz, with modulating constant $k_f = 1{,}000$. Plot the time-and frequency-domain FM signal if $m(t) = 0.8\sin(2\pi 5{,}000t)$.

18.24 A packet of information is sent on an FM channel of frequency $f_c = 15.0$ MHz that uses a modulating constant $k_f = 6{,}000$. Plot the frequency spectrum of the FM signal.

$$m(t) = 0.4\sin(2\pi 340t) + 0.35\sin(2\pi 960t) + 0.3\sin(2\pi 1{,}345t) + 0.2\sin(2\pi 2{,}230t) + 0.1\sin(2\pi 2{,}890t) + u(t)$$

$$u(t) = \begin{cases} 1 & t \le 0.001 \\ 0 & \text{Otherwise} \end{cases}$$

18.25 WOSU-FM uses a carrier frequency of 90.5 MHz and modulating constant $k_f = 66{,}000$. The speech signal utter.au is transmitted on this channel. Plot the frequency spectrum of the FM speech signal.

18.26 Consider Example 18.13. If Channel 2 is allocated for country music and the message signal may be considered to be $m(t) = 10\cos(2\pi\ 10^3 t)$, find:

a. The carrier frequency.

b. The value of k_f.

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for "Focus on measurements" weighing scales: Media Bakery.

[1]An example of this phenomenon is the changeover from analog to digital systems in cellular telephony. Both systems coexist at the present time, but it is reasonable to forecast that in a few years all *personal communication systems* will be digital.

[2]Although negative frequencies have no physical significance, the mathematical form of Fourier series and transforms requires that we consider both positive and negative frequencies.

C H A P T E R

19

DIGITAL COMMUNICATIONS

This chapter expands upon the analog communications concepts of Chapter 18 and introduces digital communications, or the idea of using digital processing techniques to simultaneously maximize information throughput while minimizing the probability of error, with the added constraint of limited transmitting power.

The field of digital communications was ushered in with the invention of the telegraph by Samuel Morse. After overhearing two of his colleagues speaking about electromagnets, Morse was struck with the idea of communicating information by using an electromagnetic clacker to indicate the flow of current in a long wire loop, where the flow of current was controlled by a switch at the other end (see Figure 19.1). Because such a machine can only communicate using the presence ("high") or absence ("low") of electric current, any message passed through the machine needed to be coded using these two states. To solve this problem, Morse invented a simple digital coding technique consisting of variable-length high and low states. High states that were shorter in duration were called *dots*, those that were longer in duration were called *dashes*, and combinations of the two

were used to indicate each letter of the alphabet. Short low states were used to indicate breaks between letters while long low states were used to indicate breaks between words. This technique, later to become the well-known Morse code, was first publicly demonstrated Page 19-2by transmitting the phrase “What hath God wrought,” from the Supreme Court room at the Capitol to the railway depot in Baltimore on May 24, 1844.

Figure 19.1 An early telegraph key (*Ryan McVay/Getty Images*)

Digital communication improves upon analog communication in many ways. In general, digital signals can be designed to be much more resistant to the effects of noise. Digital communication can have enhanced privacy through the use of encryption. And digital communication systems are usually less expensive to implement. Most importantly, digital communication allows the engineer to optimize a system, given a limited amount of transmitter power, for maximum information throughput while having predictable control over the integrity of the information at the receiver. On the downside, digital communication systems are generally more complex to design and require more bandwidth than their analog counterparts.

19.1 A TYPICAL DIGITAL COMMUNICATIONS SYSTEM

Although there are a great variety of digital communication systems, many of them can be broadly represented by the system of Figure 19.2. The input to the system is some *information source*, which may supply information in an analog (continuous) or digital (discrete) form. In the analog case, an *analog-to-digital converter* is used to convert the information into digital format. This is followed by a *source coder* that removes unnecessary redundancy and reduces the information signal to its most basic form. A *channel coder* then reintroduces a measured amount of redundancy that is

carefully designed to improve the robustness of the information when exposed to noise. Next, the *baseband modulator* converts the digital information into an analog waveform that is shifted up to the desired carrier frequency by the *upconverter*. Once at the carrier frequency, the information is carried over the *channel*, which represents the physical medium over which we are communicating. On the receiver side, the counterparts to the transmitter components serve to recover the original information accurately.

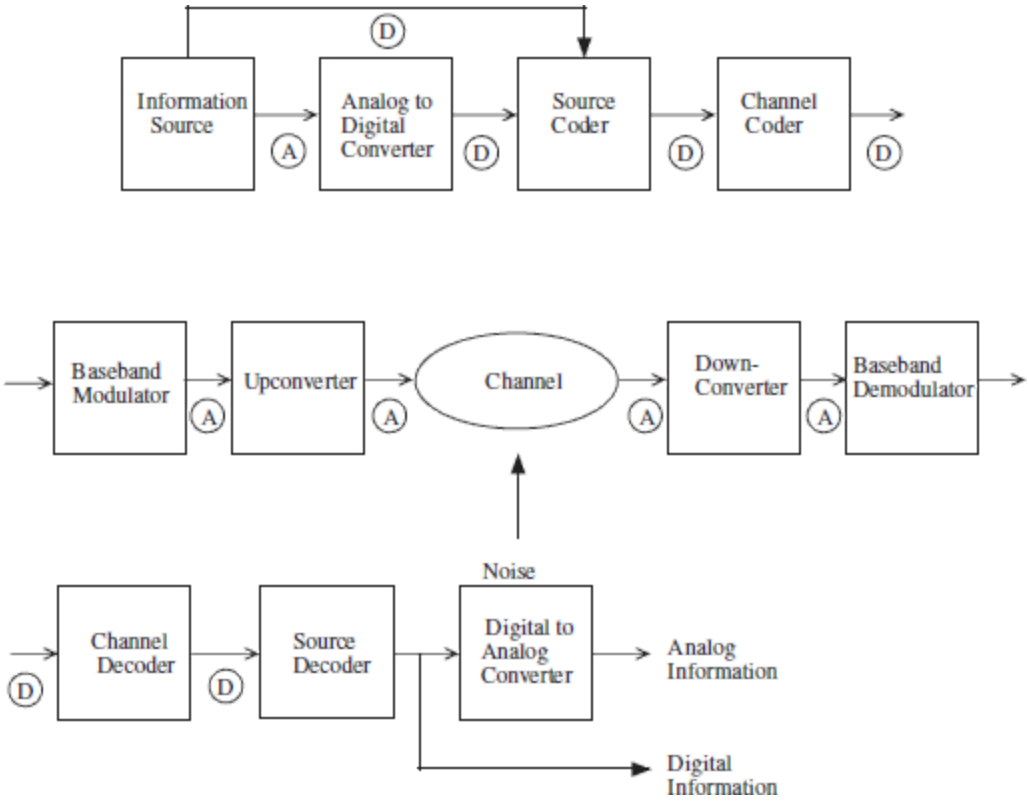


Figure 19.2 A typical communications system

19.2 INTRODUCTORY PROBABILITY

It is impossible to discuss digital communications without encountering probability theory, so in this section we present a very brief introduction to the topic. We include only enough material to allow a basic understanding of the digital communication concepts to follow. For a more thorough treatment of probability theory, the reader should examine a full text on the topic.

Random Experiments

A *random experiment* is any experiment where the outcome cannot be exactly predicted and may differ from trial to trial. Flipping a coin, tossing a die, or spinning a roulette wheel are all examples of random experiments. A random experiment is described by the *sample space* Ω of all possible *outcomes* the experiment may produce. For example, flipping a coin once has the sample space

$$\Omega = \{\text{head, tail}\}$$

while the experiment consisting of flipping a coin twice has sample space

$$\Omega = \{\{\text{tail, tail}\}, \{\text{head, tail}\}, \{\text{tail, head}\}, \{\text{head, head}\}\}$$

Rolling a six-sided die has the sample space

$$\Omega = \{1,2,3,4,5,6\}$$

The *outcome* obtained during a particular trial is denoted by ω.

An *event* E is a subset of Ω. For example, the event of getting at least one tail in two flips is

$$E = \{\{\text{tail, tail}\}, \{\text{head, tail}\}, \{\text{tail, head}\}\}$$

The event of rolling a number less than four on the die above would be

$$E = \{1,2,3\}$$

Since events are subsets of experiment outcomes, we can apply the set theory concepts of union, intersection, and complement. For example, given sets A and B, we can create the *union* event $A \cup B$ containing all outcomes that are in events A *or* B. The *intersection* event $A \cap B$ consists of all outcomes that belong to both events A and B. The *complement* event A^C is

the set of all outcomes that do not belong to A. Note that $\varnothing$ indicates the empty set, i.e., $\Omega^C = \varnothing$. Returning to the six-sided die roll example on p. 000, consider the events $E_1 = \{1, 2, 3\}$ where the roll is less than four and $E_2 = \{2, 4, 6\}$ where the roll is even. Then the union, intersection, and complements of these events are as shown in Table 19.1.

Table 19.1 Examples of set theory

Union	$E_1 \cup E_2 = \{1, 2, 3, 4, 6\}$	Outcomes where the roll is less than four *or* even
Intersection	$E_1 \cap E_2 = \{2\}$	Outcomes where the roll is less than four *and* even
Complement	$E_1^C = \{4,5,6\}$	Outcomes where the roll is *not* less than four
Complement	$E_2^C = \{1,3,5\}$	Outcomes where the roll is *not* even

Probability

Given a random experiment, the *probability* is a real function P representing the likelihood (from 0.0 to 1.0, with 1.0 being the most certain) that a given event contains the outcome of a random experiment trial. As an example, given the events E_1 and E_2 above, we have the probabilities shown in Table 19.2.

Table 19.2 Examples of probabilities

Event	Probability
$E_1 = \{1, 2, 3\}$	$P(E_1) = 1/2$
$E_2 = \{2, 4, 6\}$	$P(E_2) = 1/2$
$E_1 \cup E_2 = \{1, 2, 3, 4, 6\}$	$P(E_1 \cup E_2) = 5/6$
$E_1 \cap E_2 = \{2\}$	$P(E_1 \cap E_2) = 1/6$
$E_1^C = \{4,5,6\}$	$P(E_1^C) = 1/2$
$E_2^C = \{1,3,5\}$	$P(E_2^C) = 1/2$

Probabilities adhere to the following properties:

1. $P(E) \leq 0$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \varnothing$
4. $P(\varnothing) = 0$
5. $P(E) = 1 - P(E^C)$

6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Conditional Probability

The conditional probability $P(E_1 \mid E_2)$ is the probability of an outcome ω being in E_1 if we already know that ω is in E_2. For example, if we define the events E_1 and E_2 as above, then $P(E_1 \mid E_2)$ asks the question, "What is the probability that the outcome is less than four if we already know that the outcome is even?" In this context, you can think of E_2 as a filter of sorts, narrowing down the possible outcomes Page 19-5from which E_1 might be selected. In this case, if we know that the outcome of the die roll is even, then we can intuitively see that there is a one-in-three chance that it is less than four. This can be represented mathematically by the relation

$$P(E_1 \mid E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$$

In the example above, the intersection of the two sets is

$$P(E_1 \cap E_2) = \{2\}$$

and therefore

$$P(E_1 \mid E_2) = \frac{1/6}{1/2} = \frac{1}{3}$$

as we found above.

Random Variables

A *random variable* is a one-to-one mapping from a random experiment's sample space Ω onto the real number line. Considering the example of the coin toss given above, for example, we might map heads to +1 while mapping tails to –1. A graphical depiction of this is given in Figure 19.3. We can express this mapping mathematically by writing

$$X(\omega) = \begin{cases} +1 & \omega = \text{heads} \\ -1 & \omega = \text{tails} \end{cases}$$

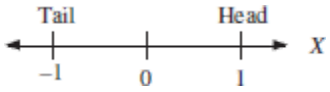


Figure 19.3 Defining a random variable for the coin toss

which is referred to as a *discrete random variable* because each outcome is mapped to a discrete number on the real axis.

Once a random variable has been defined for an experiment, its *cumulative distribution function* (CDF) is expressed as:

$$F_X(x) = P(X(\omega) \leq x) \tag{19.1}$$

which represents the probability that the value of a random variable for a given trial will be less than or equal to the argument of the CDF. For example, the CDF of the coin toss example above is given by

$$F_X(x) = \begin{cases} 0 & x < -1 \\ 1/2 & -1 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

and illustrated in Figure 19.4. From this figure, we see that there is no chance of the random value being less than –1, which is obvious because the minimum possible value of X is –1. There is a 50 percent chance ($F_X = 0.5$) that X will be less than 1, which is its value when the outcome is tails. Finally, there is a 100 percent chance ($F_X = 1.0$) that the value of X will be less than or equal to +1 since there is no possible outcome for which the random value is more than +1.

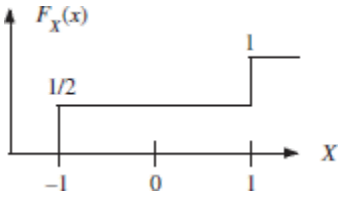


Figure 19.4 CDF of a coin toss random variable

From the CDF, one can define the *probability mass function* (PMF) given by

$$p_i = P(X = x_i)$$

The PMF tells us the probability that a discrete random variable is equal to a particular value. For the example above, we can write the PMF as

$$p_1 = P(X = -1) = 1/2$$
$$p_2 = P(X = +1) = 1/2$$

Similarly, a *continuous random variable* is one that is mapped to some continuous range of the real number line, such as if we were to measure the amount of rainfall within a particular area on any given day. Such a random experiment might have a CDF that looks like Figure 19.5.

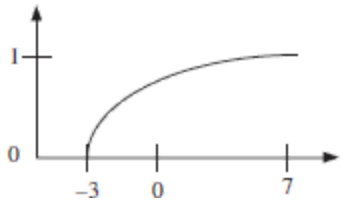


Figure 19.5 CDF of continuous random variable

For continuous random variables, it is normal to define the *probability density function* (PDF) as

$$f_X(x) = \frac{d}{dx} F_X(x) \tag{19.2}$$

One commonly encountered PDF is that for a *gaussian* random variable, given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2} \tag{19.3}$$

where m is the mean, or average value, and σ^2 is the *variance*, or spreading factor. The gaussian is encountered so often that it is frequently written as $N(m, \sigma^2)$.

The *standard normal* Φ is a special case of the gaussian where $m = 0$ and $\sigma^2 = 1$. Plots of a few gaussian distributions are given in Figure 19.6.

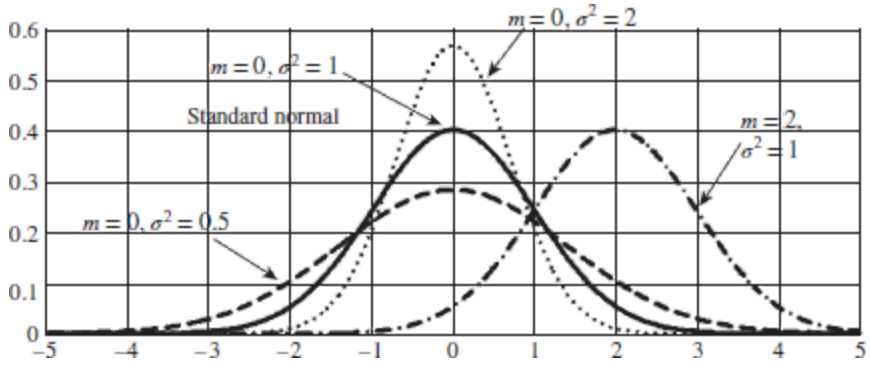


Figure 19.6 PDFs of gaussian distributions

For any random variable, we can use the PDF to find the probability that X falls between two values for a given outcome by integrating the area under the PDF for that interval. When X is the standard normal, for example, the probability that X is greater than some value a is given by the integral

$$P(\Phi > a) = \int_a^\infty f_\Phi(x)\, dx \tag{19.4}$$

The value of the integral above can be found from widely available tabulated values of the *tail function* (also called the Q function) defined by

$$Q(a) = \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-x^2/2} dx \tag{19.5}$$

Although the tail function is tabulated for the standard normal, it can be applied for any gaussian through a simple variable transformation given by

$$P(X > a) = Q\left(\frac{a - m}{\sigma}\right)$$

where X is a gaussian random variable.

19.3 PULSE-CODE MODULATION

Pulse-code modulation (PCM) is simply the process of converting an analog signal (continuous time and continuous amplitude) into a digital signal (discrete time and discrete amplitude) by applying sampling and quantization as previously described in Section 7.4 and illustrated in Figure 19.7.

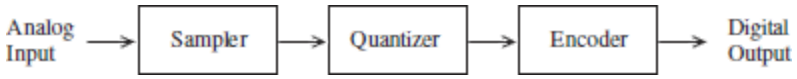


Figure 19.7 Block diagram of a pulse-code modulator

For a well-designed digital system, it is important to know the extent of any errors that are introduced into the system as a by-product of the PCM process. Assuming that the sampling rate is sufficient for the frequency content of the signal according to the Nyquist theorem, any such error introduced into the system as a result of PCM will come from quantization alone. As you recall from Section 7.4, a quantizer's resolution Δ is set by the number of bits used to represent its output.

$$\Delta = \frac{v_{\max} - v_{\min}}{2^b} \tag{19.6}$$

where $v_{\max}$ and $v_{\min}$ define the dynamic range of the quantizer and b is the number of bits. The *quantization noise* is then defined as the difference between the sampled signal and the quantized version. To find a relationship between the number of bits and the quantization noise, consider the sampled continuous-amplitude signal and its quantized version shown in Figure 19.8. From this figure, we can see that the maximum possible error due to quantization noise is $\pm\Delta/2$, meaning that the maximum average quantization noise power $P_{qn}^{\max}$ is $(\Delta/2)^2$ or

$$P_{qn}^{\max} = \frac{1}{2^{2b}} \frac{(v_{\max} - v_{\min})^2}{4}$$

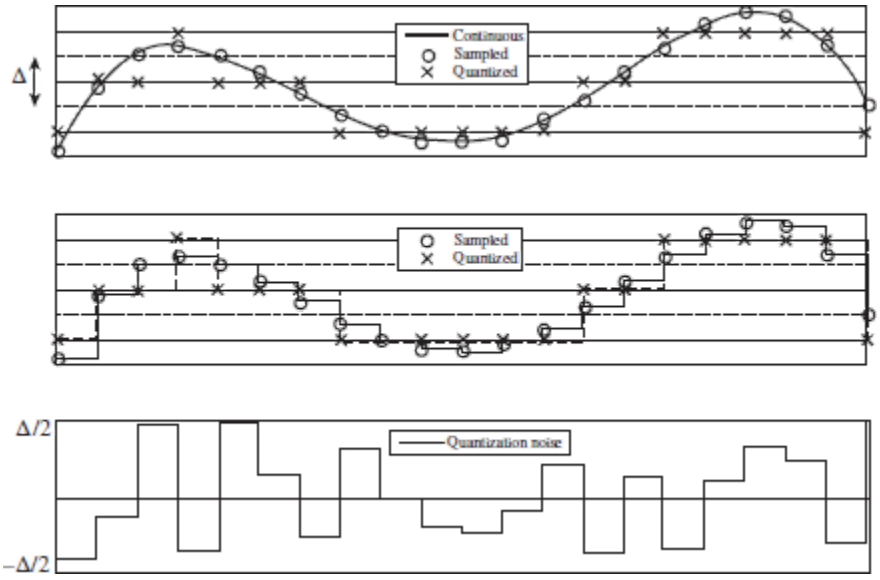


Figure 19.8 A quantized signal and its noise

Assuming some signal power P_s, the signal–to–quantization noise ratio (SQNR) is therefore proportional to the number of bits and is given by

$$SQNR = \frac{P_s}{P_{qn}} \propto 2^{2b} \qquad \text{or} \qquad \boxed{SQNR_{dB} \propto 10 \log 2^{2b} = 6b} \tag{19.7}$$

From the above equation we can discover a very important rule of thumb: each additional bit of resolution we add to a quantizer will increase the SQNR of its output by about 6 dB. Thus, we can expect the output of a 10-bit quantizer to have a SQNR of about 60 dB.

Page 19-8

19.4 SOURCE CODING

Source coding is the process of removing the redundancy of a signal and reducing it to its most basic form, retaining only the basic information that is necessary to communicate the message.

Redundancy is any message content carried by a signal beyond the fundamental information being communicated. A source of information contains redundancy if its messages can be represented with complete fidelity in a more terse form. Redundancy may or may not serve a purpose.

For example, when announcing over a public address system that a parked vehicle has its lights on, it is redundant, but useful, to announce the license plate number several times.

Human speech is one of the best examples of redundancy in nature—have you ever listened to a scratchy AM weather radio broadcast during a thunderstorm? Even if you haven't, you can imagine the quality of the audio you might hear as the electrons contained within the thunderstorm contribute noise to the signal as it travels from the radio station to our receiver. Yet even with the extra noise, the details of the weather report can usually still be understood by the listener. Even though a tiny fraction of the original signal survives the trip from the speaker's mouth to our ears, we are still able to decode the information contained within it. Signal redundancy, therefore, often serves to increase the probability that a message is understood by the receiver.

Unfortunately, the price paid for redundancy is efficiency. Consider the same radio example as above, except with a strong signal and no thunderstorm. Although the audio has much more clarity, we receive the same weather report details as before. Therefore, we can conclude that in this case redundancy is unnecessary, and with the absence of noise in the communications channel the message could be communicated just as accurately without it.

Compressibility

Source coding can also be thought of as *data compression* because it results in a message that can be represented using fewer bits. *Lossless* compression codes the signal so that it can be reconstructed exactly by the receiver while *lossy* compression Page 19-9allows some of the message to become unrecoverable in order to gain additional compression by representing the message with fewer bits. The field of *information theory* allows us to determine a signal's fundamental information content, or *entropy*, which is a measure of its average uncertainty. From the entropy we can determine the signal's *compressibility*, or how much extraneous information can be removed before some of the fundamental message contained within the signal becomes unrecoverable. Thus, the entropy of the signal allows us to establish a line between lossy and lossless compression.

Consider an experiment where the result X can be chosen from the set $\{a_1, a_2,..., a_N\}$ and the probability of that result being a_n is $p_n = P(X = a_n)$. The amount of information communicated by each possible result is inversely proportional to the probability of that result occurring. For example, knowing the stock market is going to crash tomorrow represents much more information than knowing it's not going to crash. This same idea is applicable to any experiment, not just the stock market, so we conclude that the amount of information contained by a particular result is dependent only on the probability of that result occurring and not on the result itself. This measure of *self-information* is notated by I and defined as (see Figure 19.9)

$$I(X = p_n) = \log_2\left(\frac{1}{p_n}\right) \tag{19.8}$$

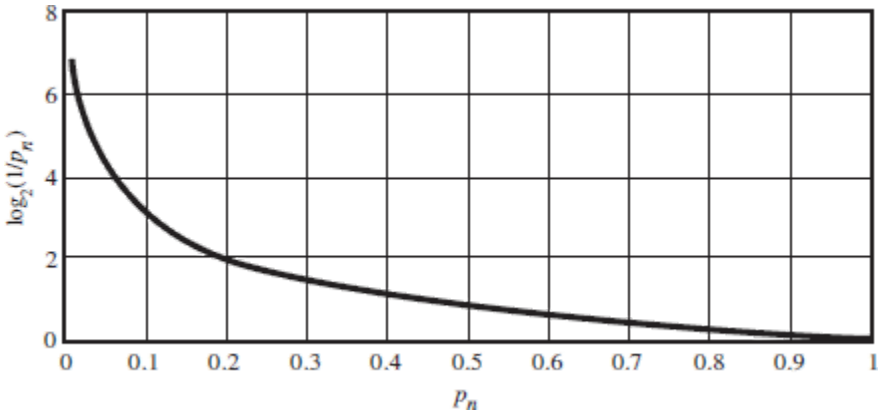


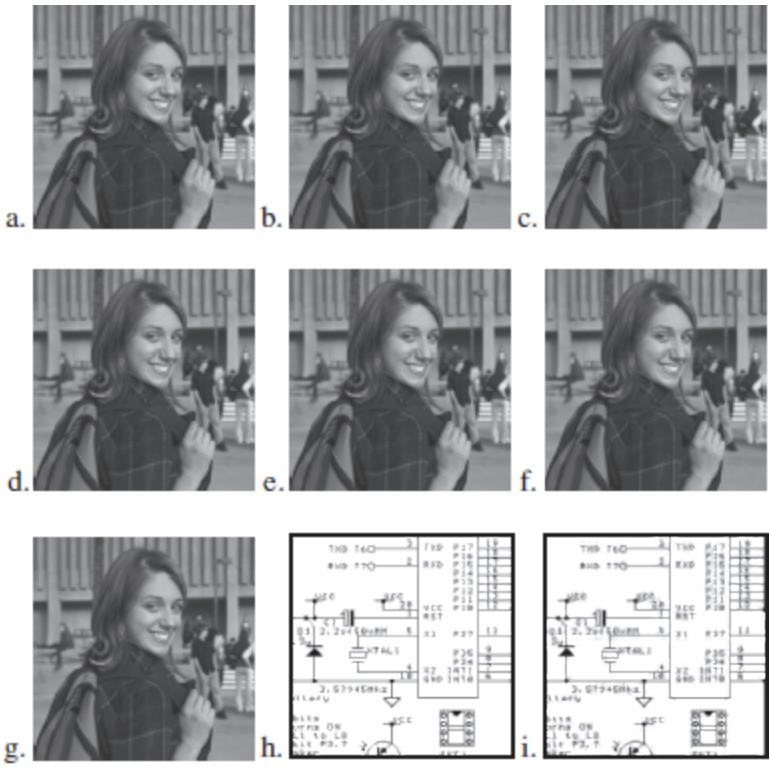
Figure 19.9 Self-information versus probability of occurrence

The total entropy $H(X)$ represented by the experiment is then the weighted average of the self-information of each result and the probability of that result occurring, as given by

$$H(X) = \sum_{n=1}^{N} p_n \log_2\left(\frac{1}{p_n}\right) \tag{19.9}$$

A good example of lossy versus lossless compression can be found by examining the storage of digital images. The two most common image storage formats in use today are the *Graphic Interchange Format* (GIF)

created by CompuServe (now America Online) in 1987, and the format named after the *Joint Photographic Experts Group* (JPEG) created by that organization in 1986. While both formats serve to compress digital images, they go about it in much different ways. The GIF standard uses an implementation of the Lempel-Ziv-Welch (LZW) algorithm which Page 19-10is lossless and therefore guarantees that the original image can be recovered exactly. The JPEG standard uses a discrete cosine transformation (DCT) to convert the image to the spectral domain where high-frequency coefficients below a threshold set by the user are filtered out. Once these coefficients are filtered out they cannot be recovered, and therefore the technique is lossy. Determining which compression technique to use usually depends on the application. For photographic images, JPEG usually results in a smaller file with little noticeable loss. However, for images of technical diagrams involving thin lines, JPEG usually results in larger files and may introduce unacceptable artifacts (see Figure 19.10).



	Format	Loss	Quality	Size (bytes)
a.	GIF	Lossless	n/a	264,358
b.	JPEG	Lossless	100	542,866
c.	JPEG	Lossy	90	73,213
d.	JPEG	Lossy	75	42,187
e.	JPEG	Lossy	50	17,170
f.	JPEG	Lossy	25	12,099
g.	JPEG	Lossy	10	10,179
h.	GIF	Lossless	n/a	3,117
i.	JPEG	Lossy	20	7,079

Figure 19.10 Examples of GIF and JPEG file sizes versus quality (*Mike Watson Images Limited/Glow Images*)

Source Coding Techniques

Once the entropy of a signal is known, a code must be devised so that individual events can be represented by sequences of bits, or *codewords*, for transmission. Consider the event sequence

$$X = \{a_1, a_2, a_1, a_4, a_3, a_1, a_1, a_2\}$$

At first glance, one might try coding each event with a unique binary codeword of fixed length, as shown in Table 19.3.

Table 19.3 Simple fixed-width coding scheme

Event	Probability	Codeword
a_1	$p_1 = 1/2$	00
a_2	$p_2 = 1/4$	01
a_3	$p_3 = 1/8$	10
a_4	$p_4 = 1/8$	11

The assembled bit sequence representing the message X would then be

$$\{00,01,00,11,10,00,00,01\} = 0001001110000001$$

with 2 bits per event for a total of 16 bits to represent the entire sequence message. By computing the entropy of the above signal, we can determine the efficiency of the simple coding scheme above. By using equation 19.9 and Table 19.3, we find the entropy to be

$$H(X) = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{8}\log_2 8 + \frac{1}{8}\log_2 8 = 1.75$$

meaning that the theoretical minimum bound is 1.75 bits per event to represent the message without loss. Since our simple proposed code uses 2 bits per pixel, we know that there must be a more efficient way to code the message. In the following sections we examine a few such coding techniques that are more efficient than our simple method above.

Huffman Source Coding

Huffman source coding is a lossless technique that represents frequently occurring events with short codewords and rarely occurring events with longer ones. In practical applications, Huffman coding is generally used as a secondary coder following some other coding technique. For example, ZIP (a popular computer compression format), JPEG, and MP3 use Huffman encoding as a secondary step.

For an example of Huffman coding, consider the sequence above with the code shown in Table 19.4.

Table 19.4 Huffman coding scheme

Event	Probability	Codeword
a_1	$p_1 = 1/2$	1
a_2	$p_2 = 1/4$	01
a_3	$p_3 = 1/8$	001
a_4	$p_4 = 1/8$	000

The scheme defined in Table 19.4 would lead to a bit pattern of

$$\{1,01,1,000,001,1,1,01\} = 10110000011101$$

with a total of 14 bits or 1.75 bits per event, which is a savings of 12.5 percent over the fixed-width code above and meets the theoretical limit given by the Page 19-12entropy computed above. Because Huffman codewords are not fixed length, however, care must be taken when choosing the sequences so that they will be *uniquely decodable*. To be uniquely decodable, codewords should be constructed in such a way that the decoder can distinguish between consecutive events in a bit sequence. The decoder knows that an event is ending when it either detects a 1 in the bit sequence or when three 0s are received in a row (indicating an a_4).

The procedure for creating a Huffman code is as follows:

1. Gather statistical data about the frequency of occurrence of each potential event.
2. Combine the two least probable events (assign them bits 0 and 1) into a new single event that occurs when either of the two original events occurs. The probability of this new event occurring is then the sum of the probabilities of the original events.
3. If only two events remain, go to step 4. Otherwise go back to step 2.
4. Assign 0 and 1 as bits representing the remaining two events. These steps are illustrated graphically in Figure 19.11.

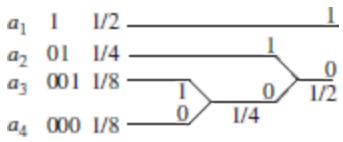


Figure 19.11 Creating a Huffman code

Lempel-Ziv-Welch Source Coding

While Huffman coding gives very good results that are close to the theoretical limit given by the entropy, it has two major disadvantages. First, a detailed knowledge of the statistical probability of each event occurring must be known beforehand. With many common information sources such as photographs, music and speech recordings, and streaming video, these statistics cannot be predicted. Secondly, Huffman coding is efficient only for codewords that represent a single event. If events are known to occur frequently in patterns (say $\{a_1, a_1, a_2\}$), then building a Huffman code library for all such patterns quickly becomes unmanageable regardless of how efficient such a code might be.

The Lempel-Ziv-Welch (LZW) code solves these two problems by building a running *dictionary* of event patterns rather than the events themselves. This dictionary is built by examining the event sequence to identify unique recurring patterns and adding them to the dictionary as they are encountered. Then, the pattern is replaced with a fixed-width codeword indicating the matching entry is the dictionary.

To keep the size of the dictionary small, new event pattern entries are represented as the concatenation of an existing dictionary entry with the "new" event. Note that unlike Huffman encoding, the LZW approach requires the dictionary to be transmitted along with the message.

As an example, consider the event sequence given by

$$X = \{a_1, a_2, a_1, a_1, a_2, a_3, a_1, a_2, a_1, a_1, a_4\}$$

To create a LZW code for this sequence, we add unique patterns to the dictionary as they are encountered. The resulting grouped sequence and dictionary are given by

$$X = \{\{a_1\}, \{a_2\}, \{a_1, a_1\}, \{a_2, a_3\}, \{a_1, a_2\}, \{a_1, a_1, a_4\}\}$$

By substituting codes from Table 19.5, the encoded bit sequence is found to be

$$X = \{000,001,010,011,100,101\} = 0000010100111001$$

and uses a total of 18 bits or 1.64 bits per event.

Table 19.5 Example Lempel-Ziv-Welch coding scheme

Entry	Existing entry	New event	Effective sequence
000	n/a	a_1	a_1
001	n/a	a_2	a_2
010	000	a_1	a_1, a_1
011	001	a_3	a_2, a_3
100	000	a_2	a_1, a_2
101	010	a_4	a_1, a_1, a_4

LZW coding is particularly useful in applications where very long or repetitive event patterns exist, such as images of technical drawings where there may be many consecutive pixels of the same color. It is typically less useful in applications where event patterns are largely random or occur infrequently, such as with photographs.

Parametric Source Coding

Both of the coding methods described above are lossless and rely on the logical layout of the data to achieve compression rates near the entropy limit. For many types of sources these codes function very close to the theoretical limit determined by the entropy of the source. For sources such as photographs and music that have little logical structure, however, these codes usually result in poor performance.

Fortunately, it is frequently possible to achieve very efficient compression rates for such sources by leveraging knowledge about the ultimate application of the data. For example, fine details in a photograph that are imperceptible to the naked eye can be removed, as with JPEG, without significant detriment. Codes using such biometric techniques belong to a class known as *parametric source codes*. Parametric source

codes generally achieve their impressive compression ratios by transforming a sequence to be encoded into some spectral domain and removing components that are unlikely to be missed. Because of this property, such codes are almost always lossy, and unlike the Huffman and LZW codes, they must be custom tailored for the intended receiver.

One of the most famous parametric codes of recent times was created in 1987 by Professor Dieter Seitzer of the University of Erlangen in association with the *Motion Picture Experts Group* for compression of audio, technically referred to as "Layer III." This standard, commonly known as MP3, can reduce the size of music files by a factor of 10 or more by filtering out certain components of sound that are inaudible to the ear in much the same way that JPEG removes photographic details that are sharper than the eye can perceive.

EXAMPLE 19.1 Construction of a Huffman Code

Problem

Considering the events and probabilities given in Table 19.6, compute the entropy, create an efficient Huffman code, and find its average code length.

Table 19.6

Event	Probabilities
a_1	$p_1 = 0.6$
a_2	$p_2 = 0.3$
a_3	$p_3 = 0.1$

Solution

The entropy of this experiment is given by

$$H(X) = 0.6\log_2\frac{1}{0.6} + 0.3\log_2\frac{1}{0.3} + 0.1\log_2\frac{1}{0.1} = 1.3$$

The Huffman code is created as shown in Figure 19.12. The average codeword length is then computed from Table 19.7 as

$$E(L) = 0.6 \cdot 1 + 0.3 \cdot 2 + 0.1 \cdot 2 = 1.4$$

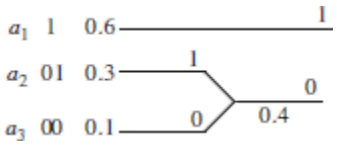


Figure 19.12 Creating a Huffman code for Example 19.1

Table 19.7 Example Huffman code

Event	Probability	Codeword
a_1	$p_1 = 0.6$	1
a_2	$p_2 = 0.3$	01
a_3	$p_3 = 0.1$	00

which is close to the theoretical minimum of 1.3 bits per event given by the entropy computation.

EXAMPLE 19.2 Construction of a Lempel-Ziv-Welch Code Problem

For the event sequence

$$X = \{a_3, a_3, a_3, a_2, a_3, a_3, a_1, a_1, a_2, a_3\}$$

create a LZW dictionary and encode the sequence to bits.

Solution

First, the sequence is broken into unique patterns as they occur. This results in the grouped sequence and dictionary given by

$$X = \{\{a_3\},\{a_3,a_3\},\{a_2\},\{a_3,a_3,a_1\},\{a_1\},\{a_2,a_3\}\}$$

From Table 19.8, the encoded bit sequence is then

$$\{000,\ 001,\ 010,\ 011,\ 100,\ 101\} = 000001010011100101$$

Note that the resulting bit sequence is identical to the sequence found in the preceding text discussion. Only the dictionary has changed.

Table 19.8 Example Lempel-Ziv-Welch coding scheme

Entry	Existing entry	New event	Effective sequence
000		a_3	a_3
001	000	a_3	a_3, a_3
010		a_2	a_2
011	001	a_1	a_3, a_3, a_1
100		a_1	a_1
101	010	a_3	a_2, a_3

19.5 DIGITAL BASEBAND MODULATION

We live in a world that is physically analog, so information must be in an analog form before being transmitted across the physical medium. However, this is not to be confused with analog communication systems where the information to be transmitted and the waveform that is transmitted are both analog (e.g., an analog voice signal). In digital communication systems, the information to be sent from the transmitter to the receiver is digital, but it is analog while traveling over the physical medium. Baseband digital modulation converts a binary signal into an analog signal that can be transmitted over the physical communication channel.

Digital *baseband modulation* transmits bits of information across a communication channel by mapping those bits to particular analog message waveforms to indicate the states of those bits. The analog waveforms used could represent voltage on a coaxial cable, electric-field potential in a wireless link, acoustic pressure waves in an underwater link, or light pulse amplitude in a fiber-optic line. By stringing together a number of these waveforms, one after the other, an entire sequence of bits can be transmitted. The receiver, using *baseband demodulation*, then decodes the message by matching the received waveforms to an internal library of previously agreed upon waveforms to recover the original bits.

Pulse-Position Modulation

One set of waveforms that is commonly used for digital baseband modulation is shown in Figure 19.13, where a 0 bit is represented by $m_0(t)$ and a 1 is represented by $m_1(t)$. Because the state of the bit is represented by the position of the pulse within the overall duty cycle T, these are known as *pulse-position modulated* (PPM) waveforms.

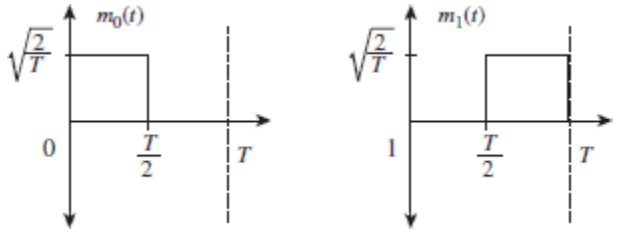


Figure 19.13 Binary pulse-position modulated (PPM) message waveforms with duty cycle T

To send information in chunks of M simultaneous bits, the set with $K = 2^M$ message waveforms can be used, each with a duration of T seconds. For example, to send 2 bits at a time we can use four message waveforms to convey all the possible bit states, as illustrated in Figure 19.14.

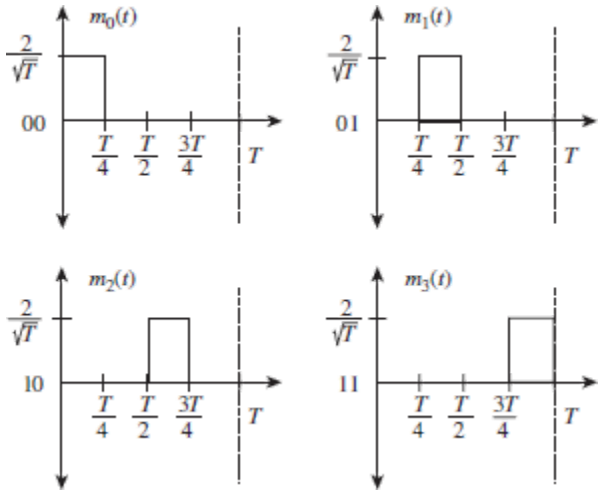


Figure 19.14 PPM for 2 bits per message waveform

The transmission rate of a digital system depends on the duty cycle T as well as on the number of bits that are being transmitted simultaneously. The *baud rate* is $1/T$, or the number of message waveform transmissions per second. The *bit rate* is M/T, or M times the baud rate since M bits are sent per transmission. For example, the V.90 telephone modem standard specifies a baud rate of 8,000 message waveform transmissions per second and a bit rate of 56,000 bits/s. Thus, each message waveform conveys 7 bits by transmitting one out of a set of 128 message waveforms.

Page 19-16

Noise added to the signal as it travels through the channel can corrupt the message that arrives at the receiver. Mathematically, the received signal $r(t)$ is written

$$r(t) = m_i(t) + w(t)$$

where $m_i(t)$ is one of K message waveforms and $w(t)$ is an additive noise signal. For example, Figure 19.15(a) shows a noise signal $w(t)$, and Figure 19.15(b) shows the message waveform corrupted by the noise $w(t)$, where the message waveform is chosen from the binary PPM waveform set shown in Figure 19.13. By examining the received signal in Figure 19.15(b), we see that it more closely resembles $m_1(t)$ than $m_0(t)$ in Figure 19.13. Thus, we conclude that a 1 was sent.

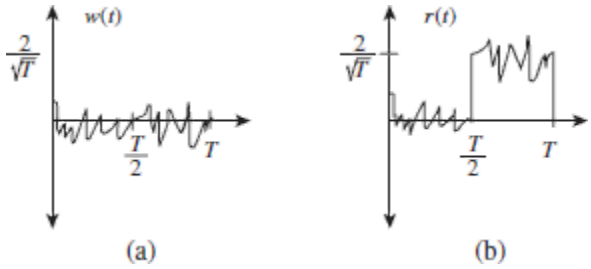


Figure 19.15 (a) Gaussian noise; (b) signal corrupted by noise

In using our intuition to determine which waveform was transmitted, we eventually come to the conclusion that the received signal $r(t)$ was somehow "closer" to one message waveform than the other. By measuring this "distance" between waveforms, the receiver automatically decides which message waveform was probably transmitted. One approach is to calculate the *root-mean-squared* (RMS) *distance* by adding up the squared absolute pointwise differences between the waveforms. The RMS distance d between $r(t)$ and a given waveform $m_i(t)$ is given by

$$d[r(t), m_i(t)] = \sqrt{\int_0^T |r(t) - m_i(t)|^2 \, dt} \tag{19.10}$$

Once the distances between the received signal and each of the candidate waveforms are calculated, the receiver chooses the candidate with the minimum RMS Page 19-17distance from the received waveform. Detection theory tells us that if the message waveforms are equally likely to be sent, then this *minimum distance* (MD) detection algorithm is optimal in the sense of minimizing the probability of choosing the wrong waveform. Mathematically, the MD detector is written

$$\hat{i} = \arg \min_{i \in \{0,1,\ldots,K-1\}} \left[\int_0^T |r(t) - m_i(t)|^2 \, dt \right] \tag{19.11}$$

where we have removed the square root since it doesn't affect the determination of the minimum. Use of the *arg* operator returns the index of the minimum value and not the minimum value itself. A block diagram of the MD detector is shown in Figure 19.16.

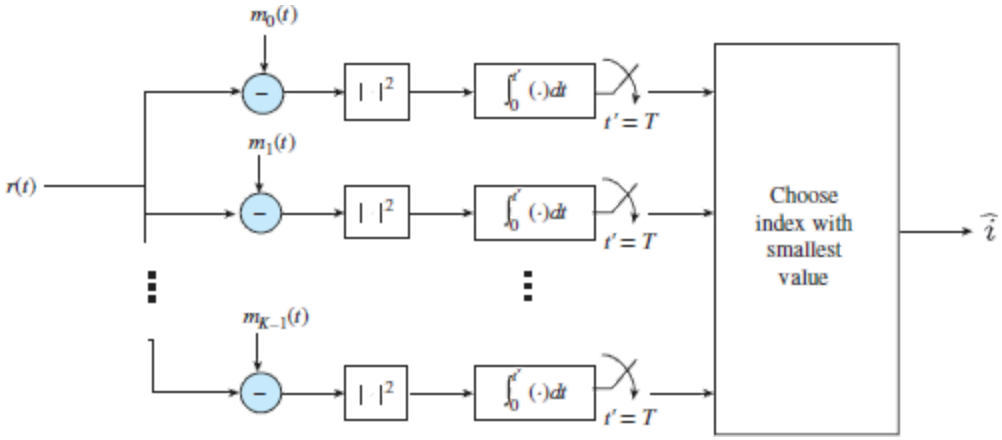


Figure 19.16 Minimum distance detector

In the absence of noise, the MD detector will find a distance of zero between the matching $m_j(t)$ and the received signal, and a nonzero distance between the received signal and all other message waveforms. As a result, the message waveforms should be designed to be as far apart from each other as possible so that noise disturbance cannot easily make the received signal closer to a message waveform that was not sent. In the ideal case there would be infinite distance between message waveforms to make the error probability go to zero. Unfortunately, the transmit power is constrained by the *average energy* E of the message waveform set, calculated as

$$E = \frac{1}{K}\sum_{i=0}^{K-1}\int_0^T |m_i(t)|^2 dt \tag{19.12}$$

If two message waveform sets have the same average energy and one has a greater minimum distance between message waveforms, then in general it will have lower error probability. This is explored further in the chapter problems.

The MD detector can be changed to an equivalent form, called the *correlation receiver*, shown in Figure 19.17. In some cases, the correlation receiver can be easier to implement than the MD detector. The correlation receiver multiplies the samples of the received signal with each message waveform, integrates the result, samples at $t' = T$ seconds, subtracts the signal energy, and then chooses the index Page 19-18of the largest value.

The multiply-integrate operation is called *correlation*. The correlation receiver can be derived from the MD detector through the following algebraic manipulation:

$$\begin{aligned}\hat{i} &= \arg \min_{i\in\{0,1,\ldots,K-1\}} \left[\int_0^T |r(t)-m_i(t)|^2\,dt\right] \\ &= \arg \min_{i\in\{0,1,\ldots,K-1\}} \left[\int_0^T |r(t)|^2 - 2r(t)m_i(t) + |m_i(t)|^2\,dt\right] \\ &= \arg \min_{i\in\{0,1,\ldots,K-1\}} \left[-2\int_0^T r(t)m_i(t) - \frac{|r(t)|^2}{2} - \frac{|m_i(t)|^2}{2}\,dt\right] \\ &= \arg \max_{i\in\{0,1,\ldots,K-1\}} \left[\int_0^T r(t)m_i(t) - \frac{E_i}{2}\,dt\right]\end{aligned} \tag{19.13}$$

where $E_i = \int_0^T |m_i(t)|^2\,dt$ is the *energy* of the ith waveform. In the last line, we changed the *min* to *max* by negating.

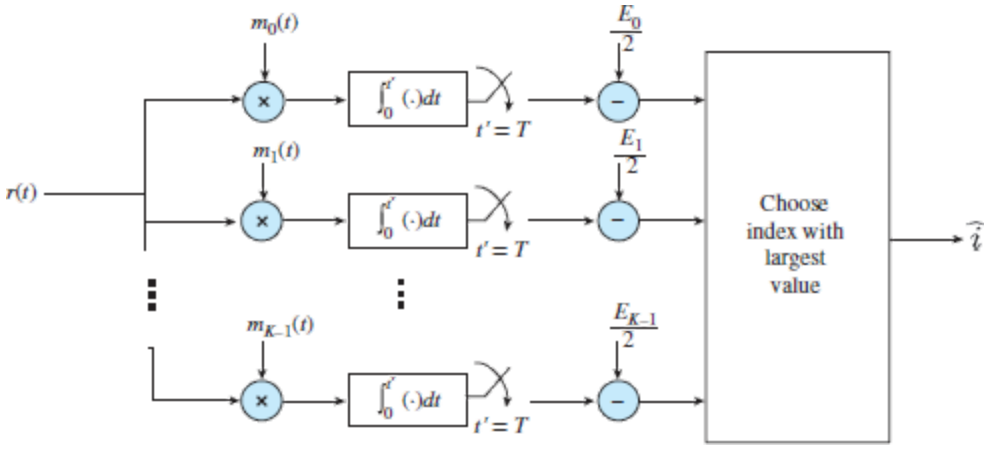


Figure 19.17 Correlation detector

Pulse-Amplitude Modulation

In *pulse-amplitude modulation* (PAM), the set of K message waveforms are built by specifying amplitudes A_i of a prototype pulse $m(t)$. Thus, the ith waveform is

$$m_i(t) = A_i m(t)$$

If we assume that the prototype pulse $m(t)$ has unit energy, i.e.,

$$\int_0^T |m(t)|^2 dt = 1$$

Page 19-19

then each message waveform has energy $E_i = A_i^2$. The correlation detector shown above can now be simplified to:

$$\begin{aligned}\hat{i} &= \arg \max_{i\in\{0,1,\ldots,K-1\}} \left[\underbrace{A_i \int_0^T r(t)m(t)dt}_{y} - \frac{A_i^2}{2} \right] \\ &= \arg \max_{i\in\{0,1,\ldots,K-1\}} \left[A_i y - \frac{A_i^2}{2} \right] \\ &= \arg \min_{i\in\{0,1,\ldots,K-1\}} \frac{1}{2}[A_i^2 - 2yA_i]\end{aligned}$$

where we changed the *max* to a *min* by negating. We complete the square and ignore terms common to all indices to write the decision as:

$$\hat{i} = \arg \min_{i\in\{0,1,\ldots,K-1\}} (A_i - y)^2 \qquad (19.14)$$

From this equation we see that when using PAM the decision is made simply by finding the amplitude that has minimum distance to the correlation output y. This result greatly simplifies the receiver structure, as seen in Figure 19.18.

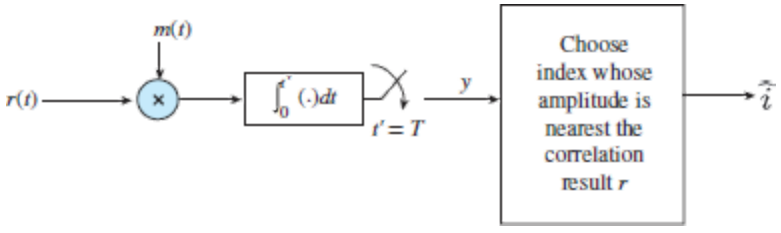


Figure 19.18 PAM detector

For a moment, consider a PAM signal set with a unit energy square prototype pulse as shown in Figure 19.19.

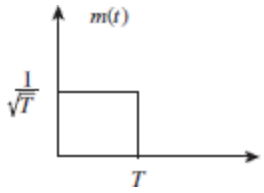


Figure 19.19 Unit energy prototype pulse for PAM signal set

For binary PAM, we choose amplitudes $A_0 = +1$ and $A_1 = -1$. The resulting message waveforms are shown in Figure 19.20. In this case, the correlation output y will be

$$y = \pm 1 + z$$

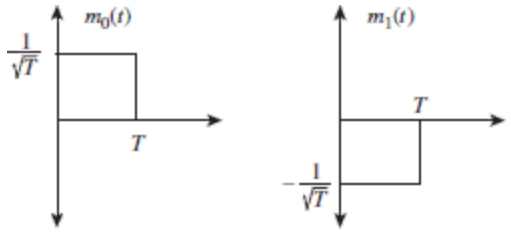


Figure 19.20 Binary PAM message waveforms

where

$$z = \int_0^T m(t)\, w(t)\, dt \tag{19.15}$$

is the noise. So far in our discussion we have avoided discussing the specific nature of the noise signal z. As it turns out, noise sources such as thermal noise, atmospheric interference, and channel distortion can be modeled as *white noise*, or noise with a constant power at all frequencies. Specifically, the power spectral density for such noise is given as

$$S_n(f) = \frac{kT}{2} \tag{19.16}$$

where k is *Boltzmann's constant* (1.38×10^{-23}J/K) and T is the *noise temperature* in kelvin. This model is actually non-physical because a constant power spectral density at all frequencies would mean that the total power of the noise signal is infinite. As we saw in Chapter 18, however, engineers are usually only interested in a very narrow band of the spectrum at a time. As a result, as long as the white noise model is applied within that narrow band we can avoid breaking any physical laws. The ratio of signal power to noise power is known as the *signal-to-noise ratio* (SNR).

Additive white gaussian noise (AWGN) is also a signal with constant power density at all frequencies, but with the additional property that for any particular sample n_0 its value is a random variable Z with gaussian PDF. Because the value of the signal z for a given sample n_0 is described by a random variable, the noise signal itself is known as a *random process*.

To see the effects of AWGN on transmission of bits over a channel, consider a system where a single bit is transmitted using PAM with $A_0 = -1.0$ V to represent a 1 bit and $A_1 = -1.0$ V to represent a 0 bit. After the addition of AWGN, the output y for bit n_0 will have a PDF given by

$$y[n_0] = \pm 1 + N(0, \sigma^2) = N(\pm 1, \sigma^2) \tag{19.17}$$

and resemble Figure 19.21, where the output for bit n_0 can take on any value according to the probability prescribed by the gaussian distribution.

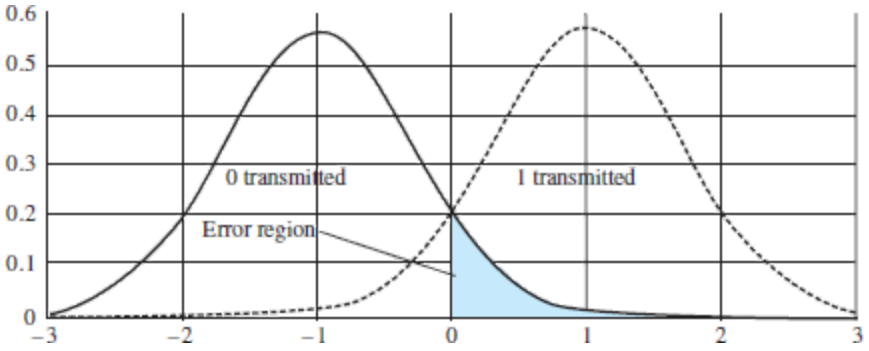


Figure 19.21 Likelihood probability curves (signal + noise) of bits transmitted over an AWGN channel

Now, consider a decoder that is designed to recover the intended value of the transmitted bit. Such a decoder $d[n]$ might use a simple comparator to decide Page 19-21whether the output bit is a 1 or 0, such as

$$d[n] = \begin{cases} 0 & y[n] \geq 0 \\ 1 & y[n] < 0 \end{cases}$$

The decoder interprets an output less than 0.0 V to indicate a 0 bit and more than 0.0 V to indicate a 1 bit. Unfortunately, the output of this decoder will be in error whenever the noise signal causes the output to cross over the zero threshold, as identified by the shaded region in the figure for a transmitted 0 bit (–1.0 V). The probability of this occurring is the area of the shaded region, as given by

$$\begin{aligned} P(d = 1|x = -1) &= P(y > 0|x = -1) \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \int_0^\infty e^{-(x+1)^2/2\sigma^2} dx \\ &= Q\left(\frac{1}{\sigma}\right) \end{aligned}$$

Through symmetry, the same probability can be found for $P(d = 0 \mid x = 1)$.

Binary Symmetric Channel Model

The *binary symmetric* channel model simplifies the AWGN model by encapsulating the error mechanism into a single parameter, $P(y \mid x)$, or the probability that event y was received given that x was transmitted. Mathematically, the binary symmetric channel is described by specifying

$$\begin{aligned} 1 - \epsilon &= P(0|0) = P(1|1) \\ \epsilon &= P(1|0) = P(0|1) \end{aligned}$$

where ϵ is the *crossover probability*, meaning that a transmission error occurred such that the output does not equal the input; in other words, the probability that the bit was "flipped" (see Figure 19.22). This simpler model is often used to model computer storage or transmission systems.

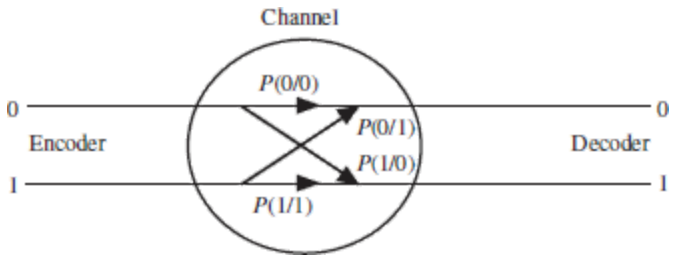


Figure 19.22 Binary symmetric channel

19.6 CHANNEL CODING

In a previous section, we showed how source coding is used to represent a sequence of events using the minimum possible number of bits. We also showed how the entropy equation set a limit on lossless compression and could be used to compute the efficiency of the coder. As we will see in the following section, a similar theorem exists to show that for a given transmitting power and bandwidth, there is a transmission rate at which one can accomplish effectively error-free transmission Page 19-22of data over a noisy channel. This transmission rate is called the *channel capacity*. Just as source coding is the mechanism that allows sources to approach the theoretical limit of compressibility, *channel coding* is the mechanism used to approach this theoretical error-free transmission rate by adding sufficient redundancy to improve the likelihood that the message can pass accurately through the communications channel over which it is transmitted or stored.

As an example, consider storage of music on a compact disc (CD). From the listener's perspective such discs can store about 74 min of audio, or the equivalent of 650 Mbytes of data, in the form of microscopic pits in a layer of fragile metal foil encased in a piece of plastic. To reproduce the music, the CD player must provide the means to accurately recover the CD's data even in the presence of normal wear and tear which may include contamination of its surface with dust, oil, or scratches. To provide this functionality, an additional 200 Mbytes of redundant data is stored on the CD to allow the player to perform automatic error detection and correction as the data bits are read from the disc. For CDs used for computer data storage, the percentage of redundant data is even higher.

The Channel Model and Additive White Gaussian Noise

Before discussing channel coding schemes, we must characterize the nature of the communication channel over which communication will take place. Although there are many complicated methods for modeling communication channels, the most common (and simplest) is the *additive channel*. Using the system diagram of Figure 19.23, we can characterize the output $y[n]$ of an additive channel as the linear superposition of the message signal $x[n]$ with some noise signal $z[n]$, expressed mathematically as:

$$y[n] = x[n] + z[n]$$

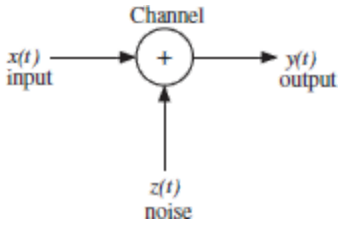


Figure 19.23 Modeling the additive communications channel

Shannon Noisy Channel Capacity

Just as the entropy of the message defined the absolute limit on lossless compression, a similar limit exists for the maximum error-free transmission rate that can be achieved over a given channel. This limit is defined by the *Shannon noisy channel coding theorem*, which for a simple channel with AWGN results in

$$C = W\log_2\left(1 + \frac{P}{kTW}\right) \tag{19.18}$$

where C is the transmission rate in bits/s, W is the available bandwidth of the channel, and P is the power transmitted per event.

Linear Block Channel Coding

Linear block coding is the simplest of all the channel coding techniques. As with source coding, linear block coding relies on a dictionary of codes that

will be substituted for the input presented to the coder. However, while source coders use dictionaries that reduce the total number of bits, channel codes use carefully chosen codewords that actually increase the number of bits and therefore add redundancy. As previously mentioned, adding redundancy does not come for free—the additional bits require either more time to transmit or increased bandwidth to accommodate the higher bit rate, not to mention increased system complexity.

Page 19-23

In general, (n, k) linear block codes map 2^k input sequences of k bits into n-bit output sequences. Consider the following very simple (2, 1) code as described by the dictionary in Table 19.9. This code converts each input bit into two output bits, essentially doubling the number of bits that must be transmitted. This simple code improves the reliability of the code by providing *error-detection* capability. In this case, the extra bit in the pair functions as a *parity bit*, meaning that the received code should always contain an even number of 1s, or *even parity*. If it doesn't, then the decoder knows that one of the bits was distorted in the channel. This simple code doesn't protect against two errors occurring in a single codeword (meaning both bits were flipped in the channel), and it also doesn't provide any capability to correct errors. However, it does give the receiver a rudimentary ability to detect when errors occur.

Table 19.9 An even parity error-detecting channel code

Input	Codeword
0	00
1	11

To examine a code that provides *error-correction* capability, consider the (5, 2) code of Table 19.10. This code converts each pair of input bits into five output bits. From the receiver's perspective it is easy to decode a received bit pattern if it exactly matches one in the dictionary. But what happens when an error in the channel causes the received pattern not to match any of the available codewords? In this case we begin by computing the *Hamming distance*, or the number of nonmatching bits between the

received pattern and each of the available codewords. Then, the codeword with the smallest Hamming distance is taken as the received output. For example, assuming that the decoder received the pattern 01111, the Hamming distances are computed as shown in Table 19.11. Taking the smallest Hamming distance, the decoder would choose 01110 as the received codeword and therefore 01 as the decoded bit sequence.

Table 19.10 An error-correcting channel code

Input	Codeword
00	00000
01	01110
10	10011
11	11101

Table 19.11 Hamming distances for received bit pattern of 01111

Input	Codeword	Hamming distance
00	00000	4
01	01110	1
10	10011	3
11	11101	2

EXAMPLE 19.3 Capacity of an AWGN Channel

Problem

For a channel with an available bandwidth of 10 kHz, find the maximum bit rate at which an event sequence can be transmitted if AWGN is present with a noise temperature of 20°C and the transmit power is 1.0 μW/event.

Solution

From the AWGN channel capacity theorem, the channel capacity for the parameters given is:

$$C = 10^4 \log\left[1 + \frac{10^{-6}}{(1.38 \times 10^{-23})(274 + 20) \times 10^4}\right] \approx 345 \text{ kbits/s.}$$

Page 19-24

19.7 ADVANCED TOPICS

One of the technologies that has rapidly expanded the field of digital communication is the cellular telephone. In this case the channel is the atmosphere and all cellular traffic within a geographic area must therefore share the same allocated frequency band, or *multiplex*, often at the same time. In the United States, 1.9 GHz is one such band. The difficulty becomes how to multiplex without causing interference between two phones, or at least how a phone might understand the signal despite the interference.

The answer to these problems is channel sharing, used by wired and wireless systems alike to support multiple simultaneous users. There are several techniques by which the channel can be shared among users to limit interference. Two of the most common ones are frequency-division multiple access and time-division multiple access.

Frequency-Division Multiple Access

As the name suggests, *frequency-division multiple access* (FDMA) transmits multiple signals by separating them in frequency, assuming that the channel bandwidth is much larger than the actual bandwidth required for transmitting the signal. In other words, $B_C \gg B_U$ where the bandwidth of the channel is B_C and the bandwidth of each user is B_U. Under these conditions, we can transmit approximately (B_C/B_U) users on the channel simultaneously. As an example, consider a 3-kHz voice signal on a 100-kHz channel. Then we can theoretically transmit (100/3) ≈ 33 users on this channel.

In practice, a frequency gap called a *guard band* is introduced between each pair of consecutive channels to ensure there is no interference between the two. Each of the signals is individually modulated and upconverted to the frequency of transmission. If the start of the channel's frequency band is f_C, then in our example, the first user would transmit at f_C + 3 kHz with a guard band from f_C + 3 kHz to f_C + 4 kHz. The second user would transmit between f_C + 4 kHz and f_C + 7 kHz and so on. In this way 25 channels actually could be transmitted on the channel. Figure 19.24 illustrates the frequency layout of a typical FDMA system.

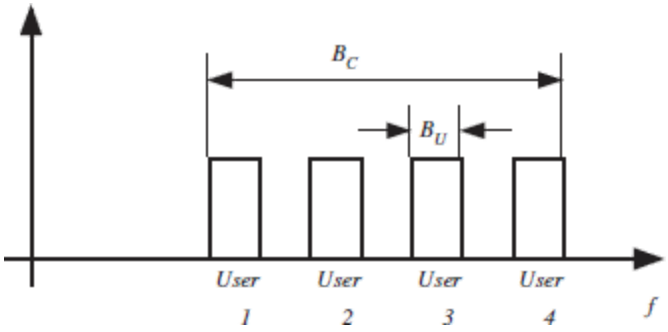


Figure 19.24 Frequency-division multiple access

Examples of systems using FDMA are analog radio, analog television and cable service, and until recently the long distance telephone system. The low cost and high quality of available FDMA equipment, especially that intended for television signals, make it a reasonable choice for many purposes. The downside is that the required bandwidth increases as more and more users are added to the system. Since the channel often has a fixed bandwidth, it is not possible to add users limitlessly.

Time-Division Multiple Access

Another way to achieve the sharing of the medium is through *time-division multiple access* (TDMA). With this approach multiple users are isolated from each other by transmitting at different points in time. Assuming that all users' signals require the same bandwidth, say B_U Hz, we know from the Nyquist theorem that each of them requires the same sampling frequency f_S

$= 2B_U$ and therefore two samples from the same user are $T_S = 1/f_S$ seconds apart. If we make the duration between user samples $T_U = (T_S/N)$, then we can accommodate the samples of all N users in one *time slot* of T_S seconds by sampling each user's signal T_U seconds apart and interleaving the samples of all users. After interleaving, this multiplexed signal is modulated and upconverted to the final transmission frequency. Figure 19.25 illustrates the timing for a typical TDMA system.

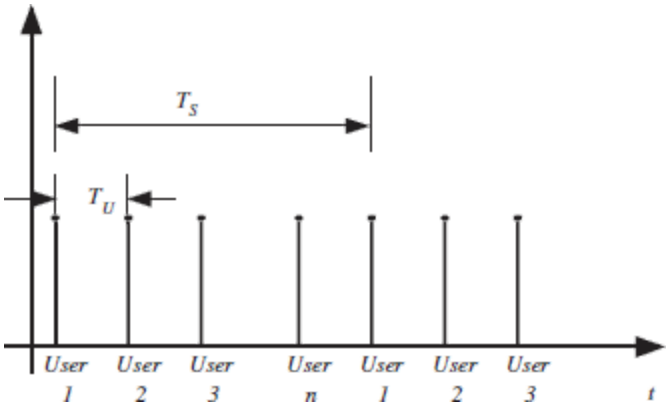


Figure 19.25 Time-division multiple access

The Global System for Mobile Communications (GSM) cellular standard uses TDMA, with individual channels that use 200 kHz of bandwidth for one-way or *simplex* (i.e., mobile to base station or base station to mobile) communication. In all, there are 124 pairs of such simplex channels, where each simplex channel can be used to support eight separate users by TDMA. The total number of users that can be supported by this system is therefore 992.

TDMA is more efficient, easier to operate, less complex, and less expensive than FDMA. However, the duration T_U cannot be chosen arbitrarily. This is because under some conditions the channel may suffer from *dispersion*, meaning that even though a signal of a user is transmitted at a particular time instant, the channel itself may cause it to "spread" in time and spill over to the next user's signal. So T_U must be carefully chosen to ensure that these channel effects do not cause interference between users.

Spread-Spectrum and Code-Division Multiple Access

Spread spectrum had its beginnings during World War II when beautiful Hollywood movie star Hedi Lamarr and talented pianist George Antheil patented a covert system for guiding torpedoes. Their idea was to "hop" the torpedo control signal between many different frequencies to make it difficult for an enemy to jam. The torpedo receiver would be synchronized to the transmitter and know exactly when the next frequency hop would occur. Unfortunately for Lamarr and Antheil, the Page 19-26U.S. Navy was not impressed by their invention, and they never saw a dime for their efforts. Today, their basic spread-spectrum idea is used in millions of mobile phones across the world.

A wireless communications system whose transmission occupies a bandwidth W much greater than the information rate R of the data source is called a *spread-spectrum system*. For example, a voice channel in a cellular IS-95 system has an information rate of R = 9.6 kbits/s, but occupies a bandwidth of W = 1.2288 MHz, a bandwidth expansion of W/R = 128. The ratio W/R is often called the *processing gain* of the system. Third-generation cellular standards use spread-spectrum waveforms that occupy bandwidths of up to 5 MHz and have variable processing gains to enable transmission of low and high data rates.

There are several ways to spread the bandwidth of the data source. Two of the most popular include *frequency-hopped* spread spectrum and *direct-sequence* spread spectrum. Frequency hopping is straightforward: a narrowband transmit signal hops in a pseudorandom sequence from frequency to frequency over a wide bandwidth. The instantaneously occupied bandwidth is small, but on average, the signal bandwidth is large. Direct-sequence spread-spectrum systems use signals whose instantaneous bandwidth is large. One can construct such signals with pseudorandom spreading code sequences called *pseudorandom* (PN) codes.

Since the transmit energy is spread over a wide bandwidth, the power per unit of bandwidth is very low. An enemy with a spectral analyzer may not even know the spread-spectrum signal is present, and even if he did, he would have to know the PN spreading sequence to recover the transmitted information. In addition, the pseudorandom nature of the spread-spectrum

signal makes it difficult for an enemy to jam. The anti-jamming attribute also implies that multiple spread-spectrum users can access the same bandwidth at the same time, where each spread-spectrum user is considered a jammer to the others. When direct-sequence spread spectrum is used for simultaneous access to common bandwidth, this is called *code-division multiple access* (*CDMA*).

Ultrawideband

Ultrawideband (UWB) digital communication is a very recent development using waveforms that distribute their power across a very wide bandwidth. Because UWB is a form of extreme spread spectrum, dozens of devices can operate in the same location at the same time. In addition, the transmitted power per hertz is so low (below the noise floor, in many cases) that data from the devices is extremely difficult for a third party to detect or intercept.

There are currently two design philosophies proposed for UWB communication system standardization. The first is *impulse* radio, using waveforms with subnanosecond duty cycle and occupying several gigahertz of bandwidth. Impulse radio is mainly intended for longer-range low-bit-rate communication systems, though efforts are under way to increase transmission rates. The second proposal is to transmit a *multiband* waveform where each carrier occupies greater than 500 MHz. This approach is intended for shorter-range higher-rate communication systems. Both proposals have their benefits and drawbacks, but it is hoped that a consensus between them can be reached quickly.

Potential applications for UWB radio communication range from military covert communications to commercial indoor wireless links. UWB devices can operate with low power for low-cost short-range communications and may some day wirelessly Page 19-27connect hand-held digital devices with personal computers at rates of up to 1 Gbit/s. Another application for UWB technology is wirelessly streaming multimedia content between components of a home entertainment system. In addition to being used for communications purposes, UWB signals have already been used for applications such as ground-penetrating radar, location sensing, and through-wall imaging.

Because UWB signals occupy several gigahertz of bandwidth, it is inevitable that they will overlap with incumbent narrowband systems such as cellular telephony and the Global Positioning System (GPS). Studies show that the interference is minimal because UWB transmission power per megahertz is so low. Even so, GPS is so critical to many applications that as an extra precaution the FCC has mandated that UWB devices may not emit energy in the GPS band. A large push for commercial applications came after the Federal Communications Commission (FCC) amended its Part 15 rules in February 2002 to allow unlicensed UWB spectral emissions.

19.8 DATA TRANSMISSION IN DIGITAL INSTRUMENTS

A necessary aspect of data acquisition and control systems is the ability to transmit and receive data. Often, a microprocessor-based data acquisition system is interfaced to other digital devices, such as digital instruments or other microprocessors. In these cases it is necessary to transfer data directly in digital form. In fact, it is usually preferable to transmit data that are already in digital form, rather than analog voltages or currents. Among the chief reasons for the choice of digital over analog is that digital data are less sensitive to noise and interference than analog signals: In receiving a binary signal transmitted over a data line, the only decision to be made is whether the value of a bit is 0 or 1. Further, digital data are often coded in such a way that many transmission errors may be detected and corrected. Finally, storage and processing of digital data are accomplished much more readily than with analog signals.

Digital signals in a microprocessor are carried by a bus, consisting of a set of parallel wires, each carrying 1 bit of information. In addition to the signal-carrying wires, there are control lines that determine under what conditions transmission may occur. A typical computer data bus consists of eight parallel wires corresponding to 1 byte; digital data are encoded in binary according to one of a few standard codes, such as the BCD code described in Chapter 11 or the ASCII code (see Appendix D). This bus configuration is usually associated with **parallel transmission,** whereby all the bits are transmitted simultaneously, along with some control bits.

Figure 19.26 depicts the general appearance of a parallel connection. Parallel data transmission can take place in one of two modes: **synchronous** or **asynchronous**. In synchronous transmission, a timing clock pulse is transmitted along with the data over a control line. The arrival of the clock pulse indicates that valid data have also arrived. While parallel synchronous transmission can be very fast, it requires the added complexity of a synchronizing clock and is typically employed only for internal computer data transmission and only over short distances (less than 4 m). Asynchronous data transmission, on the other hand, does not take place at a fixed clock rate, but requires a **handshake protocol** between sending and receiving ends. The handshake protocol consists of the transmission of *data ready* and *acknowledge* signals over two separate control wires. Whenever the sending device is ready to transmit data, it sends a pulse over the *data ready* line. When this signal reaches the receiver, and if the receiver is ready to receive the data, an *acknowledge* pulse is sent back, indicating that the transmission may occur.

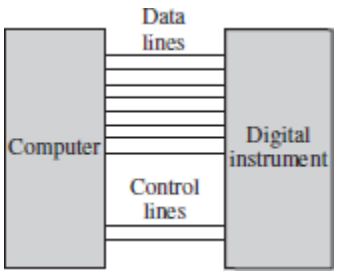


Figure 19.26 Parallel data transmission

Page 19-28

Perhaps the most common parallel interface is the **general-purpose instrument bus** (**GPIB**), which is based on the **IEEE 488 standard**. The hugely popular and widely applied **Universal Serial Bus** (**USB**) and the **Controller Area Network** (**CAN**) are two other important digital communication interfaces.

Parallel Communication and IEEE 488

The IEEE 488 bus, depicted in Figure 19.27, is an 8-bit parallel asynchronous interface that has found common application in digital

instrumentation applications. The physical bus consists of 16 lines, of which 8 are used to carry data, 3 for the handshaking protocol, and the rest to control data flow. The bus permits connection of up to 15 instruments over a 20-m length at data rates up to 1 Mbyte/s. It is possible to extend the operating distance of an IEEE 488 bus beyond the standard 20-m maximum by means of *bus extenders*. The signals transmitted are TTL compatible and employ negative logic, whereby a logic 0 corresponds to a TTL high state (>2 V) and a logic 1 to a TTL low state (<0.8 V). Often, the 8-bit word transmitted over an IEEE 488 bus is coded in ASCII format.

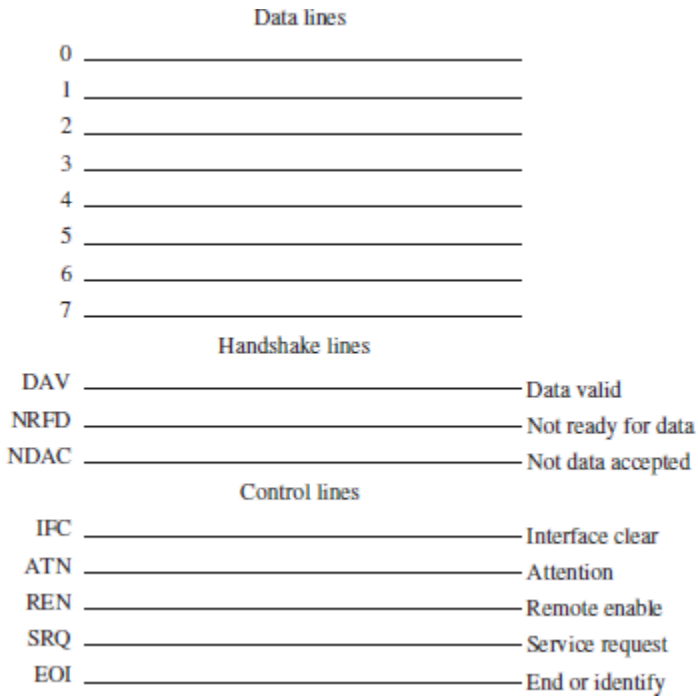


Figure 19.27 IEEE 488 bus

In an IEEE 488 bus system, devices may play different roles and are typically classified as *controllers*, which manage data flow; *talkers* (e.g., a digital voltmeter), which send data; *listeners* (e.g., a printer), which receive data; and *talkers/listeners* (e.g., a digital oscilloscope), which receive and send data. The simplest system configuration might consist of just a talker and a listener. If more than two devices are present on the bus, a controller is necessary to determine when and how data transmission occurs. For example, only one talker can transmit at a time; however, several listeners

may be active on the bus simultaneously. The talker will transmit at the slowest listener data rate so that all data is received correctly.

The **protocol** is the set of rules by which the controller determines the order of talking and listening. One aspect of the protocol is the handshake procedure, which Page 19-29enables the transmission of data. Since different devices (with different data rate capabilities) may be listening to the same talker, the handshake protocol must take into account these different capabilities. The three handshake lines used in the IEEE 488 have important characteristics that give the interface system wide flexibility, allowing interconnection of multiple devices that may operate at different speeds. The slowest active device controls the rate of data transfer, and more than one device can accept data simultaneously. The timing diagram of Figure 19.28 and the list below illustrate the sequence in which the handshake and data transfer are performed.

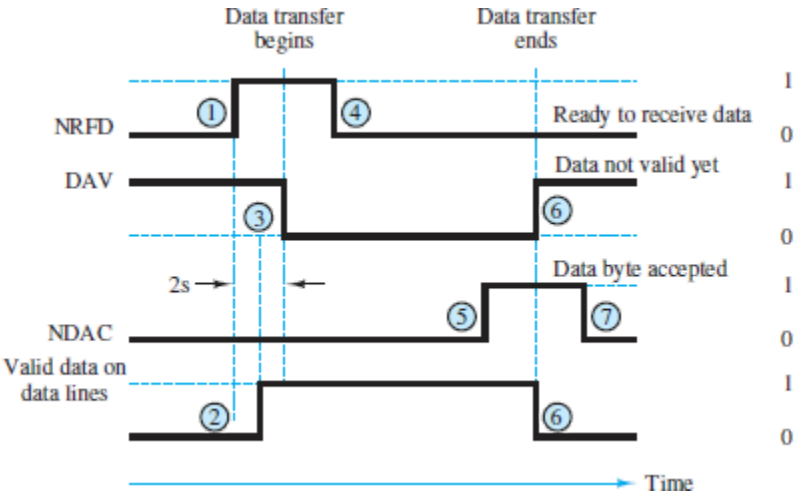


Figure 19.28 IEEE 488 data transmission protocol

1. All active listeners use the *not ready for data* (NRFD) line to indicate their state of readiness to accept a new piece of information. Nonreadiness to accept data is indicated if the NRFD line is held at 0 V. If even one active listener is not ready, the NRFD line of the entire bus is kept at 0 V and the active talker will not transmit the next byte. When all active listeners are ready and they have released the NRFD line, it goes high.

2. The designated talker drives all eight data input/output lines, causing valid data to be placed on them.
3. Two microseconds after putting valid data on the data lines, the active talker pulls the *data valid* (DAV) line to 0 V and thereby signals the active listeners to read the information on the data bus. The 2-μs interval is required to allow the data put on the data lines to reach (settle to) valid logic levels.
4. After the DAV is asserted, the listeners respond by pulling the NRFD line back down to zero. This prevents any additional data transfers from being initiated. The listeners also begin accepting the data byte at their own rates.
5. When a listener has accepted the data, it releases the *not data accepted* (NDAC) line. Only when the last active listener has released its hold on the NDAC line will that line go to its high-voltage-level state.
6. a. When the active talker sees that NDAC has come up to its high state, it stops driving the data line.
   b. At the same time, the talker releases the DAV line, ending the data transfer. The talker may now put the next byte on the data bus.
7. The listeners pull down the NDAC line back to 0 V and put the byte "away."

Each of the instruments present on the data bus is distinguished by its own Page 19-30address, which is known to the controller; thus, the controller determines who the active talkers and listeners are on the bus by *addressing* them. To implement this and other functions, the controller uses the five control lines. Of these, ATN (attention) is used as a switch to indicate whether the controller is addressing or instructing the devices on the bus, or whether data transmission is taking place: when ATN is logic 1, the data lines contain either control information or addresses; with ATN = 1, only the controller is enabled to talk. When ATN = 0, only the devices that have been addressed can use the data lines. The IFC (interface clear) line is used to initialize the bus, or to clear it and reset it to a known condition in case of incorrect transmission. The REN (remote enable) line enables a remote instrument to be controlled by the bus; thus, any function that might normally be performed manually on the instrument (e.g., selecting a range or mode of operation) is now controlled by the bus via the data lines. The SRQ (service request) line is used by instruments on the bus whenever the

instrument is ready to send or receive data; however, it is the controller that decides when to service the request. Finally, the EOI (end or identify) line can be used in two modes: When it is used by a talker, it signifies the end of a message; when it is used by the controller, it serves as a *polling* line, that is, a line used to interrogate the instrument about its data output.

Serial Communication and RS-232

The primary reason why parallel transmission of data is not used exclusively is the limited distance range over which it is possible to transmit data on a parallel bus. Although there are techniques that permit the range to be extended for parallel transmission, these are complex and costly. Therefore, **serial transmission** is frequently used, whenever data are to be transmitted over a significant distance. Since serial data travel along one single path and are transmitted 1 bit at a time, the cabling costs for long distances are relatively low; further, the transmitting and receiving units are also limited to processing just one signal and are also much simpler and less expensive. Two modes of operation exist for serial transmission: **simplex**, which corresponds to transmission in one direction only; and **duplex**, which permits transmission in either direction. Simplex transmission requires only one receiver and one transmitter, at each end of the link; on the other hand, duplex transmission can occur in one of two manners: **half-duplex** and **full-duplex**. In the former, although transmission can take place in both directions, it cannot occur simultaneously in both directions; in the latter case, both ends can simultaneously transmit and receive. Full-duplex transmission is usually implemented by means of four wires.

The data rate of a serial transmission line is measured in bits per second since the data are transmitted 1 bit at a time. The unit of 1 bit/s is 1 **baud**. Like parallel transmission, serial transmission can also occur either synchronously or asynchronously. Asynchronous transmission is less costly but not as fast. A handshake protocol is also required for asynchronous serial transmission. The most popular data-coding scheme for serial transmission is ASCII, consisting of a 7-bit word plus a **parity bit**, for a total of 8 bits per character. The role of the parity bit is to permit error detection in the event of erroneous reception (or transmission) of a bit. In serial asynchronous systems, handshaking is performed by using start and stop bits at the beginning and end of each character. The beginning of a

serial asynchronous word is announced by the *start* bit, which is always a 0 state bit. For the next five to eight successive bit times the line is switched to the 1 and 0 states Page 19-31required to represent the character being sent. Following the last data bit the parity bit is 1 bit or more in the 1 state, indicating "idle." The time period associated with this transmission is called the *stop bit interval*.

If noise pulses affect the transmission line, it is possible that a bit in the transmission could be misread. Thus, following the 5 to 8 data bits, there is a parity bit that is used for error detection. If the transmitter keeps track of the number of 1s in the word being sent, it can send a parity bit, a 1 or a 0, to ensure that the total number of 1s sent is always even (even parity) or odd (odd parity). Similarly, the receiver keeps track of the 1s received to see whether there was a transmission error. If an error is detected, retransmission of the word is requested. No error detection scheme is perfect; robust schemes generally impose significant overhead on the net transmission rate. By contrast, parity checking has low overhead but limited effectiveness.

A useful example of a serial data transmission protocol is the once widely popular **RS-232 standard**. This standard is based on the transmission of voltage pulses at a preselected baud rate; the voltage pulses are in the range –3 to –15 V for a logic 0 and in the range +3 to +15 V for a logic 1. It is important to note that this amounts to a negative logic convention and that the signals are *not* TTL compatible. The distance over which such transmissions can take place is up to approximately 17 m (50 ft). The RS-232 standard was designed to make the transmission of digital data compatible with telephone lines, which were originally designed to carry analog voice signals. The RS-232 standard describes the mechanical and electrical characteristics of the interface between *data terminal equipment* (DTE) and *data communication equipment* (DCE). DTE consists of personal computers, digital instruments, and related peripherals; DCE includes all those devices that are used to encode digital data in a format that permits their transmission over telephone lines. Thus, the standard specifies how data should be presented by the DTE to the DCE so that digital data can be transmitted over analog voice lines.

A typical example of DCE is the **modem**, which stands for *mo*dulate-*dem*odulate. A modem uses digital pulses to modulate a sinusoidal carrier

for transmission and demodulates the transmitted signal to recover the digital pulses at reception. Three methods are commonly used for converting digital pulses to an audio signal: **amplitude-shift keying**, **frequency-shift keying**, and **phase-shift keying**, depending on whether the amplitude, phase, or frequency of the sinusoid is modulated by the digital pulses. Figure 19.29 depicts the essential block of a data transmission system based on the RS-232 standard as well as examples of digital data encoded for transmission over a voice line.

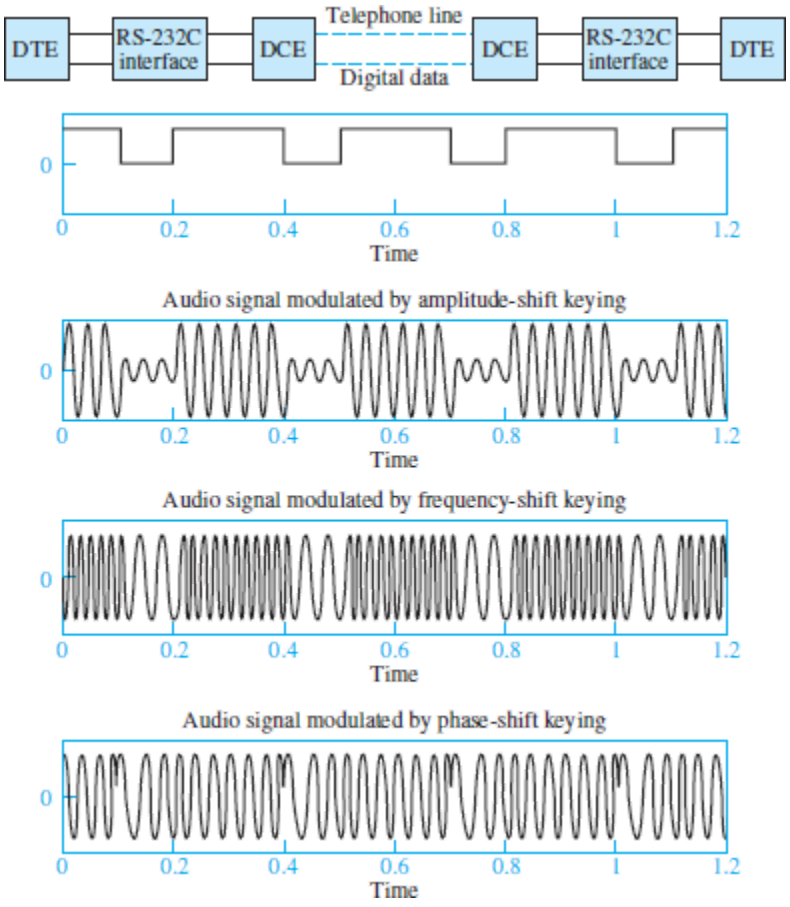


Figure 19.29 Digital data encoded for analog transmission

The Universal Serial Bus

For many devices, the RS-232 standard has been replaced by the USB standard, which was motivated by the need for a more flexible protocol for interfacing peripheral devices to personal computers. The USB technical specification describes the bus attributes, protocol definition, types of

transactions, bus management, and programming interface that are consistent with the requirements of digital devices. The original goals of the USB standard were:

1. Ease of use for PC peripheral expansion.
2. Low-cost solution that supports transfer rates up to 480 Mbits/s.
3. Full support for real-time data for voice, audio, and video.

Page 19-32

4. Protocol flexibility for mixed-mode isochronous data transfers and asynchronous messaging.
5. Integration in commodity device technology (e.g., portable media storage, printers).

The USB standard provides electrical and mechanical specifications for cables and connectors, as well as communication protocols, to accomplish these and other more recent goals. The original USB 1.0 standard, released in January 1996, specified two data transfer rates, low speed (1.5 Mbits/s) and full speed (12 Mbits/s), to accommodate the requirements of different peripheral devices. The USB 2.0 standard, released on April 27, 2000, and named USB High-Speed, increased the maximum data transfer rate by a factor of 40 to 480 Mbits/s. This version of the standard also introduced type-A and type-B connectors. With the advent of smart phones, digital cameras, and other portable digital devices, a version of USB that would allow two devices to communicate without a host PC was needed. In December 2001, the USB On-The-Go (OTG) standard was introduced to allow any two USB devices to be connected directly, with one device acting as the host and the other as the client. A wireless USB standard was introduced in 2005. More recently, the USB 3.0 standard, released on November 12, 2008, and named USB SuperSpeed, further increased the maximum data transfer rate to 5 Gbits/s.

Among the many qualities of the USB standard is the compatibility provided across all three versions of the standard. For instance, any USB 2.0 device can Page 19-33interface with a USB 3.0 device. In fact, in most respects the difference between USB 2.0 and USB 3.0 devices is transparent to the end user. One apparent change in the USB 3.0 standard was the introduction of the Micro-A and -B type connectors that are now commonly

found on many handheld devices such as digital cameras and smart phones. Both the USB 2.0 and 3.0 communication standards are based on the concepts of **endpoints**, **pipes**, and transfer types. However, USB 3.0 added several important new capabilities, including support for **bursting** whereby data **packets** are sent continuously to an endpoint buffer without requiring handshaking between each packet.

USB 3.0 also uses dual-simplex **unicasting** to concurrently send and receive data packets along direct paths from the host to the endpoint. Endpoints receive only those packets addressed to them. Hubs in the network use routing information to direct data packets to the intended endpoint.

Finally, USB 3.0 employs asynchronous notifications, which, along with its unicasting feature, allows a USB 3.0 device to initiate multiple concurrent send transactions on the same bus. These two features also allow a USB 3.0 network to send an inactive device into a low-power sleep state until a notification involving that device is generated. The result is additional power savings compared to a USB 2.0 network, which continuously polls connected devices to determine if the devices are still idle.

Ethernet

Another popular communication protocol is **Ethernet**, which is based on IEEE Standard 802.3, and is commonly used in **local area networks** (LANs) such as office networks. Ethernet was developed at Xerox PARC by Robert Metcalfe and others between 1972 and 1976 and published in 1980. Eventually, Ethernet was standardized in 1985. Many Ethernet installations rely on **twisted-pair** cables to reduce sensitivity to noise and compensate for stray capacitance. However, Ethernet is also implemented in fiber optic networks. Depending upon the physical media, data transmission rates can be as high as 100 Gbits/s. Common installations of Ethernet are known as **10BASE-T**, **100BASE-T**, and **1000BASE-T**, where the numeric prefix indicates its nominal speed in Mbits/s.

Control Area Network

Yet another popular bus architecture is the control area network, or CAN, protocol. CAN is a message-oriented transmission protocol used in *broadcast communication*. Each transmission contains a message identifier that is unique within the whole network and defines the content and priority of the message. This is important when several *stations* are competing for bus access. The CAN protocol supports two message frame formats, the only essential difference being the length of the identifier. The *CAN standard frame*, also known as **CAN 2.0 A**, supports a length of 11 bits for the identifier; and the *CAN extended frame*, also known as **CAN 2.0 B**, supports a length of 29 bits for the identifier. The structure of a CAN message is shown in Figure 19.30.

Figure 19.30 Structure of CAN message

Page 19-34

It is very easy to add stations to an existing CAN network without making any hardware or software modifications to the existing stations as long as the new stations are purely receivers. This allows the concept of modular electronics and also permits multiple reception and the synchronization of distributed processes: data needed as information by several stations can be transmitted via the network in such a way that it is unnecessary for each station to have to know who is the producer of the data. This makes it easy to service and upgrade networks, as data transmission is not based on the availability of specific types of stations.

CAN was originally developed for passenger car applications. CAN networks used in engine management connect several electronic control units (ECUs). In addition, some passenger cars are equipped with CAN-based multiplex systems connecting body ECUs. These multiplex networks link door and roof control units as well as lighting control units and seat control units. In some passenger cars, a CAN-based diagnostic interface is implemented. The different CAN-based in-vehicle networks are connected via gateways. In many system designs, the gateway functionality is implemented in the dashboard. In the future, the dashboard itself may use a

local CAN network to connect the different display and control units. The following microcontrollers, commonly used in automotive applications, support CAN interfaces: Infineon C16XC series, Motorola 683xx series and MPC5xx series, and Hitachi SH-2. Because of its widespread use in automobiles, the popularity of the CAN bus has grown significantly in recent years, and CAN has become a de facto standard for the automotive industry.

EXAMPLE 19.4 ASCII to Binary Data Conversion over IEEE 488 Bus

Problem

Determine the actual binary data sent by a digital voltmeter over an IEEE 488 bus.

Solution

Known Quantities: Digital voltmeter reading V.

Find: Binary data sequence.

Schematics, Diagrams, Circuits, and Given Data: V = 3.405 V. ASCII conversion table.

Assumptions: Data are encoded in ASCII format. Sequence is sent from most to least significant digit.

Analysis: Using an ASCII conversion table:

Control character	ASCII (hex)
3	33
.	2E
4	34
0	30
5	35

The actual binary data sent can therefore be determined by converting the hexadecimal ASCII sequence into binary data:

33 2E 34 30 35 ↔ 001100110010111000110100001100000011010 1

Comments: Note that the ASCII format is not very efficient; if you directly performed a base-10 to binary conversion only 8 bits (plus the decimal point) would be required.

CHECK YOUR UNDERSTANDING

Determine the actual binary data sent by a digital voltmeter reading of 17.06 V over an IEEE 488 bus if the data are encoded in ASCII format. Assume that the sequence is from most to least significant digit.

Answer: 31 37 2E 30 36 ↔ 00110011 00110111 00101110
00110000 00110110

Conclusion

1. Digital systems offer powerful capabilities for reliably transmitting coded information by minimizing error associated with noise. Effective digital communication systems rely on statistical methods to manage the probabilistic nature of many sources of noise.
2. Pulse-code modulation is the process by which an analog signal is sampled and quantized into a discrete sequence of digital data.
3. Source coding is the process of removing redundant and/or unnecessary information in a signal so as to produce the minimum signal necessary to communicate a message. Data compression

schemes, such as Huffman codes and Lempel-Ziv-Welch codes, are examples of source coding techniques.

4. Digital baseband modulation refers to various methods for embedding digital information in an analog signal prior to transmission through a communication channel.
5. Communication channels have theoretical limits on the rate at which data may be transmitted error-free through a channel. The limit for a particular channel is known as its channel capacity. Channel coding techniques are designed to produce transmission rates that approach the channel capacity.
6. Digital communication standards, such as IEEE 488, RS232, and USB, provide rules and guidelines for developing parallel and series communication devices.

HOMEWORK PROBLEMS

Section 19.2: Introductory Probability

19.1 The Q-function is often used in the communications literature to express the probability of error. Unfortunately, Matlab® does not have a built-in Q-function; however, the built-in *complementary error function* erfc(x) can be used to obtain the Q-function:

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2}\, dt$$

Obtain an expression for $Q(x)$ in terms of erfc(x).

Page 19-36

Section 19.3: Pulse-Code Modulation

19.2 Consider that you have been assigned the task of quantizing for transmission a voice signal whose bandwidth is 3.4 kHz. You are constrained by a channel that can support a maximum bit rate of 35,000 bits/s. What is the maximum number of bits that you can

employ for quantization? Also find the corresponding sampling rate of the signal.

Section 19.4: Source Coding

19.3 A remote color sensor observes a color manufacturing process and transmits observations to a control center. The observed colors occur with the following probabilities:

$$P(\text{Red}) = 3/4$$
$$P(\text{Green}) = 1/8$$
$$P(\text{Blue}) = 1/16$$
$$P(\text{Yellow}) = 1/16$$

A computer at the sensor encodes the color observations according to the following code:

$$\text{Red} \rightarrow 00$$
$$\text{Green} \rightarrow 01$$
$$\text{Blue} \rightarrow 10$$
$$\text{Yellow} \rightarrow 11$$

These codewords are then sent to the control center (leftmost bit first).

a. What is the average length of the code?

b. The sequence (00000001000000110000000000000 0100) was received at the control center with the leftmost bit received first. What is the observed sequence of colors?

c. Notice that the observed sequence from part b is highly redundant. To reduce this redundancy, devise a code that has an average length less than two. What is the average length of your code?

d. Encode the following sequence of color observations (leftmost observed first) with your code from part c: {Red, Red, Red, Green, Red, Red, Red, Yellow, Red, Red, Red, Red, Red, Red, Green, Red}. Do you see an improvement?

19.4 In a musical CD recording, each of the two stereo signals is sampled with a 16-bit ADC at 44.1 kHz.

a. What is the ratio of the output signal to the quantization noise for a sinusoidal input signal?

b. The music bit stream is appended with error-correction bits, clock extraction bits, and display and control bits. Assume that these bits correspond to an overhead of 50 percent. What is the playback bit rate of a CD?

c. Assume a CD records an hour of music. Determine the number of bits recorded on a CD.

d. Now let's compare data storage to the storage of music on a CD. Consider a history textbook that contains 1,000 pages, 50 lines per page, 15 words per line, 6 letters per word, and 7 bits per word on average. Determine the number of bits required to digitally store this textbook. Now estimate the number of such textbooks that can be stored on a CD.

19.5 A sinusoidal signal with an amplitude of 2 is to be quantized such that the ratio of the signal to quantization noise is at least 10 dB. What is the minimum number of bits required to perform this quantization?

19.6 High definition TVs use a format in which there are 1,920 × 1,080 pixels on the TV screen. Each of these pixels has 16 different brightness levels, and pictures are repeated at the rate of 30 frames per second. Assuming that all brightness levels are equally likely to occur, calculate the average rate of information conveyed by the TV. [*Hint:* If the rate at which source X emits symbols is r symbols/s, the information rate of the source is given by $R = rH(X)$ bits/s. You can think of the TV screen as a source and all the brightness levels of the pixels as the set of possible outcomes of the source.]

Section 19.5: Digital Baseband Modulation

19.7 Suppose a cable modem communications standard specifies a baud rate of one million message waveform transmissions per second. Further, suppose there are 256 waveforms in the message waveform set. What is the bit rate?

19.8 Referring to the waveforms of Figure P19.8:

a. Compute the average energy of each of the three message waveform sets. Note that for all sets the duty cycle is $T = 1$.

b. Compute the average RMS distance between the waveforms in each set.

c. Which waveforms would you choose for binary signaling? Why?

Page 19-37

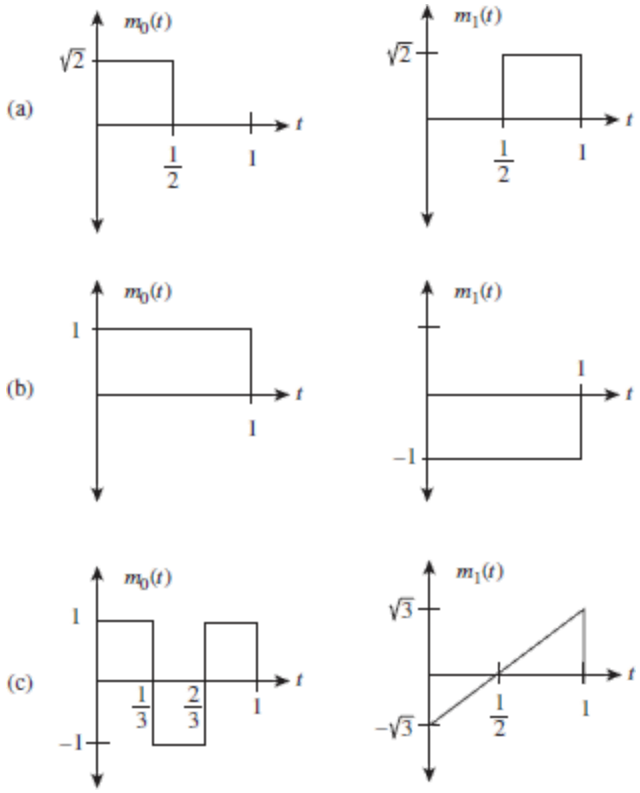


Figure P19.8 Three sets of binary message waveforms

Section 19.6: Channel Coding

19.9 It can be shown that the theoretical probability of error for a digital communication system is given by

$$P_e = Q\sqrt{\frac{2E_b}{kT}}$$

where $Q(x)$ is the Q-function, E_b is the energy per bit, and T is the noise temperature of the additive white gaussian noise. We would like to verify this formula by performing the following experiment:

a. Generate a vector of equiprobable random bits {0, 1}.

b. Map each bit to a transmitted symbol in the following manner:

$$0 \rightarrow \sqrt{E_b}$$
$$1 \rightarrow -\sqrt{E_b}$$

where E_b is the energy of the symbol.

c. Add gaussian noise with zero mean and variance $kT/2$ to the transmitted symbol.

d. Perform detection by checking the sign of the received signal. If the sign is greater than zero, then decide that a 0 was sent; otherwise, decide that a 1 was sent.

e. Count the number of errors made.

f. Repeat until you have counted at least 100 errors. (The more errors you count, the closer will be your estimate of the BER to the theoretical results.)

The estimated probability of error is

$$P_e \approx \frac{\text{number of errors counted}}{\text{number of symbols generated}}$$

Do steps a to f for signal-to-noise ratios (SNRs) of E_b/kT = 0, 2,..., 10 dB. Once you are done collecting the estimated probabilities of error, plot your results on a log scale versus SNR in decibels, i.e., 10 $\log_{10}(E_b/N_o)$. Show each data point as an X mark. Also plot the theoretical probability of error P_e for comparison. Check your results with Figure P19.9.

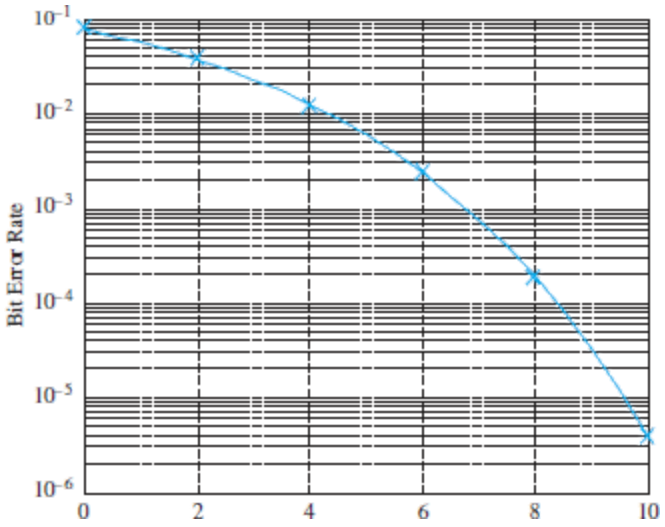


Figure P19.9 Theoretical bit error rate

19.10 You have been given the task of designing a point-to-point wireless communications link between two buildings spaced 2 km apart. Tests show that the signal-to-noise ratio P/kT at the receiver is 20 dB.

a. What is the minimum amount of bandwidth needed to support a 1 Mbit/s link?

b. Suppose you have a bandwidth limitation of 5 MHz. Is the SNR sufficient to support a data rate of 5 Mbits/s?

19.11 Your task is to transmit a voice signal of bandwidth 3.4 kHz over an AWGN channel. Assume Page 19-38the signal is sampled at 1.5 times the Nyquist rate and each sample is quantized into one of 256 equally likely levels.

a. What is the rate at which this voice source is generating information?

b. Can you transmit the output of this source without error over an AWGN channel with a bandwidth of 10 kHz and an SNR of 20 dB? (*Hint:* The SNR is P/kT.)

c. What is the minimum SNR required for error-free transmission of this information over the channel?

d. Determine the minimum AWGN channel bandwidth required for error-free transmission of the output of this source if the SNR is 20 dB.

19.12 A binary symmetric channel (BSC) has two inputs $x_1 = 0$ and $x_2 = 1$ and two outputs $y_1 = 0$ and $y_2 = 1$. The channel is symmetric because the probability of receiving a 1 if a 0 is transmitted is the same as the probability of receiving a 0 if a 1 is sent. This common transition probability is denoted as b. Consider a simple repetition coding scheme over the BSC in which each bit is repeated n times where $n = 2m + 1$ is an odd integer. For decoding the received bits, a majority rule is employed. In other words, if in a block of n received bits the number of 0s is greater than the number of 1s, the decoder will decide that a 0 was transmitted. Otherwise, the decoder will decide that a 1 was transmitted. Therefore an error occurs whenever $m + 1$ or more bits out of n bits are received incorrectly.

a. For this coding-decoding scheme, derive an expression for the probability of bit error P_e.

b. Calculate P_e when $b = 0.05$ and $n = 3, 5, 7$.

19.13 Figure P19.13 compares the *bit error rate* (BER) performance of BPSK transmission over an AWGN channel with and without error-control coding.

a. At a BER of 10^{-5} what is the SNR required to transmit bits over the channel without coding?

b. What is the SNR required to obtain BER 10^{-5} over the channel with coding employed?

c. The *coding gain* is defined as the difference between the SNRs required to obtain a certain BER using an uncoded and a coded system. The coding gain represents the reduction in signal power when using coding to achieve the same level of performance as that of a scheme without coding. What is the coding gain of this coding scheme at a BER 10^{-5}?

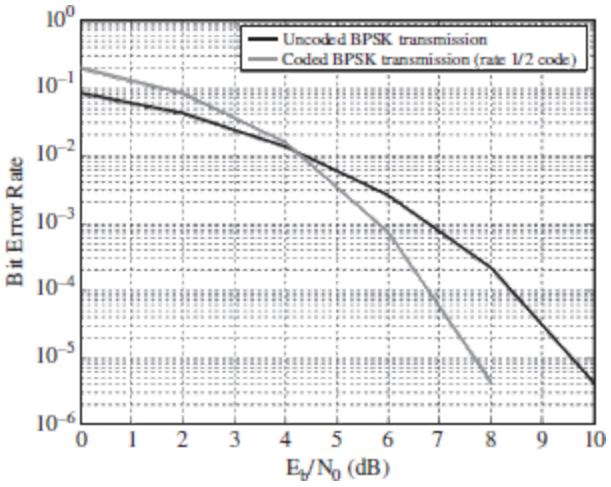


Figure P19.13

19.14 Your task is to transmit data from one computer to another through an AWGN channel. For this purpose you employ a bipolar binary signal that is a +1 V or a –1 V pulse during an interval [0, T] depending on whether the information bit is a 0 or a 1. The power spectral density of the noise added to the signal in the channel is $kT/2 = 10^{-3}$ W/Hz. Since you are transmitting important messages over this system, you require that you obtain at most 100 errors in every 125 Kbyte file. Determine the maximum rate at which you can transmit bits over this channel. [*Hint*: The average probability of error of this system or BER is (number of bits in error)/(total number of bits sent). Compute the energy per bit E_b. Then the average probability of error is given by $P_e = Q\sqrt{2E_b/kT}$.]

Section 19.7: Advanced Topics

19.15 In this problem, we study a T1 carrier system used in digital telephony. Voice signals are usually filtered using a low-pass filter of cut-off frequency 3.4 kHz and then sampled at 8 kHz. This system multiplexes samples from 24 voice signals in a single sampling period. Voice signals are coded with 8-bit PCM, and each frame (one sampling period) consists of 24 samples plus a single bit added for synchronization.

a. Calculate the duration of each bit. Note that this system multiplexes 24 samples each with 8 bits plus a single bit onto the

time duration between two consecutive samples $T_s = 1/f_s$.

b. Calculate the resulting transmission rate.

Page 19-39

Section 19.8: Data Transmission in Digital Instruments

19.16 An ASCII (hex) encoded message is given below. Decode the message.

41 53 43 49 49 20 64 65 63 6F 64 69 6E 67 20 69 73 20 65
61 73 79 21

19.17 An ASCII (binary) encoded message is given below. Decode the message. (*Hint:* Follow a line-by-line sequence, not column-by-column.)

1010100 1101000 1101001 1110011
1101001 1101101 1100101 0101101
1101110 1100111 0100000 1110000

0100000 1101001 1110011 0100000
1100011 1101111 1101110 1110011
1110010 1101111 1100010 1101100

1100001 0100000 1110100
1110101 1101101 1101001
1100101 1101101 0101110

19.18 Express the following decimal numbers in ASCII form:

a. 12

b. 345.2

c. 43.5

19.19 Express the following words in ASCII form:

a. Digital

b. Computer

c. Ascii

d. ASCII

19.20 Explain why data transmission over long distances is usually done via a serial scheme rather than a parallel one.

19.21 A certain automated data-logging instrument has 16K words of on-board memory. The device samples the variable of interest once every 5 min. How often must data be downloaded and the memory cleared to avoid losing any data?

19.22 Explain why three wires are required for the handshaking technique employed by IEEE 488 bus systems.

19.23 A CD-ROM can hold 650 Mbytes of information. Suppose the CD-ROMs are packed 50 per box. The manufacturer ships 100 boxes via commercial airliner from Los Angeles to New York. The distance between the two cities is 2,500 mi by air, and the airliner flies at a speed of 400 mi/h. What is the data transmission rate between the two cities in bits per second?

Design Credits: Mini DVI cable adapter isolated on white background: Robert Lehmann/Alamy Stock Photo; Balance scale: Alex Slobodkin/E+/Getty Images; Icon for “Focus on measurements” weighing scales: Media Bakery.

A P P E N D I X

A

LINEAR ALGEBRA AND COMPLEX NUMBERS

A.1 SOLVING SIMULTANEOUS LINEAR EQUATIONS, CRAMER'S RULE, AND MATRIX EQUATION

The solution of simultaneous equations, such as those that are often seen in circuit theory, may be obtained relatively easily by using Cramer's rule. This method applies to 2 × 2 or larger systems of equations. Cramer's rule requires the use of the concept of determinant. Linear, or matrix, algebra is valuable because it is systematic, general, and useful in solving complicated problems. A determinant is a scalar defined on a square array of numbers, or matrix, such as

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \tag{A.1}$$

In this case the matrix is a 2 × 2 array with two rows and two columns, and its determinant is defined as

$$\det = a_{11}a_{22} - a_{12}a_{21} \tag{A.2}$$

Page 936

A third-order, or 3 × 3, determinant such as

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{A.3}$$

is given by

$$\begin{aligned} \det = {} & a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\ & + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned} \tag{A.4}$$

For higher-order determinants, you may refer to a linear algebra book. To illustrate Cramer's method, a set of two equations in general form will be solved here. A set of two linear simultaneous algebraic equations in two unknowns can be written in the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{A.5}$$

where x_1 and x_2 are the two unknowns. The coefficients a_{11}, a_{12}, a_{21}, and a_{22} are known quantities. The two quantities on the right-hand sides, b_1 and b_2, are also known (these are typically the source currents and voltages in a circuit problem). The set of equations can be arranged in matrix form, as shown in equation A.6.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{A.6}$$

In equation A.6, a coefficient matrix multiplied by a vector of unknown variables is equated to a right-hand-side vector. Cramer's rule can then be applied to find x_1 and x_2, using the following formulas:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \tag{A.7}$$

Thus, the solution is given by the ratio of two determinants: the denominator is the determinant of the coefficient matrix, while the numerator is the determinant of the same matrix with the right-hand-side vector ($[b_1 \ b_2]^T$ in this case) substituted in place of the column of the coefficient matrix corresponding to the desired variable (i.e., first column for x_1, second column for x_2, etc.). In a circuit analysis problem, the coefficient matrix is formed by the resistance (or conductance) values, the vector of unknowns is composed of the mesh currents (or node voltages), and the right-hand-side vector contains the source currents or voltages.

In practice, many calculations involve solving higher-order systems of linear equations. Therefore, a variety of computer software packages are often used to solve higher-order systems of linear equations.

CHECK YOUR UNDERSTANDING

A.1 Use Cramer's rule to solve the system

$$\begin{aligned} 5v_1 + 4v_2 &= 6 \\ 3v_1 + 2v_2 &= 4 \end{aligned}$$

Page 937

A.2 Use Cramer's rule to solve the system

$$\begin{aligned} i_1 + 2i_2 + i_3 &= 6 \\ i_1 + i_2 - 2i_3 &= 1 \\ i_1 - i_2 + i_3 &= 0 \end{aligned}$$

A.3 Convert the following system of linear equations into a matrix equation as shown in equation A.6, and find matrices A and b.

$$
\begin{aligned}
2i_1 - 2i_2 + 3i_3 &= -10 \\
-3i_1 + 3i_2 - 2i_3 + i_4 &= -2 \\
5i_1 - i_2 + 4i_3 - 4i_4 &= 4 \\
i_1 - 4i_2 + i_3 + 2i_4 &= 0
\end{aligned}
$$

Answer: A1: $v_1 = 2,\ v_2 = -1$;. A2: $i_1 = 1,\ i_2 = 2,\ i_3 = 1$;. A3:

$$A = \begin{bmatrix} 1 & -4 & 1 & 2 \\ 5 & -1 & 4 & -4 \\ -5 & 3 & -2 & 1 \\ 1 & -2 & 3 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 4 \\ -2 \\ -10 \end{bmatrix}$$

A.2 INTRODUCTION TO COMPLEX ALGEBRA

From your earliest training in arithmetic, you have dealt with real numbers such as $4, -2, \frac{5}{9}, \pi, e$, etc., which may be used to measure distances in one direction or another from a fixed point. However, a number that satisfies the equation:

$$x^2 + 9 = \tag{A.8}$$

is not a real number. Imaginary numbers were introduced to solve equations such as equation A.8. Imaginary numbers add a new dimension to our number system. To deal with imaginary numbers, a new element, j, is added to the number system having the property:

$$j^2 = -1 \tag{A.9}$$

or

$$j = \sqrt{-1}$$

Thus, we have $j^3 = -j$, $j^4 = 1$, $j^5 = j$, etc. Using equation A.9, you can see that the solutions to equation A.8 are $\pm j3$. In mathematics, the symbol i is

used for the imaginary unit, but this might be confused with current in electrical engineering. Therefore, the symbol j is used in this book.

A complex number (indicated in boldface notation) is an expression of the form:

$$\mathbf{A} = a + jb \tag{A.10}$$

where a and b are real numbers. The complex number $\mathbf{A}$ has a real part a and an imaginary part b, which can be expressed as

$$\begin{aligned} a &= \text{Re } \mathbf{A} \\ b &= \text{Im } \mathbf{A} \end{aligned} \tag{A.11}$$

It is important to note that a and b are both real numbers. The complex number $a + jb$ can be represented on a rectangular coordinate plane, called the *complex plane*, by interpreting it as a point (a, b). That is, the horizontal coordinate is a on the real axis, and the vertical coordinate is b on the imaginary axis, as shown in Figure A.1. The complex number $\mathbf{A} = a + jb$ can also be Page 938uniquely located in the complex plane by specifying the distance r along a straight line from the origin and the angle θ, which this line makes with the real axis, as shown in Figure A.1. From the right triangle of Figure A.1, we can see that:

$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \theta &= \tan^{-1}\left(\frac{b}{a}\right) \\ a &= r\cos\theta \\ b &= r\sin\theta \end{aligned} \tag{A.12}$$

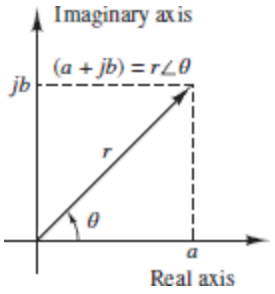


Figure A.1 Polar form representation of complex numbers

Then we can represent a complex number by the expression:

$$\mathbf{A} = re^{j\theta} = r\angle\theta \tag{A.13}$$

which is called the polar form of the complex number. The number r is called the magnitude (or amplitude), and the number θ is called the angle (or argument). The two numbers are usually denoted by $r = |\mathbf{A}|$ and $\theta = \arg \mathbf{A} = \angle\mathbf{A}$.

Given a complex number $\mathbf{A} = a + jb$, the *complex conjugate* of $\mathbf{A}$, denoted by the symbol $\mathbf{A}^*$, is defined by the following equalities:

$$\begin{aligned} \text{Re } \mathbf{A}^* &= \text{Re } \mathbf{A} \\ \text{Im } \mathbf{A}^* &= -\text{Im } \mathbf{A} \end{aligned} \tag{A.14}$$

That is, the sign of the imaginary part is reversed in the complex conjugate.

Finally, two complex numbers are equal *if and only if* the real parts are equal and the imaginary parts are equal, which is equivalent to stating that two complex numbers are equal only if their magnitudes are equal and their arguments are equal.

The following examples and exercises should help clarify these explanations.

EXAMPLE A.1

Convert the complex number $\mathbf{A} = 3 + j4$ to its polar form.

Solution

$$r = \sqrt{3^2 + 4^2} = 5 \qquad \theta = \tan^{-1}\left(\frac{4}{3}\right) = 53.13^\circ$$

$$\mathbf{A} = 5\angle 53.13^\circ$$

EXAMPLE A.2

Convert the number $\mathbf{A} = 4\angle(-60^\circ)$ to its complex form.

Solution

$$a = 4\cos(-60°) = 4\cos(60°) = 2$$
$$b = 4\sin(-60°) = -4\sin(60°) = -2\sqrt{3}$$

Thus, $\mathbf{A} = 2 - j2\sqrt{3}$

Page 939

Addition and *subtraction* of complex numbers are governed by the following rules:

$$\begin{aligned}(a_1 + jb_1) + (a_2 + jb_2) &= (a_1 + a_2) + j(b_1 + b_2)\\ (a_1 + jb_1) - (a_2 + jb_2) &= (a_1 - a_2) + j(b_1 - b_2)\end{aligned} \tag{A.15}$$

Multiplication of complex numbers in polar form follows the law of exponents. That is, the magnitude of the product is the product of the individual magnitudes, and the angle of the product is the sum of the individual angles, as shown below.

$$\mathbf{AB} = (Ae^{j\theta})(Be^{j\phi}) = ABe^{j(\theta+\phi)} = AB\angle(\theta + \phi) \tag{A.16}$$

If the numbers are given in rectangular form and the product is desired in rectangular form, it may be more convenient to perform the multiplication directly, using the rule that $j^2 = -1$, as illustrated in equation A.17.

$$\begin{aligned}(a_1 + jb_1)(a_2 + jb_2) &= a_1a_2 + ja_1b_2 + ja_2b_1 + j^2b_1b_2\\ &= (a_1a_2 + j^2b_1b_2) + j(a_1b_2 + a_2b_1)\\ &= (a_1a_2 - b_1b_2) + j(a_1b_2 + a_2b_1)\end{aligned} \tag{A.17}$$

Division of complex numbers in polar form follows the law of exponents. That is, the magnitude of the quotient is the quotient of the magnitudes, and the angle of the quotient is the difference of the angles, as shown in equation A.18.

$$\frac{\mathbf{A}}{\mathbf{B}} = \frac{Ae^{j\theta}}{Be^{j\phi}} = \frac{A\angle\theta}{B\angle\phi} = \frac{A}{B}\angle(\theta - \phi) \tag{A.18}$$

Division in the rectangular form can be accomplished by multiplying the numerator and denominator by the complex conjugate of the denominator. Multiplying the denominator by its complex conjugate converts the denominator to a real number and simplifies division. This is shown in Example A.4. Powers and roots of a complex number in polar form follow the laws of exponents, as shown in equations A.19 and A.20.

$$\mathbf{A}^n = (Ae^{j\theta})^n = A^n e^{jn\theta} = A^n \angle n\theta \qquad \text{(A.19)}$$
$$\mathbf{A}^{1/n} = (Ae^{j\theta})^{1/n} = A^{1/n} e^{j1/n\theta}$$

$$= \sqrt[n]{A}\angle\left(\frac{\theta + k2\pi}{n}\right) \qquad k = 0, \pm 1, \pm 2, \ldots \qquad \text{(A.20)}$$

EXAMPLE A.3

Perform the following operations, given that $\mathbf{A} = 2 + j3$ and $\mathbf{B} = 5 - j4$.

(a) $\mathbf{A} + \mathbf{B}$ (b) $\mathbf{A} - \mathbf{B}$ (c) $2\mathbf{A} + 3\mathbf{B}$

Solution

$$\mathbf{A} + \mathbf{B} = (2 + 5) + j[3 + (-4)] = 7 - j$$
$$\mathbf{A} - \mathbf{B} = (2 - 5) + j[3 - (-4)] = -3 + j7$$

For part c, $2\mathbf{A} = 4 + j6$ and $3\mathbf{B} = 15 - j12$. Thus, $2\mathbf{A} + 3\mathbf{B} = (4 + 15) + j[6 + (-12)] = 19 - j6$

Page 940

EXAMPLE A.4

Perform the following operations in both rectangular and polar form, given that $\mathbf{A} = 3 + j3$ and $\mathbf{A} = 3 + j3$

(a) $\mathbf{AB}$ (b) $\mathbf{A} \div \mathbf{B}$

Solution

(a) In rectangular form:

$$\begin{aligned}\mathbf{AB} &= (3+j3)(1+j\sqrt{3}) = 3+j3\sqrt{3}+j3+j^2 3\sqrt{3} \\ &= (3+j^2 3\sqrt{3})+j(3+3\sqrt{3}) \\ &= (3-3\sqrt{3})+j(3+3\sqrt{3})\end{aligned}$$

To obtain the answer in polar form, we need to convert **A** and **B** to their polar forms:

$$\begin{aligned}\mathbf{A} &= 3\sqrt{2}e^{j45^\circ} = 3\sqrt{2}\angle 45^\circ \\ \mathbf{B} &= \sqrt{4}e^{j60^\circ} = 2\angle 60^\circ\end{aligned}$$

Then

$$\mathbf{AB} = (3\sqrt{2}e^{j45^\circ})\sqrt{4}e^{j60^\circ} = 6\sqrt{2}\angle 105^\circ$$

(b) To find **A** ÷ **B** in rectangular form, we can multiply **A** and **B** by **B***.

$$\frac{\mathbf{A}}{\mathbf{B}} = \frac{3+j3}{1+j\sqrt{3}}\frac{1-j\sqrt{3}}{1-j\sqrt{3}}$$

Then

$$\frac{\mathbf{A}}{\mathbf{B}} = \frac{(3+3\sqrt{3})+j(3-3\sqrt{3})}{4}$$

In polar form, the same operation may be performed as follows:

$$\frac{\mathbf{A}}{\mathbf{B}} = \frac{3\sqrt{2}\angle 45^\circ}{2\angle 60^\circ} = \frac{3\sqrt{2}}{2}\angle(45^\circ - 60^\circ) = \frac{3\sqrt{2}}{2}\angle(-15^\circ)$$

Euler's Identity

Euler's formula extends the usual definition of the exponential function to allow for complex numbers as arguments.

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{A.21}$$

All the standard trigonometry formulas in the complex plane are direct consequences of Euler's formula. The two important formulas are

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2} \qquad \sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{2j} \tag{A.22}$$

Page 941

EXAMPLE A.5

Using Euler's formula, show that

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2}$$

Solution

Using Euler's formula gives

$$e^{j\theta} = \cos\theta + j\sin\theta$$

Extending the above formula, we can obtain

$$e^{-j\theta} = \cos(-\theta) + j\sin(-\theta) = \cos\theta - j\sin\theta$$

Thus,

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2}$$

CHECK YOUR UNDERSTANDING

A.4 In a certain AC circuit, $\boldsymbol{V} = \boldsymbol{IZ}$, where $\boldsymbol{Z} = 7.75\angle 90°$ and $\boldsymbol{I} = 2\angle -45°$. Find $\boldsymbol{V}$.

A.5 In a certain AC circuit, $\boldsymbol{V} = \boldsymbol{IZ}$, where $\boldsymbol{Z} = 5\angle 82°$ and $\boldsymbol{V} = 30\angle 45°$. Find $\boldsymbol{I}$.

A.6 Show that the polar form of $\boldsymbol{AB}$ in Example A.4 is equivalent to its rectangular form.

A.7 Show that the polar form of $\boldsymbol{A} \div \boldsymbol{B}$ in Example A.4 is equivalent to its rectangular form.

A.8 Using Euler's formula, show that $\sin\theta = (e^{j\theta} - e^{-j\theta})/2j$.

Answer: A4: $\boldsymbol{V} = 15\angle 45°$; A5: $\boldsymbol{I} = 6\angle(-37°)$

A P P E N D I X

B

THE LAPLACE TRANSFORM

The transient analysis methods illustrated in Chapter 4 for first- and second-order circuits can become rather cumbersome when applied to higher-order circuits. Moreover, solving the differential equations directly does not reveal the strong connection that exists between the transient response and the frequency response of a circuit. The aim of this appendix is to introduce an alternate solution method based on the concepts of complex frequency and of the **Laplace transform**. The concepts presented will demonstrate that the frequency response of linear circuits is but a special case of the general transient response of the circuit, when analyzed by means of Laplace methods. In addition, the use of the Laplace transform method reveals *systems* concepts, such as poles, zeros, and transfer functions.

B.1 COMPLEX FREQUENCY

In Chapter 3, we considered circuits with sinusoidal excitations such as

$$v(t) = A\cos(\omega t + \phi) \tag{B.1}$$

which we also wrote in the equivalent phasor form:

$$\mathbf{V}(j\omega) = A\,e^{j\phi} = A\angle\phi \tag{B.2}$$

The two expressions just given are related by

$$v(t) = \mathrm{Re}(\mathbf{V}\,e^{j\omega t}) \tag{B.3}$$

As was shown in Chapter 3, phasor notation is extremely useful in solving AC steady-state circuits, in which the voltages and currents are *steady-state sinusoids*. We now consider a different class of waveforms, useful in the transient analysis of circuits, namely, *damped sinusoids*. The most general form of a damped sinusoid is

$$v(t) = A\,e^{\sigma t}\cos(\omega t + \tag{B.4}$$

As one can see, a damped sinusoid is a sinusoid multiplied by a real exponential $e^{\sigma t}$. The constant σ is real and is usually zero or negative in most practical circuits. Figure B.1(a) and (b) depict the case of a damped sinusoid with negative σ and with positive σ, respectively. Note that the case of $\sigma = 0$ corresponds exactly to a sinusoidal waveform. The definition of phasor voltages and currents given in Chapter 3 can easily be extended to account for the case of damped sinusoidal waveforms by defining a new variable s, called the *complex frequency:*

$$s = \sigma + j\omega \tag{B.5}$$

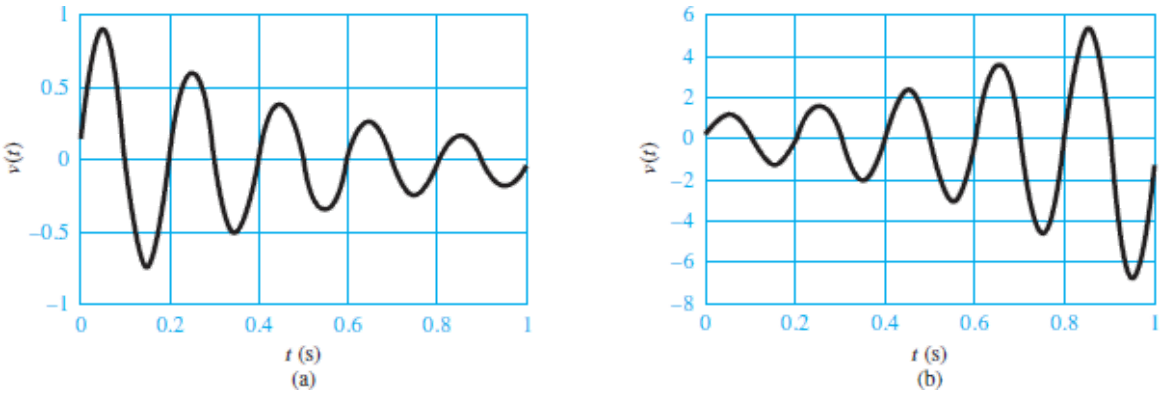


Figure B.1 Damped sinusoid: (*a*) exponential decay, negative σ; (b) exponential growth, positive σ

Note that the special case of $\sigma = 0$ corresponds to $s = j\omega$, that is, the familiar steady-state sinusoidal (phasor) case. We shall now refer to the complex variable $\mathbf{V}(s)$ as the **complex frequency domain** representation of $v(t)$. It should be observed that from the viewpoint of circuit analysis, the use of the Laplace transform is analogous to phasor analysis; that is, substituting the variable s wherever $j\omega$ was used is the only step required to describe a circuit using the new notation.

CHECK YOUR UNDERSTANDING

B.1 Find the complex frequencies that are associated with

a. $5e^{-4t}$ b. $\cos 2\omega t$ c. $\sin(\omega t + 2\theta)$ d. $4e^{-2t}\sin(3t - 50°)$ e. $e^{-3t}(2 + \cos 4t)$

B.2 Find s and $\mathbf{V}(s)$ if $v(t)$ is given by

a. $5e^{-2t}$ b. $5e^{-2t}\cos(4t + 10°)$ c. $4\cos(2t - 20°)$

Page 945

B.3 Find $v(t)$ if

a. $s = -2$, $\mathbf{V} = 2\angle 0°$ b. $s = j2$, $\mathbf{V} = 12\angle -30°$ c. $s = -4 + j3$, $\mathbf{V} = 6\angle 10°$

Answers: B.1: a. -4; b. $\pm j2\omega$; c. $\pm j\omega$; d. $-2 \pm j3$; e. -3 and $-3 \pm j4$.
B.2: a. -2, $5\angle 0°$; b. $-2 + j4$, $5\angle 10°$; c. $j2$, $4\angle -20°$. **B.3:** a. $2e^{-2t}$; b. $12\cos(2t - 30°)$; c. $6e^{-4t}\cos(3t + 10°)$

All the concepts and rules used in AC network analysis (see Chapter 3), such as impedance, admittance, KVL, KCL, and Thévenin's and Norton's theorems, carry over to the damped sinusoid case exactly. In the complex frequency domain, the current $\mathbf{I}(s)$ and voltage $\mathbf{V}(s)$ are related by the expression:

$$\mathbf{V}(s) = \mathbf{Z}(s)\mathbf{I}(s) \tag{B.6}$$

where $\mathbf{Z}(s)$ is the familiar impedance, with s replacing $j\omega$. We may obtain $\mathbf{Z}(s)$ from $\mathbf{Z}(j\omega)$ by simply replacing $j\omega$ by s. For a resistance R, the impedance is

$$\mathbf{Z}_R(s) = R \tag{B.7}$$

For an inductance L, the impedance is

$$\mathbf{Z}_L(s) = sL \tag{B.8}$$

For a capacitance C, it is

$$\mathbf{Z}_C(s) = \frac{1}{sC} \tag{B.9}$$

Impedances in series or parallel are combined in exactly the same way as in the AC steady-state case, since we only replace $j\omega$ by s.

EXAMPLE B.1 Complex Frequency Notation

Problem

Use complex impedance ideas to determine the response of a series RL circuit to a damped exponential voltage.

Solution

Known Quantities: Source voltage, resistor, inductor values.

Find: The time-domain expression for the series current $i_L(t)$.

Schematics, Diagrams, Circuits, and Given Data: $v_s(t) = 10e^{-2t}\cos(5t)$ V; $R = 4\ \Omega$; $L = 2$ H.

Assumptions: None.

Analysis: The input voltage phasor can be represented by the expression

$$\mathbf{V}(s) = 10\angle 0\ \text{V}$$

The impedance seen by the voltage source is

$$\mathbf{Z}(s) = R + sL = 4 + 2s$$

Thus, the series current is

$$\mathbf{I}(s) = \frac{\mathbf{V}(s)}{\mathbf{Z}(s)} = \frac{10}{4+2s} = \frac{10}{4+2(-2+j5)} = \frac{10}{j10} = j1 = 1\angle\left(-\frac{\pi}{2}\right)$$

Finally, the time-domain expression for the current is

$$i_L(t) = e^{-2t}\cos(5t - \pi/2) \qquad \text{A}$$

Comments: The phasor analysis method illustrated here is completely analogous to the method introduced in Chapter 3, with the complex frequency $j\omega$ (steady-state sinusoidal frequency) *replaced* by s (damped sinusoidal frequency).

Transfer functions $\boldsymbol{H}(s)$ can be defined as a ratio of a voltage to a current, a ratio of a voltage to a voltage, a ratio of a current to a current, or a ratio of a current to a voltage. The transfer function $\boldsymbol{H}(s)$ is a function of network elements and their interconnections. Using the transfer function and knowing the input (voltage or current) to a circuit, we can find an expression for the output either in the complex frequency domain or in the time domain. As an example, suppose $\mathbf{V}_i(s)$ and $\mathbf{V}_o(s)$ are the input and output voltages to a circuit, respectively, in complex frequency notation. Then

$$H(s) = \frac{\mathbf{V}_o(s)}{\mathbf{V}_i(s)} \tag{B.10}$$

from which we can obtain the output in the complex frequency domain by computing

$$\mathbf{V}_o(s) = H(s)\mathbf{V}_i(s) \tag{B.11}$$

If $\mathbf{V}_i(s)$ is a known damped sinusoid, we can then proceed to determine $v_o(t)$ by means of the method illustrated earlier in this section.

CHECK YOUR UNDERSTANDING

B.4 Given the transfer function $\boldsymbol{H}(s) = 3(s + 2)/(s^2 + 2s + 3)$ and the input $\mathbf{V}_i(s) = 4\angle 0°$, find the forced response $v_o(t)$ if

a. $s = -1$ b. $s = -1 + j1$ c. $s = -2 + j1$

B.5 Given the transfer function $\boldsymbol{H}(s) = 2(s + 4)/(s^2 + 4s + 5)$ and the input $\mathbf{V}_i(s) = 6\angle 30°$, find the forced response $v_o(t)$ if

a. $s = -4 + j1$ b. $s = -2 + j2$

Answers: B.4:
a. $6e^{-t}$; b. $12\sqrt{2}\,e^{-t}\cos(t + 45°)$; c. $6e^{-2t}\cos(t + 135°)$.
B.5: a. $3e^{-4t}\cos(t + 165°)$; b. $8\sqrt{2}\,e^{-2t}\cos(2t - 105°)$

B.2 THE LAPLACE TRANSFORM

The Laplace transform, named after the French mathematician and astronomer Pierre Simon de Laplace, is defined by

$$\mathcal{L}[f(t)] = \boldsymbol{F}(s) = \int_0^\infty f(t)e^{-st}\,dt \tag{B.12}$$

Page 947

The function $\boldsymbol{F}$(s) is the Laplace transform of $f(t)$ and is a function of the complex frequency $s = \sigma + j\omega$, considered earlier in this section. Note that the function $f(t)$ is defined only for $t \leq 0$. This definition of the Laplace transform applies to what is known as the **one-sided** or **unilateral Laplace transform**, since $f(t)$ is evaluated only for positive t. To conveniently express arbitrary functions only for positive time, we introduce a special function called the **unit-step function** $u(t)$, defined by the expression:

$$u(t) = \begin{cases} 0 & t < 0 \\ 1 & t > 0 \end{cases} \tag{B.13}$$

EXAMPLE B.2 Computing a Laplace Transform

Problem

Find the Laplace transform of $f(t) = e^{-at}\, u(t)$.

Solution

Known Quantities: Function to be Laplace-transformed.

Find: $F(s) = \mathcal{L}[f(t)]$.

Schematics, Diagrams, Circuits, and Given Data: $f(t) = e^{-at}\, u(t)$.

Assumptions: None.

Analysis: From equation B.12,

$$F(s) = \int_0^\infty e^{-at} e^{-st}\, dt = \int_0^\infty e^{-(s+a)t}\, dt = \frac{1}{s+a} e^{-(s+a)t}\Big|_0^\infty = \frac{1}{s+a}$$

Comments: Table B.1 contains a list of common Laplace transform pairs.

EXAMPLE B.3 Computing a Laplace Transform

Problem

Find the Laplace transform of $f(t) = \cos(\omega t)\, u(t)$.

Solution

Known Quantities: Function to be Laplace-transformed.

Find: $F(s) = \mathcal{L}[f(t)]$.

Schematics, Diagrams, Circuits, and Given Data: $f(t) = \cos(\omega t)\, u(t)$.

Assumptions: None.

Analysis: Using equation B.12 and applying Euler's identity to cos (ωt) give:

$$F(s) = \int_0^\infty \frac{1}{2}(e^{j\omega t} + e^{-j\omega t})e^{-st}\,dt = \frac{1}{2}\int_0^\infty (e^{(-s+j\omega)t} + e^{(-s-j\omega)t})\,dt$$

$$= \frac{1}{-s+j\omega}e^{-(s+j\omega)t}\Big|_0^\infty + \frac{1}{-s-j\omega}e^{-(s-j\omega)t}\Big|_0^\infty$$

$$= \frac{1}{-s+j\omega} + \frac{1}{-s-j\omega} = \frac{s}{s^2+\omega^2}$$

Comments: Table B.1 contains a list of common Laplace transform pairs.

Table B.1 Laplace transform pairs

$f(t)$	$F(s)$
$\delta(t)$ (unit impulse)	1
$u(t)$ (unit step)	$\frac{1}{s}$
$e^{-at}u(t)$	$\frac{1}{s+a}$
$\sin \omega t\, u(t)$	$\frac{\omega}{s^2+\omega^2}$
$\cos \omega t\, u(t)$	$\frac{s}{s^2+\omega^2}$
$e^{-at} \sin \omega t\, u(t)$	$\frac{\omega}{(s+a)^2+\omega^2}$
$e^{-at} \cos \omega t\, u(t)$	$\frac{s+a}{(s+a)^2+\omega^2}$
$tu(t)$	$\frac{1}{s^2}$

CHECK YOUR UNDERSTANDING

B.6 Find the Laplace transform of the following functions:

a. $u(t)$ b. $\sin(\omega t)\, u(t)$ c. $tu(t)$

B.7 Find the Laplace transform of the following functions:

a. $e^{-at} \sin \omega t\, u(t)$ b. $e^{-at} \cos \omega t\, u(t)$

Answers: B.6: a. $\frac{1}{s}$; b. $\frac{\omega}{s^2+\omega^2}$; c. $\frac{1}{s^2}$. **B.7:** a. $\frac{\omega}{(s+a)^2+\omega^2}$; b. $\frac{s+a}{(s+a)^2+\omega^2}$

From what has been said so far about the Laplace transform, it is obvious that we may compile a lengthy table of functions and their Laplace transforms by repeated application of equation B.12 for various functions of time $f(t)$. Then we could obtain a wide variety of inverse transforms by matching entries in the table. Table B.1 lists some of the more common **Laplace transform pairs**. The computation of the **inverse Laplace transform** is in general rather complex if one wishes to consider arbitrary functions of s. In many practical cases, however, it is possible to use combinations of known transform pairs to obtain the desired result.

EXAMPLE B.4 Computing an Inverse Laplace Transform

Problem

Find the inverse Laplace transform of

$$F(s) = \frac{2}{s+3} + \frac{4}{s^2+4} + \frac{4}{s}$$

Solution

Known Quantities: Function to be inverse Laplace-transformed.

Find: $f(t) = \mathcal{L}^{-1}[F(s)]$.

Page 949

Schematics, Diagrams, Circuits, and Given Data:

$$F(s) = \frac{2}{s+3} + \frac{4}{s^2+4} + \frac{4}{s} = F_1(s) + F_2(s) + F_3(s)$$

Assumptions: None.

Analysis: Using Table B.1, we can individually inverse-transform each of the elements of $\boldsymbol{F}(s)$:

$$f_1(t) = 2\mathcal{L}^{-1}\left(\frac{1}{s+3}\right) = 2e^{-3t}u(t)$$

$$f_2(t) = 2\mathcal{L}^{-1}\left(\frac{2}{s^2+2^2}\right) = 2\sin(2t)\,u(t)$$

$$f_3(t) = 4\mathcal{L}^{-1}\left(\frac{1}{s}\right) = 4u(t)$$

Thus

$$f(t) = f_1(t) + f_2(t) + f_3(t) = (2e^{-3t} + 2\sin 2t + 4)\,u(t)$$

EXAMPLE B.5 Computing an Inverse Laplace Transform

Problem

Find the inverse Laplace transform of

$$\boldsymbol{F}(s) = \frac{2s+5}{s^2+5s+6}$$

Solution

Known Quantities: Function to be inverse Laplace–transformed.

Find: $f(t) = \mathcal{L}^{-1}[\boldsymbol{F}(s)]$.

Assumptions: None.

Analysis: A direct entry for the function cannot be found in Table B.1. In such cases, one must compute a *partial fraction expansion* of the function $\boldsymbol{F}(s)$ and then individually transform each term in the expansion. A partial fraction expansion is the inverse operation of obtaining a common denominator and is illustrated below.

$$\boldsymbol{F}(s) = \frac{2s+5}{s^2+5s+6} = \frac{A}{s+2} + \frac{B}{s+3}$$

To obtain the constants A and B, we multiply the above expression by each of the denominator terms:

$$(s+2)\boldsymbol{F}(s) = A + \frac{(s+2)B}{s+3}$$

$$(s+3)\boldsymbol{F}(s) = \frac{(s+3)A}{s+2} + B$$

Page 950

From the above two expressions, we can compute A and B as follows:

$$A = (s+2)\boldsymbol{F}(s)|_{s=-2} = \left.\frac{2s+5}{s+3}\right|_{s=-2} = 1$$

$$B = (s+3)\boldsymbol{F}(s)|_{s=-3} = \left.\frac{2s+5}{s+2}\right|_{s=-3} = 1$$

Finally,

$$\boldsymbol{F}(s) = \frac{2s+5}{s^2+5s+6} = \frac{1}{s+2} + \frac{1}{s+3}$$

and using Table B.1, we compute

$$f(t) = (e^{-2t} + e^{-3t})\,u(t)$$

CHECK YOUR UNDERSTANDING

B.8 Find the inverse Laplace transform of each of the following functions:

a. $\boldsymbol{F}(s) = \dfrac{1}{s^2+5s+6}$ b. $\boldsymbol{F}(s) = \dfrac{s-1}{s(s+2)}$

c. $\boldsymbol{F}(s) = \dfrac{3s}{(s^2+1)(s^2+4)}$ d. $\boldsymbol{F}(s) = \dfrac{1}{(s+2)(s+1)^2}$

Answers: a. $f(t) = (e^{-3t} - e^{-2t})u(t)$; b. $f(t) = (\frac{3}{2}e^{-3t} - \frac{1}{2})u(t)$;
c. $f(t) = (\cos t - \cos 2t)u(t)$; d. $f(t) = (e^{-3t} + te^{-t} - e^{-t})u(t)$

B.3 TRANSFER FUNCTIONS, POLES, AND ZEROS

It should be clear that the Laplace transform is a convenient tool for analyzing the transient response of a circuit. The Laplace variable s is an extension of the steady-state frequency response variable $j\omega$ already encountered in this appendix. Thus, it is possible to describe the input-output behavior of a circuit by using Laplace transform ideas in the same way in which we used frequency response ideas earlier. Now we can define voltages and currents in the complex frequency domain as $\mathbf{V}(s)$ and $\mathbf{I}(s)$, and we denote impedances by the notation $\mathbf{Z}(s)$, where s replaces the familiar $j\omega$. We define an extension of the frequency response of a circuit, called the *transfer function*, as the ratio of any output variable to any input variable, i.e.,

$$H_1(s) = \frac{\mathbf{V}_o(s)}{\mathbf{V}_i(s)} \quad \text{or} \quad H_2(s) = \frac{\mathbf{I}_o(s)}{\mathbf{V}_i(s)} \quad \text{etc.} \tag{B.14}$$

As an example, consider the circuit of Figure B.2. We can analyze it by using a method analogous to phasor analysis by defining impedances:

$$\mathbf{Z}_1 = R_1 \qquad \mathbf{Z}_C = \frac{1}{sC} \qquad \mathbf{Z}_L = sL \qquad \mathbf{Z}_2 = R_2 \tag{B.15}$$

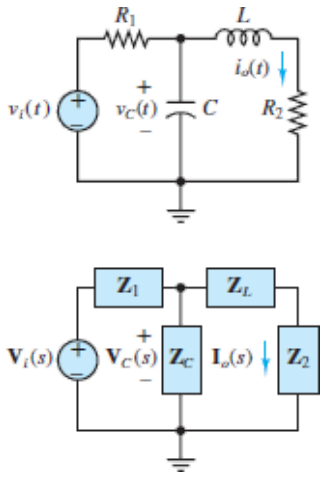


Figure B.2 A circuit and its Laplace transform domain equivalent

Page 951

Then we can use mesh analysis to determine

$$\mathbf{I}_o(s) = \mathbf{V}_i(s)\frac{\mathbf{Z}_C}{(\mathbf{Z}_L+\mathbf{Z}_2)\mathbf{Z}_C+(\mathbf{Z}_L+\mathbf{Z}_2)\mathbf{Z}_1+\mathbf{Z}_1\mathbf{Z}_C} \tag{B.16}$$

or, upon simplifying and substituting the relationships of equation B.15,

$$H_2(s) = \frac{\mathbf{I}_o(s)}{\mathbf{V}_i(s)} = \frac{1}{R_1LCs^2+(R_1R_2C+L)s+R_1+R_2} \tag{B.17}$$

If we were interested in the relationship between the input voltages and, say, the capacitor voltage, we could similarly calculate

$$H_1(s) = \frac{\mathbf{V}_C(s)}{\mathbf{V}_i(s)} = \frac{sL+R_2}{R_1LCs^2+(R_1R_2C+L)s+R_1+R_2} \tag{B.18}$$

Note that a transfer function consists of a *ratio of polynomials;* this ratio can also be expressed in factored form, leading to the discovery of additional important properties of the circuit. Let us, for the sake of simplicity, choose numerical values for the components of the circuit of Figure B.2. For example, let R_1 = 0.5 Ω, $C = \frac{1}{4}$ F, L = 0.5 H, and R_2 = 2 Ω. Then we can substitute these values into equation B.18 to obtain

$$H_1(s) = \frac{0.5s+2}{0.0625s^2+0.375s+2.5} = 8\left(\frac{s+4}{s^2+6s+40}\right) \tag{B.19}$$

Equation B.19 can be factored into products of first-order terms as follows:

$$H_1(s) = 8\left[\frac{s+4}{(s-3.0000+j5.5678)(s-3.0000-j5.5678)}\right] \tag{B.20}$$

where it is apparent that the response of the circuit has very special characteristics for three values of s: $s = -4$; $s = +3.0000 - j5.5678$; and $s = +3.0000 + j5.5678$. In the first case, at the complex frequency $s = -4$, the numerator of the transfer function becomes zero, and the response of the circuit is zero, regardless of how large the input voltage is. We call this particular value of s a **zero** of the transfer function. In the latter two cases, for $s = +3.0000 \pm j5.5678$, the response of the circuit becomes infinite, and we refer to these values of s as **poles** of the transfer function.

It is customary to represent the response of electric circuits in terms of poles and zeros, since knowledge of the location of these poles and zeros is equivalent to knowing the transfer function and provides complete information regarding the response of the circuit. Further, if the poles and zeros of the

transfer function of a circuit are plotted in the complex plane, it is possible to visualize the response of the circuit very effectively. Figure B.3 depicts the pole-zero plot of the circuit of Figure B.2; in plots of this type it is customary to denote zeros by a small circle and poles by an "×."

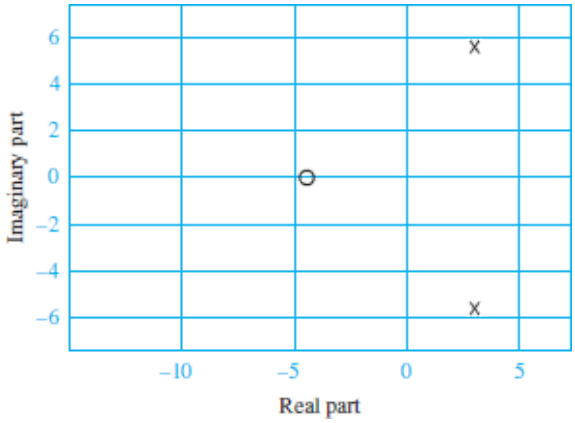


Figure B.3 Zero−pole plot for the circuit of Figure B.2

Page 952

The poles of a transfer function have a special significance, in that they are equal to the roots of the natural response of the system. They are also called the **natural frequencies** of the circuit. Example B.6 illustrates this point.

EXAMPLE B.6 Poles of a Second-Order Circuit

Problem

Determine the poles of a parallel *RLC* circuit. Express the homogeneous equation using i_L as the independent variable.

Solution

Known Quantities: Values of resistor, inductor, and capacitor.

Find: Poles of the circuit.

Assumptions: None.

Analysis: The differential equation describing the natural response of the parallel *RLC* circuit is

$$\frac{d^2i}{dt^2} + \frac{R}{L}\frac{di}{dt} + \frac{1}{LC}i = 0$$

with the characteristic equation given by

$$s^2 + \frac{R}{L}s + \frac{1}{LC} = 0$$

Now, let us determine the transfer function of the circuit, say, $\mathbf{V}_L(s)/\mathbf{V}_S(s)$. Applying the voltage divider rule, we can write

$$\frac{\mathbf{V}_L(s)}{\mathbf{V}_S(s)} = \frac{sL}{1/sC + R + sL}$$
$$= \frac{s^2}{s^2 + (R/L)s + 1/LC}$$

The denominator of this function, which determines the poles of the circuit, is identical to the characteristic equation of the circuit: The poles of the transfer function are identical to the roots of the characteristic equation!

$$s_{1,2} = -\frac{R}{2L} \pm \frac{1}{2}\sqrt{\left(\frac{R}{L}\right)^2 - \frac{4}{LC}}$$

Comments: Describing a circuit by means of its transfer function is completely equivalent to representing it by means of its differential equation. However, it is often much easier to derive a transfer function by basic circuit analysis than it is to obtain the differential equation of a circuit.

A P P E N D I X

C

FUNDAMENTALS OF ENGINEERING (FE) EXAMINATION

C.1 INTRODUCTION

The *Fundamentals of Engineering* (FE) examination[1] is one of four steps to be completed toward registering as a Professional Engineer (PE). Each of the 50 states in the United States has laws that regulate the practice of engineering; these laws are designed to ensure that registered professional engineers have demonstrated sufficient competence and experience. Each state's Board of Registration administers the exam and supplies information and registration forms.

The FE exam is offered throughout the year, except during the months of March, June, September, and December.

An examinee handbook is freely available through the NCEES website. The handbook contains information about eligibility, registration, fees,

accommodations, what to bring to the exam, the calculator policy, and answers to other questions and issues that are likely to occur. **Additional information is available on the NCEES website**.

Four steps are required to become a Professional Engineer:

1. *Education.* Usually this requirement is satisfied by completing a B.S. degree in engineering from an accredited college or university.

2. *Fundamentals of Engineering examination.* One must pass a discipline-specific examination described in Section C.2.
3. *Experience.* Following successful completion of the Fundamentals of Engineering examination, several years of engineering experience are required.
4. *Principles and practices of engineering examination.* One must pass a second examination, also known as the Professional Engineer (PE) examination, which requires in-depth knowledge of one particular branch of engineering.

This appendix provides a review of the background material in electrical engineering required in three of the discipline-specific FE exams. Those exams are prepared by the National Council of Examiners for Engineering and Surveying[2] (NCEES).

C.2 EXAM FORMAT AND CONTENT

The FE exam is offered in six specific engineering disciplines:

1. Chemical
2. Civil
3. Electrical and Computer
4. Environmental
5. Industrial and Systems
6. Mechanical

A seventh Other Disciplines exam is also offered. The 6-h, 110-question exam is offered year-round at NCEES-approved test centers. Detailed

specifications of each exam can be found online at http://ncees.org/engineering/fe/.

The passing score on the FE exam is not published by NCEES because it varies slightly across the discipline-specific exams and over time. However, data on passing rates is published and available on the NCEES website.

Of the seven exams, only three cover material presented in this book. Naturally, the Electrical and Computer Engineering exam covers nearly all the material. The Mechanical Engineering exam covers five areas of Electricity and Magnetism, namely:

- Charge, current, voltage, power, and energy
- Ohm's law and Kirchhoff's current and voltage laws
- Equivalent circuits (series and parallel)
- AC circuits
- Motors and generators

The Other Disciplines exam covers a similar set of topics as well as additional material on measuring devices, sensors, data acquisition, and data processing.

C.3 PRACTICE QUESTIONS ON ELECTRICITY AND MAGNETISM

What follows is a series of typical and relevant practice questions on FE exam material related to Electricity and Magnetism, including extensions of the theory to circuits, electronics, logic, instrumentation, communications, and electromechanics. The questions are ordered as they would be encountered in a typical engineering curriculum. Answers to the questions are provided at the end of this appendix.

Page 955

Students preparing for the FE exam should keep in mind that the actual exam time is 5 hours and 20 minutes, or 320 minutes. With 110 questions on the exam the average time per question is slightly less than 3 minutes.

Therefore, it is important to develop techniques for quickly arriving at correct answers or likely correct answers. In other words, it is not advisable to approach the FE exam as one would a typical undergraduate engineering exam. Rather it is worthwhile to develop skill at eliminating answers that are unreasonable or unlikely to be correct. For example, one can often eliminate answers due to the unreasonable scale of the answer or due to a units mismatch. It is also worthwhile to develop skill at approximating solutions. Remember, the average time per question is less than 3 minutes. The exam questions are designed with this limitation in mind. What does that tell you about the nature of many of the exam questions? When solving the practice questions below, look for ways in which you could have found the correct answer more quickly, more approximately, and/or with greater probability of correctness. And limit yourself to 3 minutes each!

Finally, the exam score is based solely on the number of correct answers. There are no deductions for wrong answers, so when in doubt, guess!

CHECK YOUR UNDERSTANDING

C.1 Determine the total charge entering a circuit element between $t = 1$ s and $t = 2$ s if the current passing through the element is $i = 5t$.

C.2 A lightbulb sees a 3-A current for 15 s. The lightbulb generates 3 kJ of energy in the form of light and heat. What is the voltage drop across the lightbulb?

C.3 How much energy does a 75-W electric bulb consume in 6 hours?

C.4 Find the voltage drop v_{ab} required to move a charge q from point a to point b if $q = -6$ C and it takes 30 J of energy to move the charge.

C.5 Two 2-C charges are separated by a dielectric with a thickness of 4 mm and with a dielectric constant $\varepsilon = 10^{-12}$ F/m. What is the force exerted by each charge on the other?

C.6 The magnitude of the force on a particle of charge q placed in the empty space between two infinite parallel plates with a spacing d and a potential difference V is proportional to:

a. qV/d^2

b. qV/d

c. qV^2/d

d. q^2V/d

e. q^2V^2/d

C.7 Assuming the connecting wires and the battery have negligible resistance, the voltage across the 25-Ω resistance in Figure C.7 is

a. 25 V

b. 60 V

c. 50 V

d. 15 V

e. 12.5 V

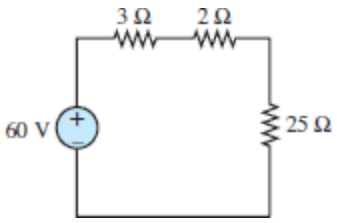


Figure C.7

C.8 Assuming the connecting wires and the battery have negligible resistance, the voltage across the 6-Ω resistor in Figure C.8 is

a. 6 V

b. 3.5 V

c. 12 V

d. 8 V

e. 3 V

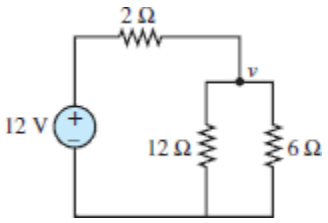


Figure C.8

C.9 A 125-V battery charger is used to charge a 75-V battery with internal resistance of 1.5 Ω, as shown in Figure C.9. If the charging current is not to exceed 5 A, the minimum resistance in series with the charger must be

a. 10 Ω

b. 5 Ω

c. 38.5 Ω

d. 41.5 Ω

e. 8.5 Ω

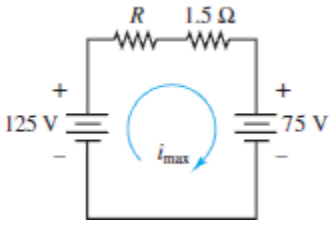


Figure C.9

C.10 A coil with an inductance of 1 H and negligible resistance carries the current shown in Figure C.10. The maximum energy stored in the inductor is

a. 2 J

b. 0.5 J

c. 0.25 J

d. 1 J

e. 0.2 J

Figure C.10

Page 956

C.11 The maximum voltage that will appear across the coil is

a. 5 V

b. 100 V

c. 250 V

d. 500 V

e. 5,000 V

C.12 A voltage sine wave of peak value 100 V is in phase with a current sine wave of peak value 4 A. When the phase angle is 60° later than a time at which the voltage and the current are both zero, the instantaneous power is most nearly

a. 300 W

b. 200 W

c. 400 W

d. 150 W

e. 100 W

C.13. A sinusoidal voltage whose amplitude is $20\sqrt{2}$ V is applied to a 5-Ω resistor. The root-mean-square value of the current is

a. 5.66 A

b. 4 A

c. 7.07 A

d. 8 A

e. 10 A

C.14 The magnitude of the steady-state root-mean-square voltage across the capacitor in the circuit of Figure C.14 is

a. 30 V

b. 15 V

c. 10 V

d. 45 V

e. 60 V

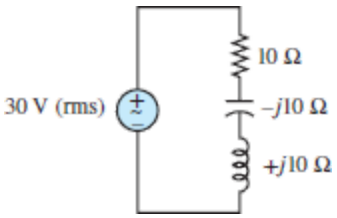


Figure C.14

The next set of questions (Exercises C.15 to C.19) pertain to single-phase AC power calculations and refer to the single-phase electrical network shown in Figure C.15. In this figure, $\mathbf{E}_S = 480\angle 0°$ V; $\mathbf{I}_S = 100\angle -15°$ A; $\omega = 120\pi$ rad/s. Further, load A is a bank of single-phase induction machines. The bank has an efficiency η of 80 percent, a power factor of 0.70 lagging, and a load of 20 hp. Load B is a bank of overexcited single-phase synchronous machines. The machines draw 15 kVA, and the load current leads the line voltage by 30°. Load C is a lighting (resistive) load and absorbs 10 kW. Load D is a proposed single-phase capacitor that will correct the source power factor to unity. This material is covered in Sections 13.1 and 13.2.

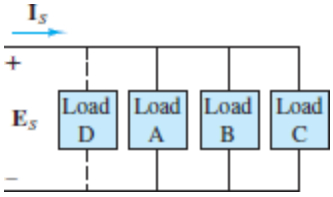


Figure C.15

C.15 The root-mean-square magnitude of load A current, denoted by I_A, is most nearly

a. 44.4 A

b. 31.08 A

c. 60 A

d. 38.85 A

e. 55.5 A

C.16 The phase angle of $\mathbf{I}_A$ with respect to the line voltage $\mathbf{E}_S$ is most nearly

a. 36.87°

b. 60°

c. 45.6°

d. 30°

e. 48°

C.17 The power absorbed by synchronous machines is most nearly

a. 20,000 W

b. 7,500 W

c. 13,000 W

d. 12,990 W

e. 15,000 W

C.18 The power factor of the system before load D is installed is most nearly

a. 0.70 lagging

b. 0.866 leading

c. 0.866 lagging

d. 0.966 leading

e. 0.966 lagging

C.19 The capacitance of the capacitor that will give a unity power factor of the system is most nearly

a. 219 μF

b. 187 μF

c. 132.7 μF

d. 240 μF

e. 132.7 pF

Page 957

$$v = \frac{\Delta w}{\Delta q} = \frac{3 \times 10^3}{45} = 66.67 \text{ V}$$

C.3: The energy used is
$w = pt = 75$ [W] $\times$ 6 [h] = 75 [W] $\times$ 6 $\times$ 3,600 [s] (450) = 1.62 MJ

C.4: The voltage drop is $v_{ab} = \frac{w}{q} = \frac{30}{-6} = -5$ V

C.5: $F = \frac{q_1 q_2}{4\pi\varepsilon r^2} = \frac{(2 \times 10^{-3}) \times (2 \times 10^{-3})}{4\pi \times 10^{-12} \times (4 \times 10^{-3})^2} = 2 \times 10^{10}$ N

C.6: Answer is a, since this is the only term that has a distance squared term in the denominator.

C.7: This problem calls for application of the voltage divider rule, discussed in Section 2.6. Applying the voltage divider rule to the circuit of Figure C.7, we have

$$v_{25\,\Omega} = 60\left(\frac{25}{3+2+25}\right) = 50 \text{ V}$$

Thus, the answer is c.

C.8: This problem can be solved most readily by applying nodal analysis (Section 3.1), since one of the node voltages is already known. Applying KCL at the node v, we obtain

$$\frac{12 - v}{2} = \frac{v}{6} + \frac{v}{12}$$

This equation can be solved to show that $v = 8$ V. Note that it is also possible to solve this problem by mesh analysis (Section 3.2). You are encouraged to try this method as well.

C.9: The circuit of Figure C.9 describes the charging arrangement. Applying KVL to the circuit of Figure C.9, we obtain

$$i_{\max} R + 1.5\, i_{\max} - 125 + 75 = 0$$

and using $i = i_{\max} = 5$ A, we can find R from the following equation:

$$5R + 7.5 - 125 + 75 = 0$$
$$R = 8.5\ \Omega$$

Thus, e is the correct answer.

C.10: The energy stored in an inductor is $W = \frac{1}{2} L i^2$ (see Section 4.1). Since the maximum current is 1 A, the maximum energy will be $W_{\max} = \frac{1}{2} L i_{\max}^2 = \frac{1}{2}$ J. Thus, b is the correct answer.

C.11: Since the voltage across an inductor is given by $v = L(di/dt)$, we need to find the maximum (positive) value of di/dt. This will occur anywhere between $t = 0$ and $t = 2$ ms. The corresponding slope is

$$\left.\frac{di}{dt}\right|_{\max} = \frac{1}{2 \times 10^{-3}} = 500$$

Therefore $v_{\max} = 1 \times 500 = 500$ V, and the correct answer is d.

C.12: As discussed in Section 7.1, the instantaneous AC power $p(t)$ is

$$p(t) = \frac{VI}{2}\cos\theta + \frac{VI}{2}\cos(2\omega t + \theta_V + \theta_I)$$

In this problem, when the phase angle is 60° later than a "zero crossing," we have $\theta_V = \theta_I = 0$, $\theta = \theta_V - \theta_I = 0$, $2\omega t = 120°$. Thus, we can compute the power at this instant as

$$p = \frac{100 \times 4}{2} + \frac{100 \times 4}{2}\cos(120°) = 300 \text{ W}$$

The correct answer is a.

Answers: **C.1**: $q = \int_{t=1}^{t=2} i\,dt = \int_{t=1}^{t=2} 5t\,dt = \left(\frac{5t^2}{2}\right)\Big|_{t=1}^{t=2} = 7.5\ \text{C}$

C.2: The total charge is $\Delta q = i\Delta t = 3 \times 15 = 45$ C. The voltage drop is

Answers: **C.13:** From Section 4.2, we know that

$$V_{\text{rms}} = \frac{V}{\sqrt{2}} = \frac{20\sqrt{2}}{\sqrt{2}} = 20 \text{ V}$$

Thus, $I_{\text{rms}} = 20/5 = 4$ A. Therefore, b is the correct answer.

C.14: This problem requires the use of impedances (Section 4.4). Using the voltage divider rule for impedances, we write the voltage across the capacitor as

$$\mathbf{V} = 30\angle 0° \times \frac{-j10}{10 - j10 + j10}$$
$$= 30\angle 0° \times (-j1) = 30\angle 0° \times 1\angle -90° = 30\angle 90°$$

Thus, the rms amplitude of the voltage across the capacitor is 30 V, and a is the correct answer. Note the importance of the phase angle in this kind of problem.

C.15: The output power P_o of the single-phase induction motor is $P_o = 20 \times 746 = 14.920$ W. The input electric power P_{in} is

$$P_{\text{in}} = \frac{P_o}{\eta} = \frac{14{,}920}{0.80} = 18{,}650 \text{ W}$$

P_{in} can be expressed as

$$P_{\text{in}} = E_S I_A \cos\theta_A$$

Therefore, the rms magnitude of the current $\mathbf{I}_A$ is found as

$$I_A = \frac{P_{\text{in}}}{E_S \cos\theta_A} = \frac{18{,}650}{480 \times 0.70} = 55.5015 \approx 55.5 \text{ A}$$

Thus, the correct answer is e.

C.16: The phase angle between $\mathbf{I}_A$ and $\mathbf{E}_S$ is

$$\theta = \cos^{-1} 0.70 = 45.57° \approx 45.6°$$

The correct answer is c.

C.17: The apparent power S is known to be 15 kVA, and θ is 30°. From the power triangle, we have

$$P = S\cos\theta$$

Therefore, the power drawn by the bank of synchronous motors is

$$P = 15{,}000 \times \cos 30° = 12{,}990.38 \approx 12.99 \text{ kW}$$

The answer is d.

C.18: From the expression for the current $\mathbf{I}_S$, we have

$$\text{pf} = \cos\theta = \cos[0° - (-15°)] = \cos 15° = 0.966 \text{ lagging}$$

The correct answer is e.

C.19: The reactive power Q_A in load A is

$$Q_A = P_A \times \tan\theta_A$$
$$\theta_A = \cos^{-1} 0.70 = 45.57°$$

Therefore,

$$Q_A = 18{,}650 \times \tan 45.57° = 19{,}025 \text{ VAR}$$

The total reactive power Q_B in load B is

$$Q_B = S \times \sin\theta_B = 15{,}000 \times \sin(-30°) = -7{,}500 \text{ VAR}$$

The total reactive power Q is

$$Q = Q_A + Q_B = 19{,}025 - 7{,}500 = 11{,}525 \text{ VAR}$$

To cancel this reactive power, we set

$$Q_C = -Q = -11{,}525 \text{ VAR}$$

and

$$Q_C = -\frac{E_S^2}{X_C} \quad \text{and} \quad X_C = -\frac{1}{\omega C}$$

Therefore, the capacitance required to obtain a power factor of unity is

$$C = -\frac{Q_C}{\omega E_S^2} = \frac{11{,}525}{120\pi \times 480^2} = 132.7 \ \mu\text{F}$$

The correct answer is c.

[1]This exam used to be called *Engineer in Training (EIT)*.

[2]P.O. Box 1686 (1826 Seneca Road), Clemson, SC 29633-1686.

Page 959

A P P E N D I X

D

ASCII CHARACTER CODE

In addition to the codes described elsewhere in the book (binary, octal, hexadecimal, binary-coded decimal), a character encoding convention adopted by all computer manufacturers is **ASCII**,[1] which maps a unique numeric value to each of 128 graphic or control characters commonly used in the display of text. The complete code is shown in Table D.1. Notice that the numeric values are shown in hexadecimal. An additional 128 nonstandard characters are often defined for any particular font implemented using the ASCII code, for a total of 256 characters in a typical font. It is no accident that 256 characters are often defined since that is the number of items that can be uniquely mapped by 8 bits or 1 byte of memory.

Page 960

Table D.1 ASCII

Graphic or control	ASCII (hex)	Graphic or control	ASCII (hex)	Graphic or control	ASCII (hex)
NUL	00	+	2B	V	56
SOH	01	,	2C	W	57
STX	02	−	2D	X	58
ETX	03	.	2E	Y	59
EOT	04	/	2F	Z	5A
ENQ	05	0	30	[	5B
ACK	06	1	31	\	5C
BEL	07	2	32	]	5D
BS	08	3	33	↑	5E
HT	09	4	34	←	5F
LF	0A	5	35	`	60
VT	0B	6	36	a	61
FF	0C	7	37	b	62
CR	0D	8	38	c	63
SO	0E	9	39	d	64
SI	0F	:	3A	e	65
DLE	10	;	3B	f	66
DC1	11	<	3C	g	67
DC2	12	=	3D	h	68
DC3	13	>	3E	i	69
DC4	14	?	3F	j	6A
NAK	15	@	40	k	6B
SYN	16	A	41	l	6C
ETB	17	B	42	m	6D
CAN	18	C	43	n	6E
EM	19	D	44	o	6F
SUB	1A	E	45	p	70
ESC	1B	F	46	q	71
FS	1C	G	47	r	72
GS	1D	H	48	s	73
RS	1E	I	49	t	74
US	1F	J	4A	u	75
SP	20	K	4B	v	76
!	21	L	4C	w	77
"	22	M	4D	x	78
#	23	N	4E	y	79
$	24	O	4F	z	7A
%	25	P	50	{	7B
&	26	Q	51	|	7C
'	27	R	52	}	7D
(	28	S	53	~	7E
)	29	T	54	DEL	7F
*	2A	U	55		

[1]American Standard Code for Information Interchange..

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Z

Index

A

B

C

D

E

F

G

H

I

7

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Z