Introduction to Communication Systems

THIRD EDITION

Abbreviations in Analog Modulation

AM	**Amplitude Modulation:** a CW modulation using amplitude variation in proportion to amplitude of the modulating signal; usually taken as DSB-LC for commercial broadcast transmissions and DSB-SC for multiplexed systems.
CW	**Continuous Wave:** a carrier signal (usually sinusoidal) used for modulation or keying.
DSB	**Double-Sideband (LC or SC):** a signal having two spectral sidebands symmetrically balanced with respect to the carrier frequency.
FM	**Frequency Modulation:** a CW modulation using frequency variation in proportion to amplitude of the modulating signal.
LC	**Large-Carrier:** a signal in which a relatively large proportion of the spectrum is concentrated at the carrier frequency (used with DSB, SSB, VSB).
NBFM	**Narrowband Frequency Modulation:** converts amplitude to frequency, but this effect is small compared with frequencies in the spectrum of the modulating signal.
NBPM	**Narrowband Phase Modulation:** converts amplitude to phase, but this effect is small compared with one radian of phase shift.
PAM	**Pulse-Amplitude Modulation:** varies amplitude in a train of constant-width pulses, in proportion to sampled values of the modulating signal.
PM	**Phase Modulation:** a CW modulation using phase variation in proportion to amplitude of the modulating signal.
PPM	**Pulse-Position Modulation:** uses constant-amplitude, constant-width pulses whose timings relative to a clock are proportional to sampled values of the modulating signal.
PWM	**Pulse-Width Modulation:** uses a train of constant-amplitude pulses whose widths are proportional to sampled values of the modulating signal.
QAM	**Quadrature AM:** uses both in-phase and quadrature AM to convey information.
SC	**Suppressed-Carrier:** a signal in which a relatively small proportion (ideally, zero) of the spectrum is concentrated at the carrier frequency (used with DSB, SSB, VSB).
SSB	**Single-Sideband (LC or SC)** ((+) indicates upper sideband, (−) indicates lower sideband): a signal having one spectral sideband which exhibits either the positive (+) or the negative (−) frequency content of the modulating signal.
VSB	**Vestigial-Sideband (LC or SC):** a signal for which the spectral density has the major portion of one sideband and a "vestige" of the second sideband.
WBFM	**Wideband Frequency Modulation:** converts amplitude to frequency, and this effect predominates compared with frequencies in the spectrum of the modulating signal.
WBPM	**Wideband Phase Modulation:** converts amplitude to phase, and this effect predominates compared with one radian of phase shift.

Abbreviations in Digital Modulation and Multiplexing

ASK	**Amplitude-Shift Keying:** switching between two amplitudes of a sinusoid to represent binary 1's and 0's.
APK	**Amplitude-Phase Keying:** combining ASK and PSK to convey information.
BPSK	**Binary Phase-Shift Keying:** switching between two phases of a sinusoid to represent binary 1's and 0's.
CPFSK	**Continuous-Phase FSK:** a version of FSK with no phase discontinuities at switching points, resulting in a decrease of the power spectral density at frequencies removed from the carrier frequency.
DM	**Delta Modulation:** DPCM using only one bit to encode differences between successive sampled values of the modulating signal.
DPCM	**Differential Pulse-Code Modulation:** uses PCM to send only the differences between successive sampled values of the modulating signal.
DPSK	**Differential PSK:** uses PSK to send only the differences between successive bits.
FDM	**Frequency-Division Multiplexing:** uses frequency offsets to convey several channels of independent information.
FSK	**Frequency-Shift Keying:** switching of instantaneous frequency of a sinusoid to represent binary 1's and 0's.
M-ary FSK	**Multistate FSK:** a version of FSK using more than two frequencies, providing efficient information transmission within a power constraint at a cost of bandwidth.
M-ary PSK	**Multistate PSK:** a version of PSK using more than two phases, providing efficient information transmission within a bandwidth constraint at a cost of power.
MSK	**Minimum-Shift Keying:** a popular version of CPFSK using minimum frequency shift for orthogonal keying.
MUX	**Multiplexer:** a system for combining several channels of information into one signal waveform.
OOK	**On-Off Keying:** a version of ASK in which the carrier signal waveform is transmitted at full amplitude ("on") or zero amplitude ("off") to represent binary 1's and 0's.
OQPSK	**Offset-keyed QPSK:** uses an offset in input data bit timing to decrease maximum phase transitions at the switching points.
QAM	**Quadrature Amplitude Modulation:** uses both in-phase and quadrature ASK to convey information.
QPSK	**Quadrature PSK:** uses both in-phase and quadrature PSK to convey information.
PCM	**Pulse-Code Modulation:** uses a digital code to designate the amplitude level of the modulating signal at each sample time.
PRK	**Phase-Reversal Keying:** switching of $(0, \pi)$ radians in phase of a sinusoid to represent binary 1's and 0's.
PSK	**Phase-Shift Keying:** switching of the phase of a sinusoid to represent binary 1's and 0's.
TDM	**Time-Division Multiplexing:** time sequencing of sampled values of two or more waveforms to convey several channels of independent information.

This book is in the
ADDISON-WESLEY SERIES IN ELECTRICAL ENGINEERING

Library of Congress Cataloging-in-Publication Data

Stremler, Ferrel G.
Introduction to communication systems / Ferrel Stremler.--3rd ed.
p. cm.
ISBN 0-201-18498-2
1. Telecommunication. 2. Signal theory (Telecommunication)
I. Title.
TK5103.S74 1990
621.382--dc20 89-38235
CIP

Reprinted with corrections October, 1992

Copyright © 1990, 1982, 1977 by Addison-Wesley Publishing Company, Inc.
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America.

16 17 18 19 20 MA 02 01 00 99

To Ruth, my wife.

Preface

This textbook presents to undergraduates an introductory explanation of communication systems, with the emphasis on signal design and modulation. The approach is therefore tailored to a careful development of the mathematical principles upon which such systems are based, using examples from a wide variety of current communication systems wherever possible. These range from commercial broadcasting and telephone systems to satellite telemetry and radar.

Extended Topical Coverage

Material added in this third edition is, primarily, about digital methods and reflects the continually increasing importance of digital signal transmission and modulation in communication systems. New material on the use of fiber optics in communications and the integrated services digital network (ISDN) has been added in Chapter 7, as well as an expanded treatment of Nyquist waveform shaping for control of ISI. Optimum filtering methods for use with ISI criteria have been included in Chapter 9. Chapter 10 has an expanded treatment of FSK and both parallel and serial MSK, as well as two new sections on spread spectrum systems. Therefore the presentation in this last chapter includes digital modulation methods that are currently being investigated for new communication systems designs.

Approach and Pedagogical Features

Because this textbook is intended for undergraduates, the material is written in as explicit a manner as possible and is clearly and liberally illustrated. Frequent use of *example problems* (117) and *drill problems* (133) with answers that, wherever possible, exemplify current practical problems helps to draw the student into an active participation in the learning process.

With the example problems worked out in the text, followed by drill problems with answers, this textbook lends itself to self-paced or individualized tutorial instruction. Each chapter ends with a summary section followed by a wide selection of problems so that the instructor can adjust the level considerably by assigning problems appropriate to the level of a specific course. Each of the problems is identified

by section number for content; thus students may refer to the appropriate text sections if they encounter difficulties.

A major change in this edition is an *expanded set of problems* for homework assignments at the end of each chapter. The focus of these problems is on analysis, verification, synthesis, extensions, illustrations of practice, and systems design considerations. Because not every problem can incorporate all aspects, there is a wide variety of problems (520 total) in this edition; 60% are new, and nearly 90% are either new or substantially revised from previous editions. It is my hope that these problems will stimulate interest as well as prove very beneficial in understanding the material presented.

There are 50 *computer problems* distributed throughout the book, with the heaviest concentration in Chapter 3 to make use of the DFT and the FFT algorithms. Although programs can be written for many of these problems, they are designed to be run using the engineering software so readily available for personal computers.[†]

A unique feature of this third edition is the inclusion of one or two special *systems design problems* at the end of each chapter. These problems have been taken from practical systems design situations and are intended to not only illustrate the relevance of the material covered but also to give students some concept of what types of problems might arise in engineering work in the area of communication systems. These problems are, by their nature, fairly broad and open-ended; students should learn that in general there is no unique answer but that their work will be judged more on their approach to the problem, taking into account the objectives and using good engineering judgment in the choices made. There is some gradual progression in the level of these problems throughout the book.

I recommend that students be encouraged to read the design problems to learn what types of problems some in this area might be working on within just a few years after taking this course. In addition, it is hoped that students will be challenged by such problems and perhaps their interests will be increased in the subject material. Another approach is to have students work in groups on a given design problem, perhaps one of their choosing. In general, there are no unique answers, but some approaches may be better than others, and some may offer better tradeoffs than others. That's what design is all about! Otherwise we are only teaching the mathematics.

Audience

Basically, the only prerequisites to a course using this textbook are a course in integral calculus and an introductory course in circuit analysis. A course in linear systems analysis would be helpful, but it is not essential.

Although written primarily for undergraduates in an electrical engineering curriculum, this text could also be used by those in other disciplines, in industry, or in

[†]Solutions to most problems designated as computer problems have been run on a personal computer using the student version of MathCAD (available from the publisher).

telecommunications practice who are interested in learning, reviewing, and updating their technical background in communication systems. For these groups the chapter arrangements and frequent examples and drill problems make the text appropriate to independent study.

Recommended auxiliary reading lists are included in the summary at the end of each chapter. Books in these lists have been carefully selected, and they should be both accessible and readable to undergraduates. They are listed in approximate order of increasing difficulty. References to specific topics are given as footnotes in the text. A majority of the references to periodicals is to the *IEEE Communication Magazine*, published monthly by the Institute of Electrical and Electronics Engineers and recommended for its excellence in original work, tutorials, and survey papers in the area of communication systems.

Organization of Text

The organization of this textbook is designed to allow maximum flexibility in the choice and presentation of subject matter. If Chapters 2, 3, and 8 constitute review for students, there is sufficient remaining material for a one-semester course. If the material in the early chapters is new to the students, some adjustment may be made by deleting the optional material in each chapter and/or not including the material on probability (Chapter 8). Optional material in each chapter has been designated by a star symbol (★).

Chapter 1 is an introduction to concepts in communication systems and an overview of the book. The Fourier methods of linear systems analysis are reviewed in Chapters 2 and 3, with particular emphasis on what will prove most useful in the succeeding chapters, such as the use of complex notation and interpretations in terms of phasors and spectral representations. These topics are not always brought out in linear systems analysis courses. Material on the numerical computation of Fourier coefficients and the discrete and fast Fourier transforms is widely used for both computational and signal processing applications, and is included as optional material in Chapters 2 and 3. Problems intended to be solved primarily using numerical methods are designated with a prefix check mark (√) in the margin.

The material in Chapters 4–7 is an introduction to the principles of communication systems. Chapter 4 covers the topics of correlation, power spectral density and thermal noise. It leads directly into the design of systems based on noise considerations. In fact, one can begin talking about satellite communication systems design considerations already in Chapter 4. Chapters 5–7 cover the topics of amplitude, angle, and pulse modulation. This organization of material has been influenced by my teaching this material at the undergraduate level. Students wish to see applications of the mathematical material. To sustain student interest, the more abstract concepts are interspersed with more practical sections that show how the concepts are being used. Thus the presentation begins with an elementary discussion of noise and quickly gets into communication systems, then proceeds through amplitude, angle,

and pulse modulation. This avoids having one or two chapters devoted entirely to signal-to-noise calculations—a topic that if prolonged fails to retain student interest at the undergraduate level.

The first part of this book does not assume a knowledge of probability theory. Presentation of the basic material without probability helps to keep the emphasis on signal design and modulation. This treatment ends with Chapter 7, and a course taught from a deterministic point of view could end here, or could conclude with some of the material in Chapter 10.

For those students for whom the first few chapters are review, there is time in a semester to take up the material in the last three chapters. If students, in addition, have had prior background in probability theory, Chapter 8 can be omitted or used for review. There is ample material in Chapters 9 and 10, in addition to Chapters 4–7, for a one-semester course if the optional sections are covered in each chapter.

Chapter 8 is an introduction to the subject of probability and random processes and is presented in such a way that students progress rapidly to the probability-density function and its use in the analysis of communication systems. Chapter 9 builds on this knowledge toward an introduction of such topics as quantization noise and probability of error in baseband transmission. Sections on partial-response signaling, equalization, M-ary signaling, power spectral densities of data waveforms, and coding for reliable communication can be covered if there is time.

Chapter 10 is a fairly complete discussion of digital modulation methods, beginning with amplitude-, frequency-, and phase-shift keying and progresses to modern methods of M-ary digital modulation such as quadrature phase-shift keying, minimum-shift keying, and amplitude-phase keying. New sections have been added on spread spectrum systems. The chapter concludes with geometric representations of digital waveforms and an introduction to maximum likelihood detection. After completing Chapter 10 the student will, it is hoped, be interested in taking an advanced course in communication theory that will employ more statistical concepts.

The appendixes from the second edition are included in this third edition because instructors found they were useful and readily available reference sources. The appendixes on commercial radio and television transmissions have been revised somewhat because of student interest in these topics and, pedagogically, add breadth in background. New sections on stereo television and high definition television have been added.

The material in Chapters 4–9 (to Section 9.7), plus the first four sections of Chapter 10 (but omitting starred sections) has been used for a one-semester course at the University of Wisconsin–Madison at the junior/senior level in electrical engineering. Many of the students in this course are not intending to major in communications but take the course for breadth and because it is recommended for such areas as signal processing, photonics, etc. The remaining material in Chapters 9 and 10 is covered in a succeeding course in communications for majors. Another variation might be to summarize some of the material (e.g., the signal-to-noise sections) in Chapters 4–7 so that more attention can be given to the material in Chapters 9 and 10. If all the material in the textbook is covered, there is ample material for a two-quarter course sequence.

Acknowledgments

I am indebted to many for their advice and assistance in this third edition. Suggestions and criticisms by reviewers for Addison–Wesley have been most helpful. In particular, I wish to thank Professor S. Hossein Mousavinezhad, Western Michigan University, for his helpful comments, and suggestions for problems in Chapter 9. The comments and reviews of Professors Joseph L. LoCicero, Illinois Institute of Technology, Sungwon Park, Texas A & I University, and Wesley W. Shelton, Jr., Florida Institute of Technology, were instructive and useful. I express my thanks to Don Fowley and Tom Robbins at Addison–Wesley for their support and encouragement to write this third edition. I appreciate the comments and suggestions made by graduate students that have improved the accuracy and clarity of the text, and for working through some advanced versions of my design problems. Also, I wish to thank Professors W. P. Birkemeier, J. A. Bucklew, and B. E. A. Saleh of the University of Wisconsin, and Mr. W. C. Luplow, Executive Director of Electronic Systems R&D, Zenith Electronics Corp., for their comments and suggestions on specific portions of the text changes and additions. The encouragement of Professor J. L. Shohet, Chairman of the Department of Electrical and Computer Engineering, is sincerely appreciated. My appreciation is extended also to those who were so helpful in two previous editions of this textbook.

Finally, I express my thanks for the constructive feedback and support of my students. Your comments and fresh insights continue to amaze me, and make teaching so enjoyable.

Madison, Wisconsin F. G. S.
December 1989

Additional Acknowledgments

My appreciation is extended also to those who were so helpful in the first and second editions of this textbook. In that connection, I found helpful certain concepts and illustrative examples from portions of Professor B. P. Lathi's out-of-print book *Communications Systems* (1968), one of the valuable and influential reference sources in the field.

Contents

CHAPTER 1 **Introduction** 1

CHAPTER 2 **Orthogonality and Signal Representations** 8

2.1 Signals and Systems 8
2.2 Classification of Signals 10
2.3 Classification of Systems 12
★ 2.4 Signals and Vectors 13
★ 2.5 Orthogonal Functions 15
★ 2.6 Choice of a Set of Orthogonal Functions 21
2.7 The Exponential Fourier Series 23
2.8 Complex Signals and Representations 26
2.9 The Trigonometric Fourier Series Representation 29
2.10 Extension by Periodicity 33
2.11 Parseval's Theorem for Power Signals 36
2.12 The Frequency Transfer Function 38
2.13 Steady-State Response to Periodic Signals 40
2.14 Harmonic Generation 42
2.15 The Fourier Spectrum and Examples 44
★ 2.16 Numerical Computation of Fourier Coefficients 50
★ 2.17 Effects of Alias Terms 53
2.18 Singularity Functions 59
2.19 Impulse Response 65
★ 2.20 Convergence of the Fourier Series 66
2.21 Summary 69
Problems 71

CHAPTER
3 The Fourier Transform and Applications 82

3.1 Representation of an Aperiodic Function Over the Entire Real Line **82**
3.2 The Spectral Density Function **85**
★ 3.3 Existence of the Fourier Transform **87**
3.4 Parseval's Theorem for Energy Signals **88**
3.5 Some Fourier Transforms Involving Impulse Functions **89**
3.6 Properties of the Fourier Transform **96**
3.7 Some Convolution Relationships **109**
3.8 Graphic Interpretation of Convolution **111**
3.9 Filter Characteristics of Linear Systems **114**
3.10 Transversal Filters **116**
3.11 Bandwidth of a System **118**
3.12 Requirements for Distortionless Transmission **119**
3.13 Time Response of Filters **120**
★ 3.14 Minimum Time-Bandwidth Product **124**
3.15 The Sampling Theorem **126**
★ 3.16 Aliasing Effects in Sampling **131**
★ 3.17 The Discrete Fourier Transform **135**
★ 3.18 The Fast Fourier Transform **141**
3.19 Summary **145**
Problems **147**

CHAPTER
4 Spectral Density and Correlation 163

4.1 Energy Spectral Density **163**
4.2 Power Spectral Density **168**
4.3 Time-Averaged Noise Representations **176**
4.4 Correlation Functions **179**
4.5 Some Properties of Correlation Functions **184**

4.6 Correlation Functions for Finite-Energy Signals **187**
4.7 Band-Limited White Noise **188**
4.8 Summary **208**
Problems **209**

CHAPTER
5 Amplitude Modulation **219**

5.1 Amplitude Modulation: Suppressed Carrier **219**
5.2 Amplitude Modulation: Large Carrier (AM) **238**
5.3 Frequency-Division Multiplexing (FDM) **249**
5.4 Single-Sideband (SSB) Modulation **255**
5.5 Vestigial-Sideband (VSB) Modulation **266**
5.6 A Time-Representation of Bandpass Noise **268**
5.7 Signal-to-Noise Ratios in AM Reception **271**
5.8 Propagation Effects **277**
5.9 Comparison of Various AM Systems **281**
5.10 Summary **282**
Problems **284**

CHAPTER
6 Angle Modulation **298**

6.1 FM and PM **298**
6.2 Narrowband FM **301**
6.3 Wideband FM **306**
6.4 Average Power in Angle-Modulated Waveforms **317**
6.5 Phase Modulation **319**
6.6 Generation of Wideband FM Signals **321**
6.7 Demodulation of FM Signals **327**

6.8 Signal-to-Noise Ratios in FM Reception **345**
6.9 Threshold Effect in FM **351**
6.10 Signal-to-Noise Improvement Using Deemphasis **355**
6.11 Summary **359**
Problems **361**

CHAPTER
7 Pulse Modulation 371

7.1 Pulse-Amplitude Modulation (PAM) **372**
7.2 Time-Division Multiplexing (TDM) **377**
7.3 Pulse Shaping and Intersymbol Interference **386**
7.4 Other Types of Analog Pulse Modulation: PWM and PPM **390**
7.5 Signal-to-Noise Ratios in Analog Pulse Modulation **399**
7.6 Pulse-Code Modulation (PCM) **402**
7.7 Fiber Optic Communication Systems **413**
7.8 Use of Parity and Redundancy in PCM **417**
7.9 Time-Division Multiplexing of PCM Signals **419**
7.10 Integrated Services Digital Network (ISDN) **429**
7.11 The Matched Filter **431**
7.12 Matched-Filter Codeword Detection **437**
7.13 Pseudonoise (PN) Sequences **439**
7.14 Summary **443**
Problems **445**

CHAPTER
8 Probability and Random Variables 457

8.1 Probability **457**
8.2 Conditional Probability and Statistical Independence **461**

8.3 The Random Variable and Cumulative Distribution Function **463**
8.4 The Probability Density Function **466**
8.5 Statistical Averages **471**
8.6 Some Probability Distributions **477**
★ 8.7 The Histogram **485**
★ 8.8 Transformations of Random Variables **487**
8.9 Joint and Conditional Density Functions **490**
★ 8.10 Correlation Between Random Variables **493**
★ 8.11 The Bivariate Gaussian Distribution **497**
8.12 Random Processes **498**
8.13 Autocorrelation and Power Spectra **504**
★ 8.14 Numerical Computation of Power Spectra **515**
8.15 Summary **517**
Problems **519**

CHAPTER
9 Information and Digital Transmission 531

9.1 A Measure of Information **531**
9.2 Channel Capacity **536**
9.3 Ideal Demodulator Detection Gain **540**
9.4 Quantization Noise **541**
9.5 Probability of Error in Transmission **547**
9.6 S/N Performance of PCM **558**
9.7 Delta Modulation and DPCM **560**
★ 9.8 Error Analysis of PCM Repeaters **564**
★ 9.9 Power Spectral Densities of Data Waveforms **567**
9.10 Partial-Response Signaling **572**
9.11 Equalization **577**
9.12 M-ary Signaling **584**
★ 9.13 Coding for Reliable Communication **589**
9.14 Summary **594**
Problems **596**

CHAPTER
10 Digital Modulation 608

10.1 Amplitude-Shift Keying (ASK) 608
10.2 Frequency-Shift Keying (FSK) 610
10.3 Phase-Shift Keying (PSK) 618
10.4 Comparison of Binary Digital Modulation Systems 626
10.5 Direct-Sequencing Spread Spectrum Systems 627
10.6 Quadrature AM (QAM) and Quaternary PSK (QPSK) 634
10.7 Continuous-Phase FSK (CPFSK) and Minimum-Shift Keying (MSK) 641
★ 10.8 *M*-ary Orthogonal FSK 651
10.9 Frequency-Hopping (FH) Spread Spectrum Systems 655
10.10 *M*-ary PSK 659
10.11 Amplitude-Phase Keying (APK) 663
10.12 Comparison of Digital Modulation Systems 666
★ 10.13 Representation of Digital Waveforms 668
★ 10.14 Optimum Detection Algorithms 673
10.15 Summary 678
Problems 680

APPENDIXES

A Selected Mathematical Tables 693
B Decibels 698
C Broadcast Frequency Bands 700
D Commercial Television Transmissions 708
E Telephone Channels 729
F Some Commercial Preemphasis/Deemphasis Systems 732
G A Table of Bessel Functions 735
H Stereo AM 737
I A Table of Gaussian Probabilities 746

Index 749

Front Endpapers

Left Abbreviations in analog modulation

Right Abbreviations in digital modulation and multiplexing

Back Endpapers

Left Selected Fourier transform pairs
Multiples and submultiples of units

Right Some Fourier transforms corresponding to given mathematical operations
Gaussian error probabilities

Introduction to Communication Systems

THIRD EDITION

CHAPTER

1 Introduction

It is difficult to imagine what modern living would be like without ready access to reliable, economical, and efficient means of communication. Communication systems are found wherever information is to be transmitted from one point to another. Telephone, radio, and television are common, everyday examples of communication systems. More complicated communication systems guide aircraft, spacecraft, and automated trains; others provide live news coverage around the world, often via satellite—and the list of examples could go on and on. It is hardly an overstatement to say that today communication systems are not only necessary to business, industry, banking, and the dissemination of information to the public, but also essential to the national welfare and defense.

The purpose of this textbook is to present an introductory treatment of communication systems. By "communication" we mean the conveying or transmission of information from one place and/or time to another. Admittedly this definition is not very precise, but the subject of communication is very broad. It may, for example, mean anything from a telephone hookup to the use of effective gestures, emphasis, and diction in a speech; from a smoke signal to a satellite transmission. Note that the commonality in these examples is that there is information transmitted that is of importance to the recipient.

In this study of communications we shall restrict ourselves to the transmission of information over comparatively long distances. The use of electrical signals has almost completely replaced all other forms of information transmission over long distances. This has happened mainly because electrical signals are relatively easy to control (compared, for instance, to the fire for the smoke signals) and travel with velocities at or near that of the speed of light. In the broader sense, we may consider light to fall within this class because it is in the electromagnetic spectrum. Certainly for long distances, then, a study of communications via electrical signals is appropriate.

What is information transmission? This subject turns out to be more complicated than it might first appear and will, in fact, form a basis for part of our study. From a strictly intuitive viewpoint, we can say that the transmission of information requires that signals vary with time. Consider, for example, a 9-V battery; once the voltage has been established, there is little further information available without changing the voltage. Connection of the battery to a variable resistor or transistor allows such a variation. However, simply the fact that the signal can vary with time is not sufficient. Consider a 120-V, 60-Hz sinusoidal voltage; once the voltage has been established, there is little further information available without changing the amplitude or phase of the sinusoidal voltage, even though the voltage is changing at a 60-Hz rate.

We conclude not only that the transmission of information is related to signals changing with time but also that these changes must be made in an unpredictable way. Thus a necessary requirement is the use of a band of signal frequency content known as bandwidth (Chapter 3). Bandwidth is a measure of how rapidly the information-bearing portions of a signal can change, and therefore it is an important parameter in any discussion of communication systems. These intuitive ideas of information transmission are enough for understanding the basic methods used in electrical communication systems. However, a more complete mathematical description of information transmission also demands some knowledge of probability (Chapter 8) and is postponed until Chapter 9.

Communication over long distances usually requires that some alterations or other operations be performed on the electrical signal conveying the information in preparation for transmission. Upon reception, known inverse operations are performed to retrieve the information.

In the process of transmission, the signals bearing the information are contaminated by noise. Noise is generated by numerous natural and man-made events and introduces errors in the information transmission. From an engineering point of view, the communication problem consists in designing those portions of the transmission over which one can exercise some control. A criterion for doing this is to keep the information transmission as error-free as possible.

With these desired objectives in mind, we shall consider different communication systems and their basic principles of operation. The emphasis will be on the methods and not on the particular circuits or devices currently employed.

A diagram of the basic units comprising a communication system is shown in Fig. 1.1. Not every communication system makes use of all indicated operations, but each always involves a transmission medium of some kind. The encoder chooses the best form for the signal in order to optimize its detection at the output. The decoder performs the inverse operation to make the best decision, based on the available signals, that a given message was indeed sent. The design of the encoder and decoder must rely on a detailed mathematical description of information transmission. While the subject of coding often carries with it an air of secrecy, a more important motive in many modern coding systems is the improved efficiency in conveying information.

The modulator produces a varying signal at its output which is proportional in some way to the signal appearing across its input terminals. For example, a sinusoidal modulator may vary the amplitude, frequency, or phase of a sinusoidal signal in direct proportion to the voltage input. The roles of the encoder and the modulator

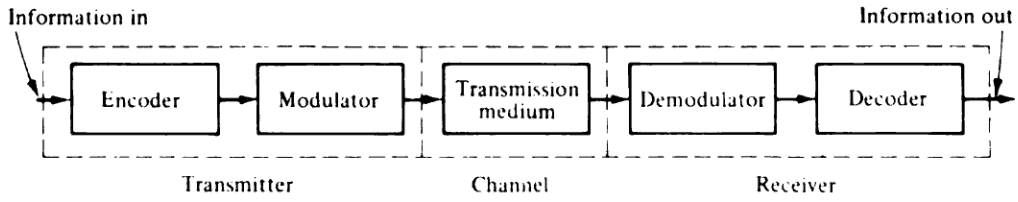


Figure 1.1 A communication system.

are similar in that both prepare the signal for more efficient transmission. However, the process of coding is designed to optimize the error-free detection that a given message is being sent, whereas the process of modulation is designed to impress the information signal onto the waveform to be transmitted. The demodulator performs the inverse operation of the modulator to recover the signal in its original form.

The transmission medium is the crucial link in the system. Without it, there would be no communication problem. The transmission medium may include the ionosphere, the troposphere, free space, or simply a transmission line. In any case, attenuation and distortion, as well as noise signals generated in the media and the transmitting and receiving equipment, are introduced. For our purposes, noise signals are any electrical signals (voltages or currents) that interfere with the error-free reception of the message-bearing signal.

Three very basic subsystems of a communication system are indicated by the dashed lines in Fig. 1.1. The central subsystem restricts the flow of information and is called the *channel*. The channel includes the effects of additive noise, interference, propagation, and distortion. It is the limiting factor in the performance of any well-designed communication system. The role of the *transmitter* is to prepare the information to be sent in such a way that it will best cope with the limitations imposed by the channel. The role of the *receiver* is to perform the inverse of the transmitter operations in order to recover the information with the least error possible. Note that, in this broad sense, the transmitter and the receiver as a pair are specifically designed to combat the deleterious effects of the channel on information transmission.

The communication system shown in Fig. 1.1 is capable of one-way transmission and is called a *simplex* (SX) transmission system. In many cases it is desirable to maintain two-way communication, or at least to be able to send a message back to its origin for possible verification, comparison, or control. One method of accomplishing this is to use the same channel alternately for transmission in each direction, as shown in Fig. 1.2. This method of transmission is called *half-duplex* (HDX). Although communication flows in both directions in half-duplex transmission, the flow of information is only one-way at any given time.

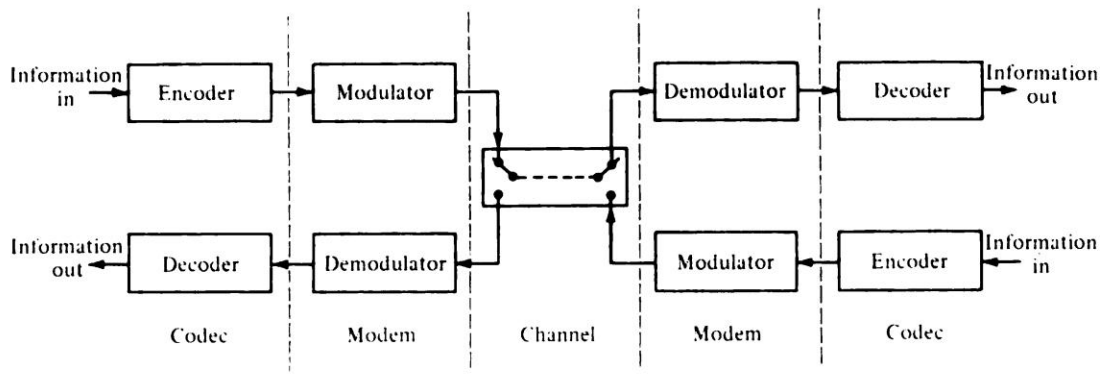


Figure 1.2 Communication system using half-duplex transmission.

A third type, *full-duplex* (FDX), is shown in Fig. 1.3. In full-duplex transmission, simultaneous communication is accomplished in both directions. Note that in both HDX and FDX transmission, the modulators and demodulators operate in pairs. This combination of a modulator and a demodulator is called a *modem* (*mo*dulator-*dem*odulator) in data transmission systems. Also, the encoders and decoders operate in pairs, suggesting the term *codec* (*co*der-*dec*oder).

Now consider a channel whose only transmission impairment is that of additive noise. As mentioned previously, a certain minimum bandwidth, B, is necessary for communication. A wider bandwidth would allow more noise to interfere with the information transmission, so it is important to keep the bandwidth in such a channel as small as possible.

We characterize the noise present by its average power, N, and the transmitted signal by its average power, S. If the average noise power is relatively small, then the signal power does not need to be very large for the receiver to determine what information is being sent. (We are, of course, also concerned with efficiency and therefore attempt to minimize the transmitted power required to convey the information to the user.) In contrast, the average signal power required must be relatively large when the noise power is large. We conclude that it is the *ratio* of the average signal power to the average noise power that is important, and not the magnitudes of S and N themselves. This S/N ratio, called the *signal-to-noise ratio*, is an important parameter in the theory and design of communication systems.

All communication systems can be judged in terms of bandwidth, signal-to-noise ratio, and economic (cost) factors. There are various trade-offs between these parameters in any given type of system, and we will examine these later in this book. Signal-to-noise required involves both transmitted power and a performance criterion, so that for some systems the concept of the three-way performance trade-offs depicted in Fig. 1.4 is more convenient. Some trade-offs between bandwidth and signal-to-noise ratio are emphasized in the following brief discussion of digital systems.

We assume (at least temporarily) that the communication systems are digital and, specifically, that they are binary. Suppose, then, that the information we wish to send at a given instant can be characterized by one sample from n possible input states, all

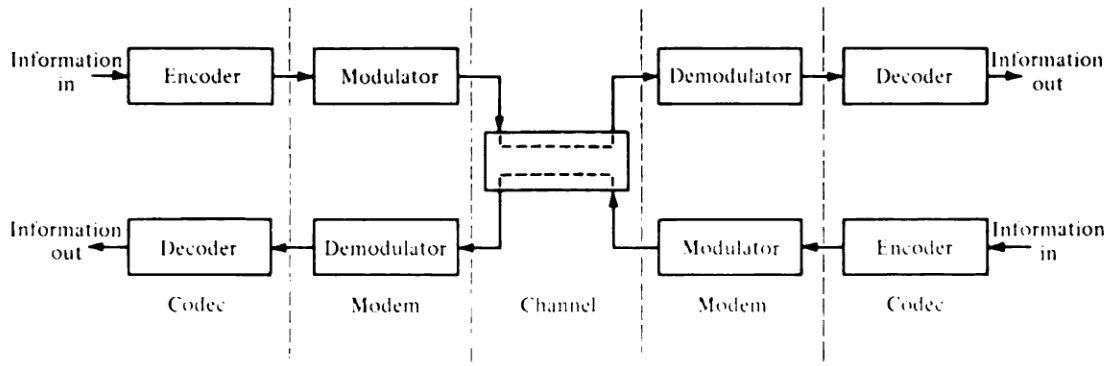


Figure 1.3 Communication system using full-duplex transmission.

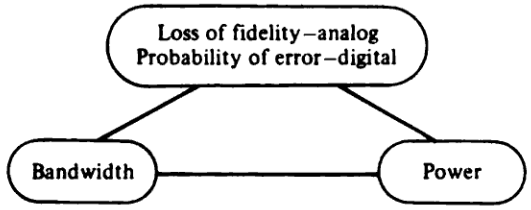


Figure 1.4 Performance trade-off factors in communication systems.

of which are equally likely to occur. For example, this sample could be one of 256 equally likely voltage levels. To send this sample using a binary system, we first generate a digital word composed of m binary symbols. Thus each binary word consists of $m = \log_2 n$ binary digits to represent the one sample out of $n = 2^m$ possibilities. When used in this way, m is called the number of *bits* (*bi*nary digi*ts*) needed to represent the one-of-n possible input states. In this example, for instance, an eight-bit word is necessary to describe one of 256 possible input states.

Next we wish to send this m-bit binary word serially through the channel. In a binary system, these bits are represented by binary symbols (e.g., $+1$ and -1) that are generated at a symbol rate of r symbols per second. The resulting information rate from the transmitter is $R = mr$ bits per second (bps). At the receiver, the transmitted signal is corrupted by the addition of noise, and as a result the receiver will make some errors. It seems reasonable that this error rate will decrease if S/N is increased. It also seems reasonable that perhaps the error rate could be decreased by a receiver designed to process signals with an increased complexity. Because we are interested in both efficiency and accuracy of communication, a question of great importance to us, then, is the following: For a given channel and a given information transmission rate, is it possible theoretically to make system improvements with the objective of reducing the error rate? The answer to this question, based on the theoretical work of Claude Shannon published in 1949,[†] is in the affirmative *if* the information transmission rate R is such that $R \leq C$, where C is the *channel capacity*. For the type of channel considered here, the channel capacity is given by the Hartley-Shannon law

$$C = B \log_2 (1 + S/N) \quad \text{bps} \tag{1.1}$$

where B is the bandwidth of the channel (in Hz) and S/N is the signal-to-noise ratio. If one tries to send information at too rapid a rate — that is, $R > C$ — then the errors begin to increase rapidly and there is no point to trying to design a system to improve the situation. On the other hand, for $R < C$ there is some hope for improvement via good system design. All communication systems to be covered in this textbook, then, will lie within the bound expressed by Eq. (1.1).

Let's pursue our intuitive reasoning farther. Suppose we decide to increase the information transmission rate by increasing the symbol rate r. Because more symbol transitions per second are possible, the required bandwidth must be increased. Thus an increase in information rate can be accommodated by an increase in bandwidth.

[†] C. E. Shannon, "Communication in the Presence of Noise," *Proceedings of the IRE*, vol. 37 (January 1949): 10–21.

This merely emphasizes a conclusion we could have obtained directly from Eq. (1.1). What is surprising, however, is that Eq. (1.1) states that bandwidth and signal-to-noise can be exchanged. Thus a smaller S/N ratio may be adequate if the bandwidth is increased, and vice versa. Note that for small S/N the potential trade-off is approximately linear,

$$\log_2 (1 + x) \approx (\log_2 e)x \qquad \text{for small } x,$$

but it is exponential for large S/N.

It turns out that the Hartley-Shannon law is applicable to continuous as well as discrete systems, and therefore it is a very powerful and far-reaching result. Its application is restricted, however, to channels with additive noise, and such effects as distortion and interference are not included.

Although the Hartley-Shannon law states that the maximum information transmission rate for a given channel is bounded and suggests that bandwidth can be traded off for S/N, it does not offer a method for the design of a system that will meet these expectations. In other words, it gives us a bound with which we can compare the performance of the systems we design, but it does not give us a procedure for designing systems whose performance will meet that bound. But we should not expect too much. It's fortunate we have the above result.

With this brief introduction to communication systems, we turn now to matters of approach and organization of this textbook.

It is our purpose in this textbook to analyze the principal characteristics of communication systems and to discuss some of their realizations in practice. Our approach will be from a systems viewpoint, in contrast to a study of circuits and devices. This will require mathematical descriptions and representations of the electrical signals (that is, the voltages and currents) that characterize such systems.

The Fourier methods of signal analysis prove to be most useful in our study. Applications of these methods often furnish valuable insights into the signal design aspects of communication systems. These methods are introduced and developed in the first part of this book. An introduction to communication systems, from both an analysis and an applications approach, then follows in the succeeding chapters.

Selected References for Further Reading[†]

1. A. B. Carlson. *Communication Systems*. Third ed. New York: McGraw-Hill, 1986. Chapter 1 of this text contains a good introductory discussion of communication, communication systems, and modulation.
2. W. D. Gregg. *Analog and Digital Communication*. New York: John Wiley & Sons, 1977. Chapter 1 has a good descriptive introduction to historical perspectives, frequency allocations, licensing, and regulations.

[†] Selected references for further reading in the material covered are included at the end of each chapter. In general, they are listed in an order of increasing difficulty.

3. M. Schwartz. *Information Transmission, Modulation, and Noise*. Third ed. New York: McGraw-Hill, 1980.
The first chapter of this book introduces the mathematical formulation of information transmission and communication systems for the interested student.

4. B. Sklar. *Digital Communications*. Englewood Cliffs, N.J.: Prentice Hall, 1988.
The first part of Chapter 1 has an interesting overview of digital communication systems, based on a two-part paper published by the author.

CHAPTER

2 Orthogonality and Signal Representations

We begin this chapter by reviewing some concepts of signal and systems representations that will prove helpful to us later. The concept of orthogonality, for instance, leads to the use of the Fourier series, which in turn is used to introduce the concepts of signal and systems representations in both time and frequency.

2.1 SIGNALS AND SYSTEMS

In a general sense, a *system* is a group of objects that can interact harmoniously and that are combined in a manner intended to achieve a desired objective. A system may in turn be a member (subsystem) of a larger system. A complete hierarchy of systems may be established, each with its identified domain.

A *signal* is an event that serves, or at least is capable, to start some action; i.e., it can incite action. Within energy and power restrictions, we are particularly interested in the concept of a signal and also in the *response* of a system to a given signal. The diagram in Fig. 2.1 illustrates the role of the signal, the system, and the response. Although the concept of a system was introduced first, it is often more convenient to use the concepts of the signal and the resulting response to describe the characteristics of a system. Sometimes a system is known only in terms of its response to given signals.

For our purposes, a signal is defined to be a single-valued function of time, which means that to each assigned instant of time (the independent variable) there is one unique value of the function (the dependent variable). This value may be a real number, in which case we have a real-valued signal, or it may be complex, and then we can speak of a complex-valued signal. In either case the independent variable (time) is real-valued.

The complex-valued notation can be used to describe signals in terms of two independent variables, for example $x(t)$ and $y(t)$. Thus the complex-valued notation is convenient for a description of two-dimensional phenomena, such as circular motion,

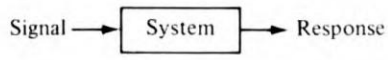


Figure 2.1 A system diagram.

plane wave propagation, etc., as a function of time. The signals used in this book, however, are expressed in terms of time and only one dependent variable (e.g., a voltage versus time). Because we are limited to one dependent variable, we insist that all signals that correspond to physically observable quantities must be real-valued quantities.

But then, why use complex notation at all if our observations are real-valued? In many analyses the mathematical models and calculations are often much simpler, and even more intuitively obvious, if the complex notation is used. For the sake of convenience, the complex notation will prove to be a definite advantage. After all operations have been completed, it is a fairly simple matter to take the real part of the resulting expression. This procedure works whenever superposition holds; we must be cautious in applying it to more general cases.

The preceding remarks can be applied to the description and analysis of physical processes in general. We shall restrict ourselves principally to the description and analysis of *electrical* signals and systems.

An electrical signal may be either a voltage or a current waveform that we are describing mathematically. We are not as interested in "voltage drops," "loop currents," etc., as we are in the time variations of signals, whether they be voltages or currents. It follows that a signal is merely a single-valued function of time that may be used to represent either a voltage or a current for a specific situation. Exceptions to this will sometimes arise, particularly in discussions involving energy and power. In this case, the assumption of a 1-ohm resistor is a convenient way to avoid any direct commitment. The value of a particular resistor will just scale the result once the identity of the waveform is established. For all calculations of energy and power a 1-ohm resistor is assumed unless otherwise explicitly specified in the given problem.

Sinusoidal signals play a major role in the analysis of communication systems. Such a signal $f(t)$ may be represented as a function of time t by the equation

$$f(t) = A \cos(\omega t + \theta), \tag{2.1}$$

where A is the amplitude, θ is the phase, and ω is the rate of phase change or frequency of the sinusoid in radians per second. It may also be expressed as f in cycles per second (Hz), where $\omega = 2\pi f$.[†]

The principle of the Fourier methods of signal analysis is to break up all signals into summations of sinusoidal components. This provides a description of a given signal in terms of sinusoidal frequencies. A major objective is a description of how signal (and response) energy and power are distributed in terms of these frequencies. Any description of a response to a given signal will, of course, bring in the characteristics of the system.

[†] The generally accepted unit for frequency in cycles per second (cps) is the *hertz*, abbreviated as Hz. Using prefix multipliers, we then have the following: 1 kHz (one kilohertz) = 10^3 cps; 1 MHz (one megahertz) = 10^6 cps; 1 GHz (one gigahertz) = 10^9 cps.

2.2 CLASSIFICATION OF SIGNALS

The most useful method of signal representation for any given situation depends upon the type of signal being considered. A few of the classifications most useful to us are discussed here.

Energy Signals, Power Signals

An energy signal is a pulselike signal that usually exists for only a finite interval of time or, even if present for an infinite amount of time, at least has a major portion of its energy concentrated in a finite time interval.

For electrical systems, a signal is a voltage or a current. The instantaneous power dissipated by a voltage $e(t)$ in a resistance R is

$$p = |e(t)|^2/R \qquad \text{watts,} \tag{2.2}$$

and for a current $i(t)$

$$p = |i(t)|^2 R \qquad \text{watts.} \tag{2.3}$$

In each instance the instantaneous power is proportional to the squared magnitude of the signal. For a 1-ohm resistance, these equations assume the same form. Thus it is customary in signal analysis to speak of the instantaneous power associated with a given signal $f(t)$ as

$$p = |f(t)|^2 \qquad \text{watts.} \tag{2.4}$$

Even though the dimensions may not appear to be correct in Eq. (2.4), the multiplication or division by an appropriate resistance is implied in this convention.

Using this convention, the energy dissipated by the signal during a time interval (t_1, t_2) is

$$E = \int_{t_1}^{t_2} |f(t)|^2\, dt \qquad \text{joules.} \tag{2.5}$$

We define an *energy signal* to be one for which Eq. (2.5) is finite even when the time interval becomes infinite; i.e., when

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt < \infty\,. \tag{2.6}$$

Examples of several energy signals are shown in Fig. 2.2.

The average power dissipated by the signal $f(t)$ during the time interval (t_1, t_2) is

$$P = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} |f(t)|^2\, dt\,. \tag{2.7}$$

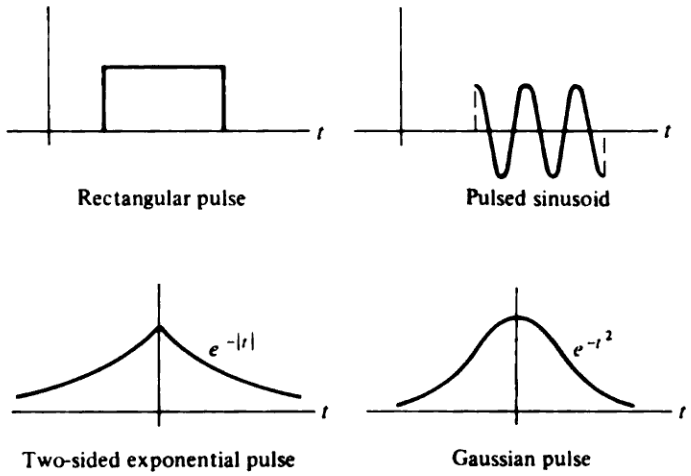


Figure 2.2 Some energy signals.

If the right-hand side of Eq. (2.7) remains finite but nonzero when the time interval becomes infinite, i.e., if

$$0 < \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2 \, dt < \infty, \tag{2.8}$$

then the signal $f(t)$ has finite average power and is called a *power signal*.

Periodic, Nonperiodic

A *periodic signal* is one that repeats itself exactly after a fixed length of time. Thus the signal $f(t)$ is periodic if there is a number T such that

$$f(t + T) \equiv f(t) \quad \text{for all } t. \tag{2.9}$$

The smallest positive number T that satisfies Eq. (2.9) is called the *period*. The period defines the duration of one complete cycle of $f(t)$. A periodic signal is a power signal if its energy per cycle is finite, and then the average power need only be calculated over one complete cycle.

Any signal for which there is no value of T satisfying Eq. (2.9) is said to be *nonperiodic*, or *aperiodic*. A borderline case between periodic and nonperiodic signals is that of the "almost periodic signal." This latter type of signal is composed of the sum of two or more periodic signals having incommensurate periods.[†] An example of such a signal is

$$f(t) = \sin t + \sin \sqrt{2}\, t. \tag{2.10}$$

This function is "almost periodic" because each term on the right-hand side is periodic, yet there is no period T in which $f(t)$ exactly repeats itself.

[†] Two periods are commensurate if their ratio can be expressed as the ratio of two integers.

Random, Deterministic

A *random signal* is one about which there is some degree of uncertainty before it actually occurs. Such a signal can be thought of as belonging to a collection of signals, each of which is different. If one of these signals is chosen (at random), it may turn out to be quite well defined, as, for example, a sinusoid of fixed frequency but uncertain starting phase. However, if a second sinusoid were chosen from this collection, we could not be certain of the same starting phase. In other cases, future values of the signal may not be predictable even after observation of past values. An example of such a random signal is the output of a radio receiver when tuned off station as it responds to noise arising from disturbances in the atmosphere and its internal circuitry.

A *nonrandom*, or *deterministic*, signal is one about which there is no uncertainty in its values. In almost all cases, an explicit mathematical expression can be written for such a signal. The signals to be discussed in the earlier part of this book are nonrandom. As new methods of analysis are developed, we shall begin to see how certain types of random signals can also be handled.

2.3 CLASSIFICATION OF SYSTEMS

Mathematically, a system is a rule used for assigning a function $g(t)$ (the output) to a function $f(t)$ (the input); that is,

$$g(t) = \mathcal{T}\{f(t)\}, \tag{2.11}$$

where $\mathcal{T}\{\ \}$ is the rule.† This rule could be in terms of an algebraic operation, a differential and/or integral equation, etc. For two systems connected in cascade, the output of the first system forms the input to the second, thus forming a new overall system:

$$g(t) = \mathcal{T}_2\{\mathcal{T}_1[f(t)]\} = \mathcal{T}\{f(t)\}. \tag{2.12}$$

As in signal analysis, we find it convenient to classify systems by some of their basic properties. Those most useful to us are discussed below.

Linear, Nonlinear

If a system is *linear* then superposition applies; that is, if

$$g_1(t) = \mathcal{T}\{f_1(t)\}, \quad \text{and} \quad g_2(t) = \mathcal{T}\{f_2(t)\},$$

then

$$\mathcal{T}\{a_1 f_1(t) + a_2 f_2(t)\} = a_1 g_1(t) + a_2 g_2(t), \tag{2.13}$$

where a_1, a_2 are constants. A system is *linear* if it satisfies Eq. (2.13); any system not meeting these requirements is *nonlinear*.

† The symbol used here is the script "t." Other script letters are used in later chapters to designate specific operators.

Time-Invariant or Time-Varying

A system is *time-invariant* if a time shift in the input results in a corresponding time shift in the output so that

$$g(t - t_0) = \mathcal{T}\{f(t - t_0)\} \quad \text{for any } t_0 . \tag{2.14}$$

The output of a time-invariant system depends on time differences and not on absolute values of time. Any system not meeting this requirement is said to be *time-varying*. A system may be linear yet time-varying and vice versa. Two examples of such systems are shown in Fig. 2.3.

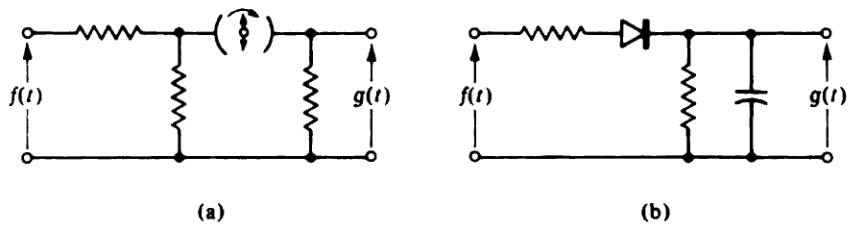


Figure 2.3 Examples of systems: (a) Linear, time-varying; (b) nonlinear, time-invariant.

Realizable, Nonrealizable

A physically realizable system cannot have an output response before an arbitrary input function is applied. Stated another way, the output of a physical system at $t = t_0$, namely, $g(t_0)$, must depend only on values of the input $f(t)$ for $t \leq t_0$. A system having this property is called *physically realizable or causal.*[†] Any system not meeting this requirement is said to be nonrealizable or noncausal.

2.4 ★ SIGNALS AND VECTORS[‡]

Consider the signal $f(t)$ which is defined for all values of time within the interval (t_1, t_2), as shown in Fig. 2.4. This signal is time-limited and has a finite energy, E:

$$E = \int_{t_1}^{t_2} |f(t)|^2 \, dt < \infty .$$

The graph in Fig. 2.4 illustrates one way of specifying $f(t)$; i.e., for each value of the independent variable t we are given the value of the dependent variable, $f(t)$. The closer we take the increments in time, the more exact we know the variations of the signal $f(t)$. Obviously it takes an infinite set of values to completely specify the graph

[†] Some texts define "realizability" in other ways; this definition will suffice for our purposes.
[‡] Optional material in each chapter is designated by a star.

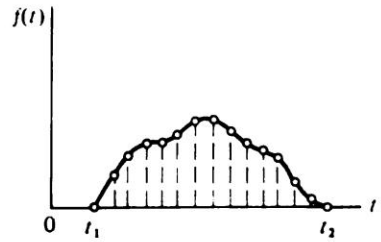


Figure 2.4 A finite-energy signal.

for $f(t)$. From experience, however, we know that as long as we have a few samples for each wiggle on the graph, we can construct a graph that is accurate enough for all practical purposes. We shall place this criterion on a more quantitative basis in Chapter 3 (the sampling theorem).

Alternatively, we may wish to specify $f(t)$ by a countable set of numbers which are not dependent on the explicit choice of the variable t.† In other words, we would like to express $f(t)$ as

$$f(t) = \sum_n f_n \phi_n(t), \tag{2.15}$$

where the $\phi_n(t)$ form a set of functions to be specified and the f_n are numbers independent of time.

This idea should not seem particularly surprising to us, for we are quite accustomed to describing a vector in terms of a countable set of numbers. For example, a vector from the origin to the point (1, 1, 1) is that vector which has a length of $\sqrt{3}$ and is inclined at 55° to each of the coordinate axes in the first quadrant. The reason that this vector notation is so convenient, of course, is that each number in the notation represents the projected length of the vector upon a given coordinate axis. These coordinate axes are chosen in such a way that they are perpendicular to each other in the vector space. Thus the projection of the vector on each coordinate axis is entirely independent of its projection on each of the other axes. This allows us to write down the equation of the vector immediately in terms of these respective projections and the reference vectors in each coordinate direction.

The analogy between vectors and signals is more than coincidental. In fact, it promises to give us some valuable insights into the description of signals. The use of the more familiar geometric ideas often prove very useful in gaining added perspective into the handling of signal waveforms.

We assume throughout this discussion that all vectors of interest have a finite length. Such vectors can be uniquely specified in space by referring them to a given set of coordinate axes. There must be one coordinate for each dimension of the vector in order to guarantee the uniqueness of the representation. Then we say that the vector space is *complete*.

The scalar ("dot") product between two vectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2$ is

$$C_{12} = \boldsymbol{\phi}_1 \cdot \boldsymbol{\phi}_2 . \tag{2.16}$$

† A countable set is a set which can be placed in a one-to-one correspondence to the set of real integers.

The scalar C_{12} is an indicator of the similarity between the vectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2$. If C_{12} is zero, we conclude that either (a) $\boldsymbol{\phi}_1$ or $\boldsymbol{\phi}_2$ have zero magnitude, or (b) $\boldsymbol{\phi}_1$ has no component along the vector $\boldsymbol{\phi}_2$. In the latter case the two vectors are mutually perpendicular and are said to be *orthogonal*. $\perp$

Suppose we generate an orthogonal vector space with the three orthogonal vectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \boldsymbol{\phi}_3$. These vectors do not necessarily have unit length; however, we can write

$$\boldsymbol{\phi}_n \cdot \boldsymbol{\phi}_m = \begin{Bmatrix} k_n & n = m \\ 0 & n \neq m \end{Bmatrix}, \tag{2.17}$$

where k_n is the squared length of $\boldsymbol{\phi}_n$. Any vector $\mathbf{A}_1$ in this vector space may be represented in the form

$$\mathbf{A}_1 = A_{11}\boldsymbol{\phi}_1 + A_{12}\boldsymbol{\phi}_2 + A_{13}\boldsymbol{\phi}_3, \tag{2.18}$$

where

$$A_{1n} = \frac{\mathbf{A}_1 \cdot \boldsymbol{\phi}_n}{\boldsymbol{\phi}_n \cdot \boldsymbol{\phi}_n} = \mathbf{A}_1 \cdot \left(\frac{\boldsymbol{\phi}_n}{\boldsymbol{\phi}_n \cdot \boldsymbol{\phi}_n}\right) = \mathbf{A}_1 \cdot \left(\frac{\boldsymbol{\phi}_n}{k_n}\right) \tag{2.19}$$

for $n = 1, 2, 3$. Note that as a result of orthogonality the computation of each scalar product in Eq. (2.19) is not dependent on the computation of any of the others. It also turns out that if the vector space is not complete, the squared length of the error vector remaining is a minimum using the above methods.[†]

We may extend these concepts to a more general N-dimensional space. Such a vector space, of course, does not exist physically in nature. On the other hand, there are many problems that have analogies to an N-dimensional space.

DRILL PROBLEM 2.4.1

Three vectors, expressed in the Cartesian coordinate system described by the unit vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, are $\mathbf{A} = \mathbf{x}_1 - \mathbf{x}_2 + 5\mathbf{x}_3$; $\mathbf{B} = -\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3$; $\mathbf{C} = 3\mathbf{x}_1 + \mathbf{x}_2 + 2\mathbf{x}_3$; $\mathbf{D} = \mathbf{x}_1 + 5\mathbf{x}_2 - 4\mathbf{x}_3$. (a) Determine which of these vectors are orthogonal to $\mathbf{D}$. (b) Represent $\mathbf{A}$ in terms of the three vectors $\mathbf{B}, \mathbf{C}, \mathbf{D}$. (c) Compute the squared length of the error vector remaining if $\mathbf{A}$ is represented in terms of $\mathbf{C}, \mathbf{D}$ only.

ANSWER: (a) $\mathbf{B}, \mathbf{C}$; (b) $\mathbf{A} = \mathbf{B} + \frac{6}{7}\mathbf{C} - \frac{4}{7}\mathbf{D}$; (c) 3.

2.5 ★ ORTHOGONAL FUNCTIONS

Returning to the original problem [cf. Eq. (2.15)], recall that we wish to express the function $f(t)$ as a set of numbers (the f_n) which, when expressed in terms of a properly chosen coordinate space [i.e., the $\phi_n(t)$], will specify the function uniquely. It is

[†] See, e.g., B. P. Lathi, *Signals, Systems and Communications*, New York: John Wiley & Sons, 1965.

highly desirable that the set so chosen be a linearly independent set; i.e., that the individual terms are not dependent on each other and that the set is formed by the totality of these terms.†

We have seen that any vector of finite length may be expressed as a sum of its components along n mutually orthogonal vectors provided that these vectors form a complete set of mutually perpendicular coordinates. In addition, this method yields a minimum length (squared) for the error vector when the set is not complete. These ideas naturally form a motivation for choosing a complete set of orthogonal functions, the $\phi_n(t)$. The particular $\phi_n(t)$ chosen are called *basis functions*.

Two complex-valued functions $\phi_1(t)$ and $\phi_2(t)$ are defined to be *orthogonal* over the interval (t_1, t_2) if

$$\int_{t_1}^{t_2} \phi_1(t)\phi_2^*(t)\,dt = \int_{t_1}^{t_2} \phi_1^*(t)\phi_2(t)\,dt = 0\,. \tag{2.20}$$

Thus if members of a set of complex-valued functions are mutually orthogonal over (t_1, t_2), then

$$\int_{t_1}^{t_2} \phi_n(t)\phi_m^*(t)\,dt = \begin{cases} 0 & n \neq m \\ K_n & n = m \end{cases}. \tag{2.21}$$

The set of basis functions $\phi_n(t)$ is said to be "normalized" if

$$K_n = \int_{t_1}^{t_2} |\phi_n(t)|^2\,dt = 1 \quad \text{for all } n.$$

If the set is both orthogonal and normalized, it is called an *orthonormal* set.

The integral of the product of two functions over a given interval is called the *inner product* of the two functions. The square root of the inner product of a function with itself is called the *norm*. Thus Eq. (2.21) is the inner product of $\phi_1(t)$, $\phi_2(t)$ and $K_n^{1/2}$ is the norm. As we shall soon see, the inner product and the norm in signal space are analogous to the dot product and length in vector space.

Now we return to the question of an approximation of a function $f(t)$ in terms of the $\phi_n(t)$ [cf. Eq. (2.15)]; for N terms we have

$$f(t) \cong \sum_{n=1}^{N} f_n \phi_n(t)\,. \tag{2.22}$$

The integral-squared error remaining in this approximation after N terms is

$$\int_{t_1}^{t_2} |\epsilon_N(t)|^2\,dt = \int_{t_1}^{t_2} |f(t) - \sum_{n=1}^{N} f_n\phi_n(t)|^2\,dt\,. \tag{2.23}$$

The integrand in Eq. (2.23) is nonnegative so that the integral-squared error will go to zero only if the integrand goes to zero.

† A set of functions is linearly independent if no one of the functions can be constructed as a weighted sum of the remaining functions in the set.

We wish to minimize Eq. (2.23) by a proper choice of the f_n. Expanding the right-hand side of Eq. (2.23) and allowing an interchange in the order of summation and integration, we have

$$\int_{t_1}^{t_2} |\epsilon_N(t)|^2\,dt = \int_{t_1}^{t_2} |f(t)|^2\,dt - \sum_{n=1}^{N}\left[f_n^* \int_{t_1}^{t_2} f(t)\phi_n^*(t)\,dt + f_n \int_{t_1}^{t_2} f^*(t)\phi_n(t)\,dt - |f_n|^2 K_n\right], \tag{2.24}$$

where use has also been made of Eq. (2.21). Completing the square within the summation, Eq. (2.24) becomes

$$\int_{t_1}^{t_2} |\epsilon_N(t)|^2\,dt = \int_{t_1}^{t_2} |f(t)|^2\,dt + \sum_{n=1}^{N}\left[\left|K_n^{1/2} f_n - \frac{1}{K_n^{1/2}} \int_{t_1}^{t_2} f(t)\phi_n^*(t)\,dt\right|^2 - \left|\frac{1}{K_n^{1/2}} \int_{t_1}^{t_2} f(t)\phi_n^*(t)\,dt\right|^2\right]. \tag{2.25}$$

All three terms on the right-hand side of Eq. (2.25) are nonnegative. Only the second term, however, is dependent on our choice of the f_n and is minimized by

$$f_n = \frac{1}{K_n} \int_{t_1}^{t_2} f(t)\phi_n^*(t)\,dt = \frac{\int_{t_1}^{t_2} f(t)\phi_n^*(t)\,dt}{\int_{t_1}^{t_2} |\phi_n(t)|^2\,dt}. \tag{2.26}$$

Note the close similarity with vector spaces [cf. Eq. (2.19)] and the analogous roles played by the inner product and the dot product.

Using Eq. (2.26) in Eq. (2.25), the minimum integral-squared error in the orthogonal series approximation to $f(t)$ over (t_1, t_2) is

$$\int_{t_1}^{t_2} |\epsilon_N(t)|^2\,dt = \int_{t_1}^{t_2} |f(t)|^2\,dt - \sum_{n=1}^{N} |f_n|^2 K_n. \tag{2.27}$$

The right-hand side of Eq. (2.27) is the difference between two nonnegative quantities (which in turn must be nonnegative). Note that the first term is the energy in $f(t)$. If, for any $f(t)$ having finite energy, i.e., for

$$\int_{t_1}^{t_2} |f(t)|^2\,dt < \infty$$

the $\phi_n(t)$ are such that

$$\lim_{N\to\infty} \int_{t_1}^{t_2} |\epsilon_N(t)|^2\,dt = 0, \tag{2.28}$$

then we say that the orthogonal set $\phi_n(t)$ is *complete* over (t_1, t_2).

For a complete orthogonal set, Eq. (2.15) can now be rewritten as

$$f(t) = \sum_{n=1}^{\infty} f_n \phi_n(t) \tag{2.29}$$

and Eq. (2.27) reduces to

$$\int_{t_1}^{t_2} |f(t)|^2 \, dt = \sum_{n=1}^{\infty} |f_n|^2 K_n . \tag{2.30}$$

This relation is known as *Parseval's theorem*.

The representation of a function $f(t)$ by an infinite set of mutually orthogonal functions is called a "generalized Fourier series representation" of $f(t)$. We shall find frequent applications for such representations in our analysis of signals and systems. A possible way of generating the coefficients for such a representation is shown in Fig. 2.5.

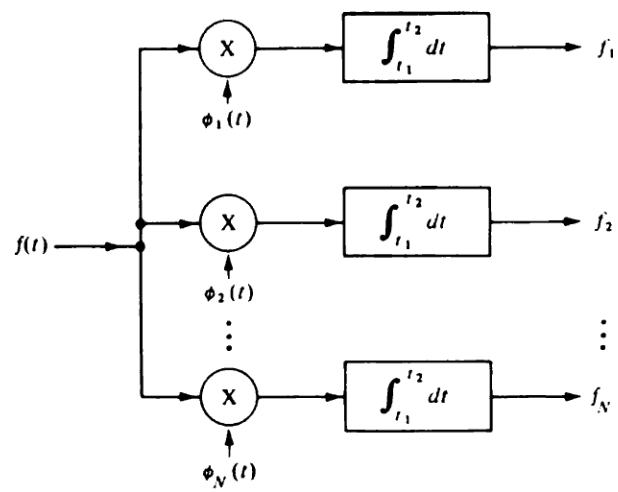


Figure 2.5 A method of generating the coefficients for the generalized Fourier series representation.

EXAMPLE 2.5.1

A given rectangular function is defined over (0, 2) by (see Fig. 2.6)

$$f(t) = \begin{cases} 1, & 0 < t < 1 \\ -1, & 1 < t < 2 \end{cases}.$$

We wish to approximate this finite-energy function using a set of functions defined by $\phi_n(t) = \sin n\pi t$, $n > 0$, over the interval (0, 2).

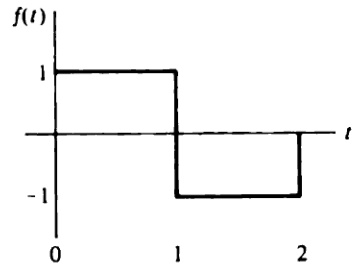


Figure 2.6 A rectangular function.

Solution We can easily show that the functions $\sin n\pi t$ and $\sin m\pi t$ are orthogonal over the interval $(0, 2)$:

$$\int_0^2 \sin n\pi t \sin m\pi t\, dt = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}.$$

[In fact, note that this set is orthonormal over $(0, 2)$.] Then we have the following representation for $f(t)$ over $(0, 2)$:

$$f(t) = \sum_{n=1}^{\infty} f_n \sin n\pi t,$$

where the f_n can be found from Eq. (2.26):

$$f_n = \frac{\int_0^2 f(t) \sin n\pi t\, dt}{\int_0^2 \sin^2 n\pi t\, dt} = \int_0^2 f(t) \sin n\pi t\, dt.$$

For this particular choice of $f(t)$,

$$f_n = \int_0^1 \sin n\pi t\, dt - \int_1^2 \sin n\pi t\, dt = \begin{cases} 4/\pi n & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}.$$

Thus the function $f(t)$ is approximated by the series representation

$$f(t) = \frac{4}{\pi}\left(\sin \pi t + \frac{1}{3} \sin 3\pi t + \frac{1}{5} \sin 5\pi t + \frac{1}{7} \sin 7\pi t + \cdots\right)$$

over the interval $(0, 2)$. Figure 2.7 shows the actual function and the approximation when the function is approximated with one, two, and three terms, respectively. The point error (instantaneous error) is also shown for each approximation.

The error energy (integral-squared error) can be calculated using Eq. (2.27):

$$\int_0^2 \epsilon_N^2(t)\, dt = 2 - \sum_{\substack{n=1 \\ n \text{ odd}}}^{N} \left(\frac{4}{\pi n}\right)^2.$$

Results for the first six nonzero terms are shown in Table 2.1. In this case the error energy quickly decreases with the first few terms in the series. In fact, the calculations show that about 95% of the energy is contained in the first four terms.

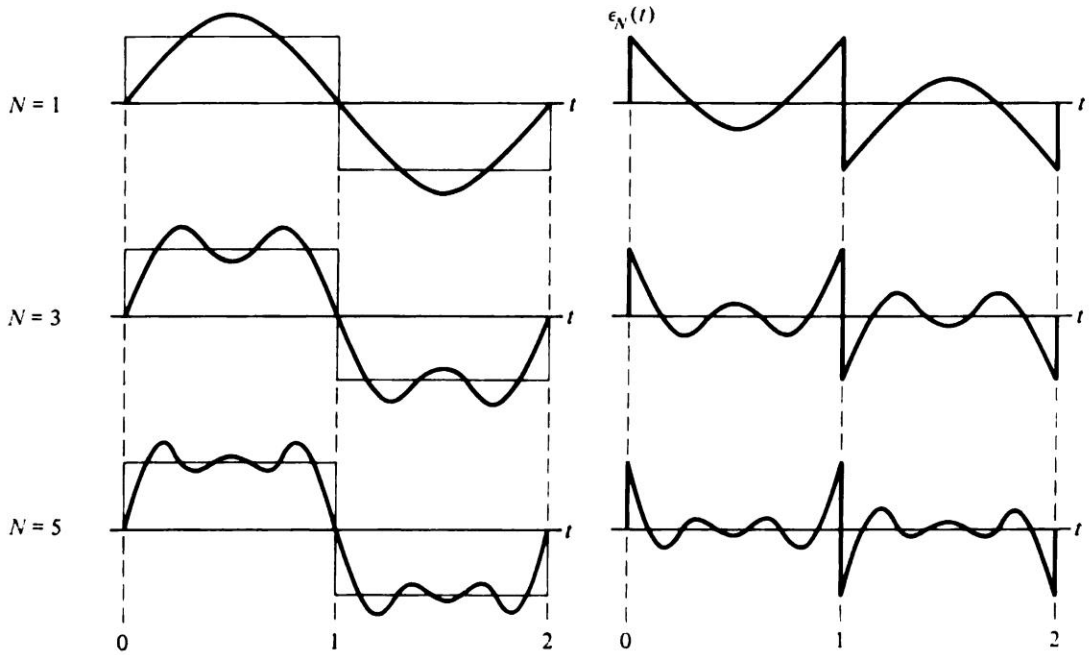


Figure 2.7 Finite sum of orthogonal functions to approximate a rectangular function and the instantaneous (point) error remaining after each successive approximation.

TABLE 2.1 Error Energy for Example 2.5.1

N	Error Energy	Percent Energy
1	0.379	18.94
3	0.199	9.94
5	0.134	6.69
7	0.101	5.04
9	0.081	4.04
11	0.067	3.37

DRILL PROBLEM 2.5.1

Find the generalized Fourier representation of the function

$$f(t) = \begin{Bmatrix} 1 & -2 < t < 2 \\ 0 & \text{elsewhere} \end{Bmatrix}$$

over the interval $(-4, 4)$ using the set of orthogonal functions

$$\phi_n(t) = \cos n\frac{\pi}{4}t, \qquad n = 0, 1, 2, \ldots.$$

ANSWER:

$$f(t) = \frac{2}{\pi}\left(\frac{\pi}{4} + \cos\frac{\pi}{4}t - \frac{1}{3}\cos\frac{3\pi}{4}t + \frac{1}{5}\cos\frac{5\pi}{4}t - \frac{1}{7}\cos\frac{7\pi}{4}t + \cdots\right).$$

EXAMPLE 2.5.2

Making use of the analogy between signals and vectors, show that, for two finite-energy signals $x(t), y(t)$,

$$\left|\int_{-\infty}^{\infty} x(t)y(t)\,dt\right|^2 \le \int_{-\infty}^{\infty} |x(t)|^2\,dt \int_{-\infty}^{\infty} |y(t)|^2\,dt.$$

Solution Consider the vector representation $\mathbf{x}$ and $\mathbf{y}$. The scalar product $\mathbf{x}\cdot\mathbf{y}$ cannot exceed the product of the lengths of $\mathbf{x}$ and $\mathbf{y}$. If θ is the included angle, then

$$|\mathbf{x}\cdot\mathbf{y}|^2 = (|\mathbf{x}|\,|\mathbf{y}|\cos\theta)^2 \le |\mathbf{x}|^2|\mathbf{y}|^2.$$

The equality holds here if and only if $\cos\theta = \pm 1$, i.e., if the vectors are collinear.

Using a signal-space analogy, the above inequality can be rewritten in terms of the inner products of $x(t)$ and $y(t)$:

$$\left|\int_{-\infty}^{\infty} x(t)y^*(t)\,dt\right|^2 \le \int_{-\infty}^{\infty} |x(t)|^2\,dt \int_{-\infty}^{\infty} |y(t)|^2\,dt.$$

Consistent with x and y being collinear, the equality condition is given by $x(t) = Ky(t)$, K = any real-valued constant.

This result is usually rewritten by letting $z(t) = y^*(t)$ so that

$$\left|\int_{-\infty}^{\infty} x(t)z(t)\,dt\right|^2 \le \int_{-\infty}^{\infty} |x(t)|^2\,dt \int_{-\infty}^{\infty} |z(t)|^2\,dt,$$

with equality if and only if $x(t) = Kz^*(t)$. This well-known and useful relationship is called the *Schwarz inequality*.

2.6 ★ CHOICE OF A SET OF ORTHOGONAL FUNCTIONS

We have discussed the generalized Fourier series representation of a function of finite energy over a given interval by a linear combination of mutually orthogonal functions. However, many sets of orthogonal functions exist and hence a given function may be expressed in terms of different sets of orthogonal functions. This should not surprise us for, in the analogy to vector spaces, we know that a given vector can be expressed uniquely in different coordinate systems. Often the choice of a specific

coordinate system in vector space will make the analysis of a given type of problem easier and a good portion of our experience in vectors is to learn to make certain judicious choices in coordinates and coordinate systems. This same conclusion holds true for signal spaces.

Some examples of sets of orthogonal functions are exponential functions, trigonometric functions, Walsh functions, and Legendre polynomials. We shall not launch into an intensive investigation of each set of orthogonal functions and its properties. (Some are introduced in the problems.) Because we are primarily concerned in applying our results, we ask, Out of all the possible sets of orthogonal functions, which set will be the most convenient one for us to choose? Again, as in vector spaces, the answer to this question depends upon the type of problem to be solved. Therefore we will digress here and, with a little foresight, attempt to make a judicious choice at the outset.

We shall be concerned with systems that are linear and those whose parameters do not vary as a function of time, at least not within given intervals of time. When a signal is applied to such a system, the output can be expressed in terms of linear differential equations with constant coefficients. As an example, consider the R-L electrical circuit shown in Fig. 2.8 describing a particular system in the above sense. The relationship between the input voltage, $v_i(t)$, and the output voltage, $v_o(t)$, can be expressed as†

$$\frac{d}{dt}v_o(t) + \frac{R}{L}v_o(t) = \frac{R}{L}v_i(t).$$

For a finite-energy input, we could then expect to solve this differential equation to find the output. Unless the input signal is a very simple function of t, however, we have difficulty solving the differential equations encountered and so we turn to writing a representation of the signal in terms of orthogonal functions. However, our choice of a representation for the input signal will affect the difficulty which we may encounter in solving the problem.

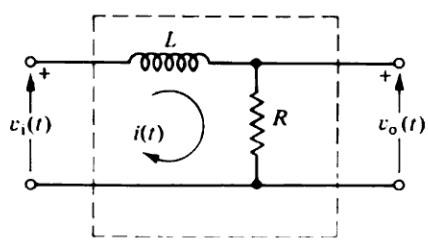


Figure 2.8 A linear system.

† Because the current is common to both R and L, we can write

$$\frac{1}{L}\int_0^t [v_i(t) - v_o(t)]\,dt = v_o(t)/R.$$

Taking the derivative of both sides and multiplying by R yields the desired differential equation.

Now we entertain an intriguing possibility: if we could choose a function that repeats itself under the operation of differentiation, then such a function at the input will yield the same function at the output multiplied by some algebraic polynomial in terms of the parameters of the function and the differential equation. Furthermore, upon cancelling out the original function from both sides of the differential equation, this algebraic equation will be *only a function of the parameters of the system and the parameters of the signal and not the input signal itself*!

But what type of function demonstrates this type of behavior? The simplest elementary function which repeats itself under the operation of differentiation is that of the complex exponential form $e^{\pm(\sigma+j\omega)t} = e^{\pm st}$, where σ, ω are parameters independent of t. It can easily be verified that such an input signal, when applied to a system that can be described in terms of a differential equation with constant coefficients, will yield an output function of an identical form multiplied by an algebraic equation in $s = \sigma + j\omega$ and the constants of the system, because

$$\frac{d^n}{dt^n}(e^{st}) = s^n e^{st}.$$

We are interested in expressing finite-energy signals that may exist for negative as well as positive time. Because the real part, σ, of the exponential may cause convergence problems in such cases, we prefer to set $\sigma = 0$ and thus use functions of the form $e^{\pm j\omega t}$. These functions, or linear combinations of them, can exist for all time and analyses using them are usually referred to as "steady-state."

The parameter ω is the rate of phase change or frequency of the complex exponential in radians per second. It may also be expressed as f in cycles per second (Hz), where $\omega = 2\pi f$. Expressions for signals in terms of frequency are said to be in the "frequency domain," whereas expressions in time are said to be in the "time domain."

DRILL PROBLEM 2.6.1

Letting $v_i(t) = V_i e^{st}$ and $v_o(t) = V_o e^{st}$, determine the ratio of V_o/V_i in terms of s for the circuit shown in Fig. 2.8.

ANSWER: $(R/L)/(s + R/L)$.

2.7 THE EXPONENTIAL FOURIER SERIES

We investigate a set of complex-valued exponential functions expressed as

$$\phi_n(t) = e^{jn\omega_0 t}, \tag{2.31}$$

where n is an integer ($n = 0, \pm 1, \pm 2, \ldots$) and ω_0 is a constant to be determined. The value of n is referred to as the harmonic number or *harmonic*.

We consider the following operation on $\phi_n(t)$:

$$\int_{t_1}^{t_2} \phi_n(t)\phi_m^*(t)\,dt = \int_{t_1}^{t_2} e^{jn\omega_0 t} e^{-jm\omega_0 t}\,dt$$

$$= \frac{1}{j(n-m)\omega_0}\left[e^{j(n-m)\omega_0 t_2} - e^{j(n-m)\omega_0 t_1}\right], \qquad n \neq m,$$

$$= \frac{1}{j(n-m)\omega_0} e^{j(n-m)\omega_0 t_1}\left[e^{j(n-m)\omega_0(t_2-t_1)} - 1\right]. \tag{2.32}$$

Except for the trivial case of $t_2 = t_1$, we can force the term within brackets to zero if we choose [since $(n - m)$ is an integer]

$$\omega_0(t_2 - t_1) = 2\pi$$

so that

$$\int_{t_1}^{t_2} e^{jn\omega_0 t} e^{-jm\omega_0 t}\,dt = \begin{Bmatrix} (t_2 - t_1) & n = m \\ 0 & n \neq m \end{Bmatrix}, \tag{2.33}$$

if

$$\omega_0 = \frac{2\pi}{(t_2 - t_1)}. \tag{2.34}$$

If two functions $\phi_n(t), \phi_m(t)$ meet the condition

$$\int_{t_1}^{t_2} \phi_n(t)\phi_m^*(t)\,dt = \begin{Bmatrix} \text{constant} & n = m \\ 0 & n \neq m \end{Bmatrix},$$

then these two functions are *orthogonal* over the interval (t_1, t_2). Therefore the set of functions

$$\phi_n(t) = e^{jn\omega_0 t} \qquad n = 0, \pm 1, \pm 2, \ldots$$

forms an orthogonal set over the interval (t_1, t_2) if $\omega_0 = 2\pi/(t_2 - t_1)$.

Next we seek to express an arbitrary signal $f(t)$ in terms of a finite set of complex exponentials by writing

$$f(t) = \sum_{n=-N}^{N} F_n e^{jn\omega_0 t} \qquad (t_1 < t < t_2)$$

where the coefficients F_n are to be determined. It can be shown that the error energy between $f(t)$ and its approximation in terms of a set of complex exponentials decreases to zero as the number of terms taken approaches infinity. When a set of $\phi_n(t)$ meets this condition, it is said to be *complete*. Forming the set, we have

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t} \qquad (t_1 < t < t_2). \tag{2.35}$$

Because this $\phi_n(t)$ forms a complete orthogonal set, it is possible to represent any arbitrary complex-valued function $f(t)$ with finite energy by a linear combination of

complex exponential functions over an interval (t_1, t_2). The representation of $f(t)$ by the exponential series as demonstrated above in Eq. (2.35) is known as the *exponential Fourier series representation* of $f(t)$ over the interval (t_1, t_2). The coefficients in this series can be found by multiplying both sides of Eq. (2.35) by $\phi_m^*(t) = e^{-jm\omega_0 t}$ and integrating with respect to t over the interval (t_1, t_2). As a result of orthogonality, all terms on the right-hand side of Eq. (2.35) vanish except for the one for $m = n$ and this yields the expression†

$$F_n = \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} f(t) e^{-jn\omega_0 t}\, dt. \tag{2.36}$$

In summary, we have found that any given function having finite energy may be expressed as a discrete sum of exponential functions: $\phi_n(t) = e^{jn\omega_0 t}$ over an interval (t_1, t_2), where n is an integer ($n = 0, \pm 1, \pm 2, \ldots$) and where $\omega_0 = 2\pi/(t_2 - t_1)$.

EXAMPLE 2.7.1

Write a representation of the function $f(t)$ given in Example 2.5.1 in terms of the complex exponential Fourier series.

Solution

$$\omega_0 = \frac{2\pi}{t_2 - t_1} = \frac{2\pi}{2} = \pi,$$

$$F_n = \frac{1}{2} \int_0^2 f(t) e^{-jn\pi t}\, dt$$

$$= \frac{1}{2} \int_0^1 e^{-jn\pi t}\, dt - \frac{1}{2} \int_1^2 e^{-jn\pi t}\, dt$$

$$= \frac{1}{2jn\pi}[-e^{-jn\pi} + 1 + e^{-j2n\pi} - e^{-jn\pi}] = \frac{1}{jn\pi}[1 - e^{-jn\pi}],$$

$$F_n = \begin{cases} 2/jn\pi & n \text{ odd} \\ 0 & n \text{ even} \end{cases},$$

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t},$$

$$f(t) = \frac{2}{j\pi}\left(e^{j\pi t} + \frac{1}{3}e^{j3\pi t} + \frac{1}{5}e^{j5\pi t} + \cdots - e^{-j\pi t} - \frac{1}{3}e^{-j3\pi t} - \frac{1}{5}e^{-j5\pi t} - \cdots\right).$$

† Note that as a result of orthogonality each coefficient in the series is independent of every other coefficient. The coefficients could be found also by a direct application of Eq. (2.26).

DRILL PROBLEM 2.7.1

Write a representation of the function $f(t)$ given in Drill Problem 2.5.1 in terms of the complex exponential Fourier series over the interval $(-4, 4)$.
ANSWER:

$$f(t) = \frac{1}{\pi}\left(\frac{\pi}{2} + e^{j(\pi/4)t} - \frac{1}{3}e^{j(3\pi/4)t} + \cdots + e^{-j(\pi/4)t} - \frac{1}{3}e^{-j(3\pi/4)t} + \cdots\right).$$

2.8 COMPLEX SIGNALS AND REPRESENTATIONS

For a signal whose instantaneous value is a complex number, the real and imaginary parts form an ordered pair of components given by (suppressing the time dependence for convenience in writing)

$$f = f_r + jf_i. \tag{2.37}$$

The complex conjugate of the signal value is

$$f^* = f_r - jf_i, \tag{2.38}$$

and therefore the real and imaginary parts are

$$f_r = \frac{1}{2}(f + f^*) \tag{2.39}$$

$$f_i = \frac{1}{2j}(f - f^*). \tag{2.40}$$

The squared magnitude of the complex-valued signal is equal to the product of the signal and its complex conjugate. Using the fact that $(j)^2 = -1$, we have

$$|f|^2 = ff^* = |f_r|^2 + |f_i|^2, \tag{2.41}$$

so that the square of the signal magnitude is the sum of the squares of the magnitudes of the real and imaginary components.

In the mathematics of linear systems analysis, we are particularly interested in signals of the complex exponential form $f(t) = e^{j\omega_0 t}$. Physically, this particular function may be thought of as describing the motion of a point on the rim of a wheel of unit radius. The wheel revolves (by convention) counterclockwise at an angular rate of ω_0 radians per second. From trigonometry (see Fig. 2.9), we see that the projection of the point on the real axis is $\cos \omega_0 t$, and the projection on the imaginary axis is $\sin \omega_0 t$. Combining, we can write

$$e^{j\omega_0 t} = \cos \omega_0 t + j \sin \omega_0 t. \tag{2.42}$$

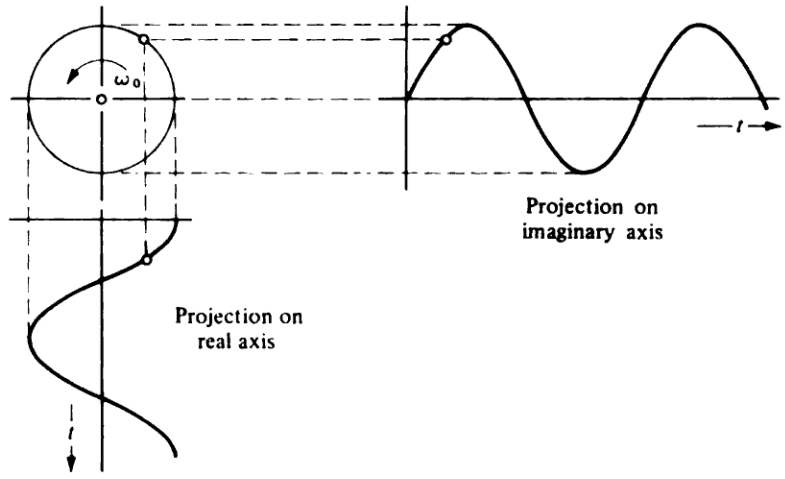


Figure 2.9 Complex notation for angular signals.

An application of Eqs. (2.39) and (2.40) yields the real and imaginary components of the left-hand side of this result. Then, equating the real and imaginary components of both sides of Eq. (2.42), we obtain "Euler's identities":

$$\cos \omega_0 t = \frac{1}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t}), \tag{2.43}$$

$$\sin \omega_0 t = \frac{1}{2j}(e^{j\omega_0 t} - e^{-j\omega_0 t}). \tag{2.44}$$

The addition of several terms, each involving circular motion with differing magnitudes and angular rates, makes the problem a little more complicated. However, we find that if an arrow is drawn from the origin to the reference point on each circle (wheel), the sum can be found by adding the arrows vectorially just as if they were vector quantities. Further, the resultant arrow may be treated just as if it were a vector and translated in coordinates so that its tail is at the origin. The projections on the translated coordinate system describe the real and imaginary components of the resulting summation.

These arrows are called *phasors*. A phasor is specified in the complex plane either by its real and imaginary components or by its magnitude and phase angle. Phasors obey the rules of vectors but they are not vectors; for this reason they are sometimes called pseudovectors.

The exponential Fourier series consists of a summation of complex exponential terms, each with its own magnitude, phase, and angular rate (frequency). Therefore summations of phasors can describe the instantaneous values of complex-valued signals using the exponential Fourier series representation. The usual convention is to let the real part of the phasor represent the real-valued function, although this certainly is not necessary. (We shall interchange the roles in the example below.)

With these added insights, let us reconsider the results of Example 2.7.1. There we obtained the complex exponential Fourier representation of a given function $f(t)$ which resulted in a solution in terms of an infinite series of complex exponentials rotating at integer multiples of π radians per second. Each term in the series is a phasor and the complex coefficients, the F_n, represent the starting angle and magnitude of each phasor. These phasors can be added using the rules of vector addition, and their sum represents the instantaneous complex amplitude and phase of the original function.

For convenience, let us retain only the first three nonzero terms of the series and consider only the interval $0 < t < 0.5$. From the results of Example 2.7.1, the first three terms for positive n are

$$\frac{2}{j\pi}\left(e^{j\pi t} + \frac{1}{3}e^{j3\pi t} + \frac{1}{5}e^{j5\pi t}\right).$$

As a result of the factor of $(-j)$, all three phasors start on the negative imaginary axis so that their projections on the real axis start at zero. For convenience in graphing, let us rotate everything by 90° and take the projections on the imaginary axis.

The phasors are shown in Fig. 2.10. At $t = 0$, all three phasors lie on the axis because the starting phase angles are all zero. As t increases, each of the phasors rotates at different angular rates and the resultant begins to curl in upon itself. The vertical components of the resultant describe the values of the signal representation at the corresponding times marked. Each phasor magnitude is labeled with the particular value of n corresponding to it.

Figure 2.10 is a graphic illustration of the way in which a vector summation of phasors, spinning at different angular rates, can add up in such a manner that the projection of their sum closely approximates the desired signal. As more and more terms are added to the series, the number of phasors in the sum becomes arbitrarily large. In this way, the tip of the resultant can move very quickly to describe the desired function.

Note that we summed over only positive values of n in the example above and then took the real part. An alternative to this procedure is to use a conjugate pair of phasors. Thus a cosine (real part) or a sine (imaginary part) function may be considered as the resultant of two conjugate phasors, rotating in opposite directions as shown in Fig. 2.11.

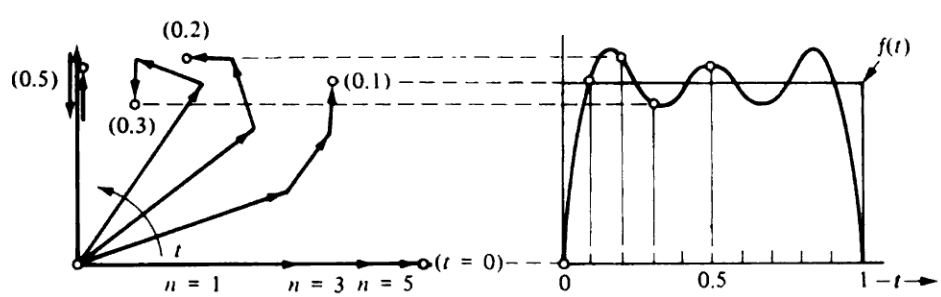


Figure 2.10 Phasor addition of complex exponential series for function described in Example 2.7.1.

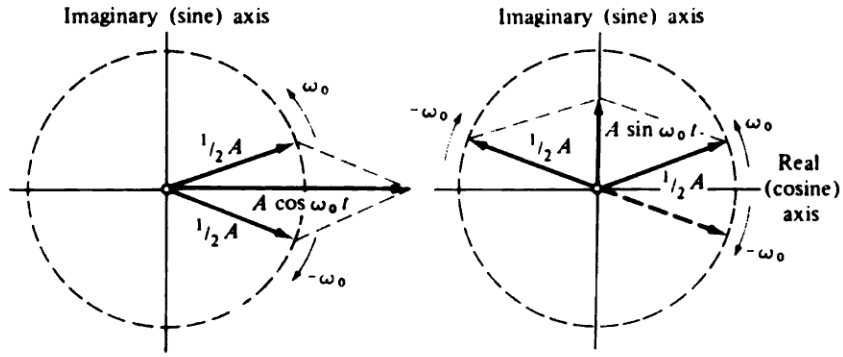


Figure 2.11 Representation of a cosine and sine by a pair of conjugate phasors.

In both cases the complex conjugate phasors rotate in opposite directions with angular frequencies of $\pm\omega_0$ radians per second to describe real-valued functions of time. Because a *positive frequency* is associated with counterclockwise uniform angular motion, this representation suggests the concept of a *negative frequency* to describe the motion of the second phasor. The use of negative frequencies is a convenient way to describe a real-valued function using pairs of complex-valued functions. It is equivalent to taking the real (or imaginary) part of the complex-valued function. The concept of negative frequencies is used throughout this book, largely for mathematical convenience. We must realize, of course, that to represent a real-valued signal the magnitudes of two conjugate phasors must be equal and their instantaneous phase angles must be equal-but-oppositely directed.

DRILL PROBLEM 2.8.1

Using Eqs. (2.37)–(2.40), show that

$$\mathcal{R}e\{f_1 f_2\} = \mathcal{R}e\{f_1\}\mathcal{R}e\{f_2\} - \mathcal{I}m\{f_1\}\mathcal{I}m\{f_2\}.$$

2.9 THE TRIGONOMETRIC FOURIER SERIES REPRESENTATION†

The question may well be asked at this point: If we know that a given function $f(t)$ is real-valued, isn't there an equivalent way of expressing a Fourier series representation of $f(t)$ using a set of real-valued orthogonal functions? Let us investigate by taking the

† J. B. J. Fourier, *Theorie Analytique de la Chaleur*, 1822; republished in English by Dover (New York). The first rigorous proof of a version of Fourier's theorem was given by Dirichlet in 1829.

real part of both sides of Eq. (2.35). Making use of the result of Drill Problem 2.8.1, we have

$$f_r(t) = \sum_{n=-\infty}^{\infty} \mathscr{R}e\{F_n\}\mathscr{R}e\{e^{jn\omega_0 t}\} - \sum_{n=-\infty}^{\infty} \mathscr{I}m\{F_n\}\mathscr{I}m\{e^{jn\omega_0 t}\}$$

$$= \sum_{n=-\infty}^{\infty} \mathscr{R}e\{F_n\} \cos n\omega_0 t - \sum_{n=-\infty}^{\infty} \mathscr{I}m\{F_n\} \sin n\omega_0 t . \tag{2.45}$$

The F_n are given by Eq. (2.36). But since we need real-valued functions, we will express new real-valued coefficients in terms of F_n and F_n^*. Therefore we investigate the F_n^* for real-valued signals:

$$F_n^* = \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} f(t) e^{jn\omega_0 t} dt = F_{-n} . \tag{2.46}$$

Recalling Eqs. (2.39) and (2.40), and using Eq. (2.46), we can write

$$\mathscr{R}e\{F_n\} = \tfrac{1}{2}[F_n + F_n^*] = \tfrac{1}{2}[F_n + F_{-n}] , \tag{2.47}$$

$$\mathscr{I}m\{F_n\} = \frac{1}{2j}[F_n - F_n^*] = \frac{1}{2j}[F_n - F_{-n}] . \tag{2.48}$$

For convenience, let us define new coefficients at this point:

$$a_0 \triangleq F_0 , \tag{2.49}$$

$$a_n \triangleq [F_n + F_{-n}] = 2\ \mathscr{R}e\{F_n\}, \qquad n \neq 0 , \tag{2.50}$$

$$b_n \triangleq j[F_n - F_{-n}] = -2\ \mathscr{I}m\{F_n\} , \tag{2.51}$$

$$F_n = \tfrac{1}{2}(a_n - jb_n), \qquad n \neq 0 . \tag{2.52}$$

From the definitions of a_n, b_n, it can be seen that a_n is an even function of n while b_n is an odd function of n. Thus everything is completely symmetric in n about $n = 0$, and we could just as well double the coefficients and sum only over positive integer values of n. Following this procedure, Eq. (2.45) can be rewritten as

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega_0 t + \sum_{n=1}^{\infty} b_n \sin n\omega_0 t \qquad \text{for } f(t) \text{ real-valued over } (t_1, t_2) . \tag{2.53}$$

Thus the set of functions $\cos n\omega_0 t$ and $\sin n\omega_0 t$ for $(n = 0, 1, 2, \ldots)$, $\omega_0 = 2\pi/(t_2 - t_1)$, forms a complete orthogonal set over the interval (t_1, t_2). The result given in Eq. (2.53) is called the *trigonometric Fourier series representation of $f(t)$ over the interval* (t_1, t_2). It is capable of describing a real-valued function of finite energy over a given interval. Note that although the functions $\sin \omega_0 t$, $\sin 2\omega_0 t$, etc., form an orthogonal set over any interval $[t_1, t_1 + (2\pi/\omega_0)]$, this set is not complete. This arises from the fact that we can exhibit a function, namely $\cos n\omega_0 t$, which is orthogonal to $\sin n\omega_0 t$ over the same interval. Hence to complete the set we must include cosine as well as sine functions.

Rather than evaluating the constants a_n and b_n from the F_n, we could have multiplied both sides of Eq. (2.53) by $\cos n\omega_0 t$ and $\sin n\omega_0 t$. Because $\cos n\omega_0 t$, $\cos m\omega_0 t$, $\sin n\omega_0 t$, and $\sin m\omega_0 t$ are all mutually orthogonal, the terms remaining after simplification are†

$$a_n = \frac{\int_{t_1}^{t_2} f(t) \cos n\omega_0 t \, dt}{\int_{t_1}^{t_2} \cos^2 n\omega_0 t \, dt} = \frac{2}{(t_2 - t_1)} \int_{t_1}^{t_2} f(t) \cos n\omega_0 t \, dt, \tag{2.54}$$

$$b_n = \frac{\int_{t_1}^{t_2} f(t) \sin n\omega_0 t \, dt}{\int_{t_1}^{t_2} \sin^2 n\omega_0 t \, dt} = \frac{2}{(t_2 - t_1)} \int_{t_1}^{t_2} f(t) \sin n\omega_0 t \, dt, \tag{2.55}$$

$$a_0 = \frac{\int_{t_1}^{t_2} f(t) \, dt}{\int_{t_1}^{t_2} dt} = \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} f(t) \, dt. \tag{2.56}$$

The trigonometric Fourier series may be represented in more compact notational form as

$$f(t) = \sum_{n=0}^{\infty} c_n \cos (n\omega_0 t + \phi_n), \tag{2.57}$$

where

$$c_n = \sqrt{a_n^2 + b_n^2}, \tag{2.58}$$

$$\phi_n = \tan^{-1} (-b_n/a_n). \tag{2.59}$$

Substituting Eqs. (2.50) and (2.51) into Eqs. (2.58) and (2.59), we find that this trigonometric representation is related to the complex exponential representation by

$$c_n = 2|F_n| = 2\sqrt{F_n F_n^*}, \qquad n \neq 0, \tag{2.60}$$

$$\phi_n = \tan^{-1} \frac{\mathcal{I}m\{F_n\}}{\mathcal{R}e\{F_n\}}, \tag{2.61}$$

and $c_0 = F_0$.

† These coefficients could be found also by a direct application of Eq. (2.26).

EXAMPLE 2.9.1

Represent $f(t) = t^2$ in a trigonometric Fourier series over the interval $(0, 2)$.

Solution In this case, $t_1 = 0$, $t_2 = 2$, and $\omega_0 = \pi$. The coefficients can be found by a direct application of Eqs. (2.54)–(2.56):[†]

$$a_0 = \frac{1}{2}\int_0^2 t^2\,dt = \frac{4}{3},$$

$$a_n = \frac{2}{2}\int_0^2 t^2 \cos n\pi t\,dt = 4/(n\pi)^2,$$

$$b_n = \frac{2}{2}\int_0^2 t^2 \sin n\pi t\,dt = -4/(n\pi).$$

Thus the trigonometric Fourier series of $f(t) = t^2$ over the interval $(0, 2)$ is

$$f(t) = \frac{4}{3} + \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\cos n\pi t - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin n\pi t.$$

Any function $f(t)$ can be expressed in terms of a sum of a corresponding even function, $f_e(t)$, and an odd function, $f_o(t)$. These functions can be formed by the following relations:

$$f_e(t) = \tfrac{1}{2}[f(t) + f(-t)], \tag{2.62}$$

$$f_o(t) = \tfrac{1}{2}[f(t) - f(-t)], \tag{2.63}$$

and thus $f_e(t) + f_o(t) = f(t)$.

The properties of even and odd functions are particularly convenient when one considers the trigonometric Fourier series over a symmetric interval $(-T/2, T/2)$. In this case Eq. (2.53) may reduce to a special form depending on whether $f(t)$ is even or odd. If $f(t)$ is even, the product $f(t)\sin n\omega_0 t$ is odd in t, the b_n's are zero, and the result is a cosine series. Similarly, if $f(t)$ is odd, then the product $f(t)\cos n\omega_0 t$ is odd in t, the a_n's are zero, and the result is a sine series. In both cases the nonzero terms can be evaluated by integrating over half the period and then multiplying by two.

DRILL PROBLEM 2.9.1

Find the trigonometric Fourier series of the triangular waveform shown in Fig. 2.12 over the interval $(-\pi, \pi)$.

[†] These integrals can be found in Appendix A.

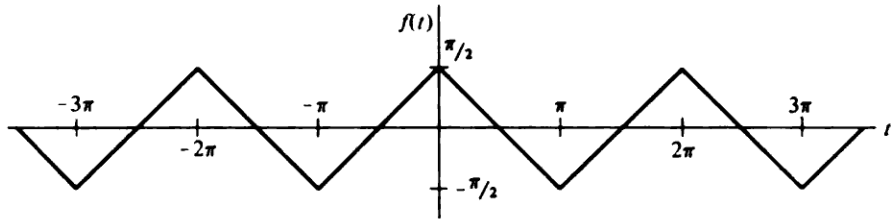


Figure 2.12 A triangular waveform.

ANSWER: $f(t) = \dfrac{4}{\pi} \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \dfrac{1}{n^2} \cos nt \qquad -\pi < t < \pi .$

DRILL PROBLEM 2.9.2

Find the expansion of the function described by $f(t) = t$ in a trigonometric Fourier series over the interval $(0, 2)$.

ANSWER: $f(t) = 1 - \dfrac{2}{\pi} \sum_{n=1}^{\infty} \dfrac{1}{n} \sin n\pi t \qquad 0 < t < 2 .$

2.10 EXTENSION BY PERIODICITY

We have been able to represent a given function with finite energy by a Fourier series over a finite interval (t_1, t_2). Outside this interval the function $f(t)$ and the corresponding Fourier series need not be equal.

Now we would like to extend this representation to periodic signals, i.e., those for which the relation

$$f(t + T) = f(t) \tag{2.64}$$

holds for all t. We shall assume that $f(t)$ has finite energy over an interval $(t_0, t_0 + T)$. As we consider longer and longer time intervals, the energy accumulates and in the limit the total energy is infinite. However, we also note that the energy in $f(t)$ in any interval T seconds long is the same as in any other given interval which is T seconds in length. It follows, then, that if we divide by T, the average *rate* of energy, or average power, is a constant.

With this slight change, we can extend all the previous analyses directly over to the case of periodic functions with finite average power simply by (1) taking the interval for finding the coefficients, etc., as the period T of the periodic function; and (2) dividing by this interval in any calculation involving the energy so that we obtain

an average energy rate or power. Such representations are said to "converge in mean square" to the periodic function $f(t)$.

It is hardly worth going back and rephrasing all our previous results for the case of the series representation of a periodic function. In fact, for the complex exponential set we chose earlier, it is easy to see that Eq. (2.64) is satisfied if the interval is $T = 2\pi/\omega_0$ because

$$e^{jn\omega_0(t+T)} \equiv e^{jn\omega_0 t} \qquad \text{for } T = 2\pi/\omega_0 .$$

It follows that the series representation

$$f(t) = \sum_n F_n e^{jn\omega_0 t}$$

will represent a periodic function over the infinite interval and the representation converges in a mean-square (or average-square) sense.

Graphically, we are saying that since the basic phasor rotates at the angular rate ω_0 rad/sec, the complex exponential Fourier series representation will describe any periodic function whose period is just that time it takes the basic phasor to make one complete revolution. This is shown in Fig. 2.13.

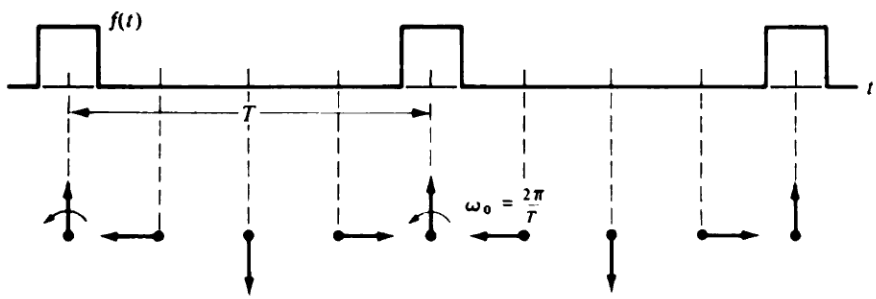


Figure 2.13 Revolution of a basic phasor in synchronism with a given periodic function.

In general the interval of integration to determine coefficients in the Fourier series is taken over one complete period T to evaluate a given coefficient. Just where the particular interval is taken, however, makes no difference when $f(t)$ is periodic. If the lower limit has an arbitrary value of t_0, then the upper limit will be $(t_0 + T)$. Often it is convenient to take the interval of integration from $-T/2$ to $+T/2$ in order to recognize possible symmetry conditions.

DRILL PROBLEM 2.10.1

Determine the trigonometric Fourier series for the symmetric square wave shown in Fig. 2.14.

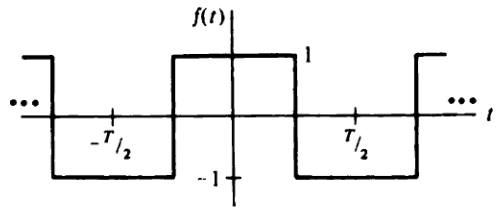


Figure 2.14 The symmetric square wave.

ANSWER:

$$f(t) = \frac{4}{\pi}\left(\cos \omega_0 t - \frac{1}{3}\cos 3\omega_0 t + \frac{1}{5}\cos 5\omega_0 t - \cdots\right) \quad \text{where } \omega_0 = 2\pi/T.$$

Fourier series coefficients for some selected waveforms commonly used are listed in Table 2.2. A useful exercise, if a computer and a graphics monitor are available, is

TABLE 2.2 Fourier Series Coefficients for Selected Waveforms

Periodic Waveform	Symmetry	F_n†
1. *Symmetric square wave*	*even*	
$\begin{cases} +1 & \lvert t\rvert < T/4 \\ -1 & T/4 \le \lvert t\rvert < T/2 \end{cases}$		$\begin{cases} \mathrm{Sa}(n\pi/2) & n \neq 0 \\ 0 & n = 0 \end{cases}$
2. *Rectangular pulse train*	*even*	
$\begin{cases} +1 & \lvert t\rvert < \tau/2 \\ 0 & \tau/2 \le \lvert t\rvert < T/2 \end{cases}$		$\frac{\tau}{T}\,\mathrm{Sa}(n\pi\tau/T)$
3. *Symmetric triangular wave*	*even*	
$1 - 4\lvert t\rvert/T, \quad \lvert t\rvert < T/2$		$\begin{cases} \mathrm{Sa}^2(n\pi/2) & n \neq 0 \\ 0 & n = 0 \end{cases}$
4. *Symmetric sawtooth wave*	*odd*	
$2t/T, \quad \lvert t\rvert < T/2$		$\begin{cases} j(-1)^n/(n\pi) & n \neq 0 \\ 0 & n = 0 \end{cases}$
5. *Half-wave rectified sinusoid*		
$\begin{cases} \sin \omega_0 t & 0 \le t < T/2 \\ 0 & -T/2 \le t < 0 \end{cases}$		$\begin{cases} \frac{1}{\pi(1-n^2)} & n \text{ even} \\ -j/4 & n = \pm 1 \\ 0 & \text{otherwise} \end{cases}$
6. *Full-wave rectified sinusoid*	*even*	
$\lvert \sin \omega_0 t\rvert$		$\begin{cases} \frac{2}{\pi(1-n^2)} & n \text{ even} \\ 0 & \text{otherwise} \end{cases}$

† $\mathrm{Sa}(x) \triangleq (\sin x)/x$.

to observe reconstructions of these waveforms from the series representations (e.g., convert to the trigonometric series and use 50 points per period, 15 harmonics for the reconstructions).

2.11 PARSEVAL'S THEOREM FOR POWER SIGNALS

The average power developed across a 1-ohm resistance is [Eq. (2.7)]

$$P = \frac{1}{T}\int_{-T/2}^{T/2} f(t)f^*(t)\,dt \qquad \text{watts}. \tag{2.65}$$

Using the exponential Fourier series and substituting in Eq. (2.65),

$$P = \frac{1}{T}\int_{-T/2}^{T/2} \sum_{m=-\infty}^{\infty} F_m e^{jm\omega_0 t} \sum_{n=-\infty}^{\infty} F_n^* e^{-jn\omega_0 t}\,dt \tag{2.66}$$

where $\omega_0 = 2\pi/T$.

Assuming that $f(t)$ is integrable over the interval t_0 to $(t_0 + T)$, we may interchange the order of summation and integration in Eq. (2.66):

$$P = \sum_{m=-\infty}^{\infty} F_m \sum_{n=-\infty}^{\infty} F_n^* \frac{1}{T}\int_{-T/2}^{T/2} e^{j(m-n)\omega_0 t}\,dt. \tag{2.67}$$

But the complex exponential functions are orthogonal over the interval t_0 to $(t_0 + T)$ so that the integral in Eq. (2.67) is zero except for the special case when $m = n$. For this specific condition the double summation reduces to a single summation and we have a new relation for the average power in terms of the magnitudes of the coefficients:

$$P = \sum_{n=-\infty}^{\infty} F_n F_n^* = \sum_{n=-\infty}^{\infty} |F_n|^2. \tag{2.68}$$

Combining Eqs. (2.65) and (2.68), we obtain a relationship which is known as *Parseval's theorem for periodic signals:*

$$P = \frac{1}{T}\int_{-T/2}^{T/2} |f(t)|^2\,dt = \sum_{n=-\infty}^{\infty} |F_n|^2. \tag{2.69}$$

If we know the time function $f(t)$, we can find the average power. Alternatively, if we know the Fourier coefficients, we can find the average power. The answer obtained in the time domain and in the frequency domain must agree.

EXAMPLE 2.11.1

Determine the average power of $f(t) = 2 \sin 100t$ using Eqs. (2.65) and (2.68).

Solution Using Eq. (2.65), we have

$$P = \frac{1}{T}\int_{-T/2}^{T/2} 4 \sin^2 100t\, dt = 2 \text{ W}.$$

The Fourier coefficients of $f(t)$ are†

$$F_1 = -j, \qquad F_{-1} = j,$$
$$F_n = 0 \quad \text{for all } n \neq \pm 1.$$

Then Eq. (2.68) gives

$$P = \sum_{n=-\infty}^{\infty} |F_n|^2 = F_1 F_1^* + F_{-1} F_{-1}^* = |j|^2 + |-j|^2 = 1 + 1 = 2 \text{ W}.$$

Equation (2.68) illustrates the fact that the power in a periodic function is distributed over discrete frequencies that are harmonically related to one another. The power contained in each frequency component is given by its respective term in Eq. (2.68). Thus if we make a graph of P versus ω, the power in $f(t)$ is located only at discrete frequencies, as depicted in Fig. 2.15. Such a graph is called the "power spectrum" of the signal $f(t)$.

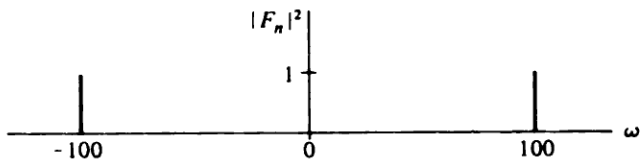


Figure 2.15 The power spectrum of the function described in Example 2.11.1

DRILL PROBLEM 2.11.1

Repeat Example 2.11.1 for $f(t) = 2 \cos 100t$ and show that its power spectrum is given by Fig. 2.15.

† A shortcut in the determination of these coefficients is to use Euler's identity for $2 \sin 100t$. The coefficients are found by equating term by term to the corresponding Fourier series.

2.12 THE FREQUENCY TRANSFER FUNCTION

In characterizing the behavior of systems, we follow two main plans of analysis, one in the frequency domain and one in the time domain. In both approaches, we use superposition to add up the responses of the system for combinations of elemental functions. Representation in the frequency domain is considered first; the time domain is considered later in Section 2.18. Following the reasoning of the previous sections, we make use of the complex exponential function.

A fundamental property of a linear time-invariant system is that the input and the output are related by linear differential equations with constant coefficients. A typical system could be described by

$$a_0 g(t) + a_1 \frac{dg}{dt} + \cdots = b_0 f(t) + b_1 \frac{df}{dt} + \cdots, \tag{2.70}$$

where the a's and b's are constants.

Now we use the input signal

$$f(t) = e^{j\omega t} \tag{2.71}$$

to test the system. A particular solution can be written as

$$g(t) = H(\omega) e^{j\omega t}. \tag{2.72}$$

Using Eqs. (2.71) and (2.72) in Eq. (2.70), we obtain

$$H(\omega) = \frac{\sum_k b_k (j\omega)^k}{\sum_m a_m (j\omega)^m}. \tag{2.73}$$

This important ratio is called the *frequency transfer function* of the system.† Note that the right-hand side of Eq. (2.73) depends only on the system.

Physically, Eq. (2.73) tells us that a way to test a linear time-invariant system is to apply a sinusoid of known amplitude, frequency, and phase to the input of the system. The output will be another sinusoid *at the same frequency* but the amplitude and phase will, in general, differ from that of the input. Taking the ratio of these two complex coefficients gives the value (in amplitude and phase) of the system transfer function at that frequency. This process may be continued for other frequencies. A plot of the system frequency transfer function may then be obtained by drawing a continuous curve through the points graphed.

We can go one step farther and incorporate the system response characteristics with the components of the system. This is done by multiplying every term giving

† The *transfer function* of a linear time-invariant system is defined as the ratio of the Laplace transform of the output to the Laplace transform of the input for zero initial conditions.

rise to a derivative by $(j\omega)$ and every term giving rise to an integral by $(1/j\omega)$. This is the method commonly used in ac linear circuit analysis. Use of these methods is demonstrated in Example 2.12.1.

EXAMPLE 2.12.1

Determine the frequency transfer function of the system shown in Fig. 2.16.

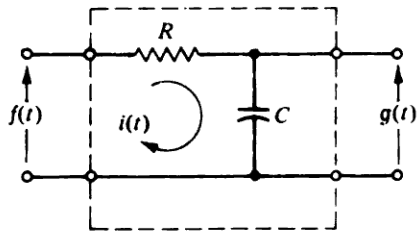


Figure 2.16 The RC lowpass filter.

Solution The differential equation describing this system is found easily by summing the currents yielding†

$$\frac{1}{R}[f(t) - g(t)] = C\frac{dg}{dt},$$

or

$$f(t) = RC\frac{dg}{dt} + g(t).$$

For the calculation of the system frequency transfer function we let $f(t) = e^{j\omega t}$ and $g(t) = H(\omega)e^{j\omega t}$ so that substitution of the particular solution in the differential equation above becomes

$$e^{j\omega t} = j\omega RCH(\omega)e^{j\omega t} + H(\omega)e^{j\omega t}.$$

The system frequency transfer function is then

$$H(\omega) = \frac{1}{j\omega RC + 1}.$$

To make the calculation even easier, we can replace C by $(j\omega C)^{-1}$ in the circuit and use the property of a voltage divider to obtain the above result by inspection. This is the method commonly used in ac linear circuit analysis.

† Usually a zero source impedance and an infinite load impedance are assumed for defining transfer functions.

DRILL PROBLEM 2.12.1

A given linear time-invariant system can be described by the differential equation $g(t) + d^4g/dt^4 = d^2f/dt^2$. Determine the system frequency transfer function.

ANSWER: $H(\omega) = -\omega^2/(1 + \omega^4)$.

In general, $H(\omega)$ is a complex-valued function of frequency and is usually expressed in the polar form

$$H(\omega) = |H(\omega)|e^{j\theta(\omega)} = \left|\frac{G(\omega)}{F(\omega)}\right| e^{j[\theta_g(\omega)-\theta_f(\omega)]}. \tag{2.74}$$

The quantity $|H(\omega)|$ is called the *magnitude response* of the system and $\theta(\omega)$ is called the *phase shift* of the system.

A device available for the rapid measurement of the magnitude response of a system is the sweep generator. Basically, the sweep generator is a voltage-controlled oscillator whose center frequency and frequency sweep can be preset. The amplitude of the resulting sinusoidal waveform is held constant over the range.

The sweep generator is connected to the input of the system under test and a magnitude-responding device is connected to the output and to the vertical deflection on an oscilloscope. A sawtooth voltage derived from the oscilloscope sweep circuit causes the sweep generator to generate sinusoids of constant amplitude but whose frequencies vary in a linear fashion as the oscilloscope trace moves horizontally. The resulting display is a plot of the magnitude of the system frequency transfer function. Because phase shift and the time rate of frequency are related, phase plots are generally not made in this manner.

Note that a transfer function may include a change of units (e.g., from voltage to current, current to voltage, etc.). It may also be used to describe systems containing active elements (gain factors) as long as they are linear and time-invariant.

2.13 STEADY-STATE RESPONSE TO PERIODIC SIGNALS

It follows from Section 2.12 that if the input signal to a linear time-invariant system is described by

$$f(t) = Ae^{j(\omega_1 t + \phi_1)}, \tag{2.75}$$

the output is

$$g(t) = AH(\omega_1)e^{j(\omega_1 t + \phi_1)}. \tag{2.76}$$

Then the system response to a periodic signal becomes very simple if the input is written as an exponential Fourier series,

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t},$$

and the resulting output is

$$g(t) = \sum_{n=-\infty}^{\infty} G_n e^{jn\omega_0 t} = \sum_{n=-\infty}^{\infty} H(n\omega_0) F_n e^{jn\omega_0 t}. \tag{2.77}$$

We can also relate the output average power to the input by using Parseval's theorem:

$$P_f = \sum_{n=-\infty}^{\infty} |F_n|^2,$$

$$P_g = \sum_{n=-\infty}^{\infty} |G_n|^2 = \sum_{n=-\infty}^{\infty} |H(n\omega_0)|^2 |F_n|^2. \tag{2.78}$$

EXAMPLE 2.13.1

Determine the output, $g(t)$, of a linear time-invariant system whose input and frequency transfer function are shown in Fig. 2.17. Calculate the input and output average power.

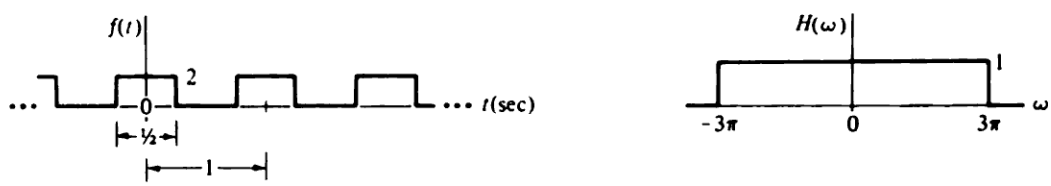


Figure 2.17 The input and transfer function for Example 2.13.1.

Solution Writing the Fourier series for the input,

$$f(t) = \sum_{n=-\infty}^{\infty} \frac{\sin(n\pi/2)}{n\pi/2} e^{jn2\pi t},$$

and therefore the output is [using Eq. (2.77)]

$$g(t) = \sum_{n=-\infty}^{\infty} H(2\pi n) \frac{\sin(n\pi/2)}{n\pi/2} e^{jn2\pi t} = 1 + \frac{4}{\pi} \cos 2\pi t.$$

The input average power is

$$P_f = \frac{1}{T} \int_{t_0}^{t_0+T} |f(t)|^2 \, dt = \int_{-1/4}^{1/4} 4 \, dt = 2 \text{ W}.$$

The output average power is [using Eq. (2.78)]

$$P_g = \sum_{n=-\infty}^{\infty} |H(n\omega_0)|^2 |F_n|^2,$$

$$P_g = \left(\frac{2}{\pi}\right)^2 + 1 + \left(\frac{2}{\pi}\right)^2 = 1.811 \text{ W}.$$

Note that, as far as energy and power calculations are concerned, it is only the magnitude of $H(\omega)$ that counts.

DRILL PROBLEM 2.13.1

The system described in Drill Problem 2.12.1 is used to filter the input signal of Example 2.13.1. Determine a series expression for the output signal; also calculate the output power.

ANSWER: $g(t) \cong \left(\frac{1}{\pi}\right)^3 \left(-\cos 2\pi t + \frac{1}{27} \cos 6\pi t - \cdots\right)$; 0.52 mW.

2.14 HARMONIC GENERATION

An important practical application of the Fourier series representation is in the measurement of the generation of harmonic content. This harmonic generation may be either intentional or unintentional and we briefly consider both types here.

In some applications it is desirable to generate a sinusoid whose frequency is an exact multiple of that of a given sinusoid. A device with a nonlinear output-input gain characteristic can be used to accomplish this. Suppose, for example, that the output-input gain characteristic is

$$e_o(t) = a_1 e_i(t) + a_2 e_i^2(t), \tag{2.79}$$

where $e_o(t)$, $e_i(t)$ are the output and input signals respectively. If we let $e_i(t) = A \cos \omega_0 t$, then we can write

$$\begin{aligned} e_o(t) &= a_1 A \cos \omega_0 t + a_2 A^2 \cos^2 \omega_0 t \\ &= \tfrac{1}{2} a_2 A^2 + a_1 A \cos \omega_0 t + \tfrac{1}{2} a_2 A^2 \cos 2\omega_0 t. \end{aligned} \tag{2.80}$$

The nonlinear output-input characteristic has therefore resulted in the generation of a second-harmonic term; such a device is called a "frequency doubler." Similarly, a third-order nonlinearity results in generation of third-harmonic content, etc. In practice, the nonlinearities may only be known graphically. In this case, one can use a power series approximation or numerical techniques using a digital computer. The latter approach is illustrated in problems at the end of this chapter.

Because harmonic content is generated by nonlinearities, the above procedure may be reversed. If an amplifier has a linear gain (i.e., output-input) characteristic, a sinusoidal input will result in a sinusoidal output *at the same frequency*. Therefore the presence of harmonic content in the output when only a single-frequency sinusoid is applied to the input represents distortion resulting from nonlinearities in the amplifier. A convenient way to measure this distortion without requiring knowledge of gain or impedance values is to take the ratio of the mean-square harmonic distortion terms to the mean-square value of the first harmonic.† This ratio is expressed in percent and is called the *total harmonic distortion* (THD) of the amplifier. Expressed in terms of the trigonometric Fourier series, the total harmonic distortion is then

$$\text{THD} = \frac{\sum_{n=2}^{\infty} (a_n^2 + b_n^2)}{a_1^2 + b_1^2}. \tag{2.81}$$

A computation of the THD of a given waveform is illustrated in the following example.

EXAMPLE 2.14.1

A given amplifier is tested with an input 500 Hz sinusoid of 2 mV peak amplitude. The output is found to be composed of the parabolic sections [t is in msec, $f(t)$ is in V]

$$f(t) = \begin{cases} (1 - 4t^2) & -0.5 < t < 0.5 \\ 4(t-1)^2 - 1 & 0.5 < t < 1.5 \end{cases}$$

etc. (periodically), as shown in Fig. 2.18. (a) Determine the total harmonic distortion (THD). (b) What is the linear gain of the amplifier?

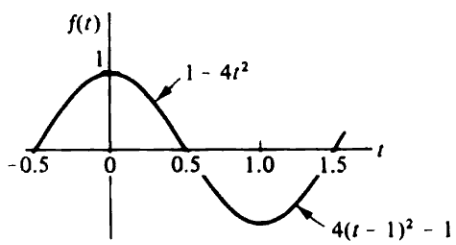


Figure 2.18 A parabolic waveform.

Solution The period is 2 msec; however, because we are only interested in the relative magnitudes of the Fourier series coefficients, we can scale by 10^3 and drop the "msec," thereby scaling everything to "sec." By inspection, the average value of $f(t)$ is zero and the b_n's are zero (as a result of even waveform symmetry).

† Zero average value is assumed; otherwise the average value must be subtracted out before taking the mean-square value.

Computation of the a_n's gives

$$a_n = 32\frac{\sin(n\pi/2)}{(n\pi)^3}.$$

(a) The mean-square value of $f(t)$ is

$$\overline{f^2} = \frac{1}{T}\int_{-T/2}^{T/2} f^2(t)\,dt = 2\int_0^{1/2}(1-4t^2)^2\,dt = \frac{8}{15}.$$

The total harmonic distortion (THD) is

$$\text{THD} = \frac{\overline{f^2} - \frac{1}{2}a_1^2}{\frac{1}{2}a_1^2} = \frac{16}{15}\left(\frac{\pi^3}{32}\right)^2 - 1 = 0.145\%.$$

(b) The linear gain is $32/2\pi^3 \times 10^3 \cong 516$.

DRILL PROBLEM 2.14.1

As an extreme case, suppose the input sinusoid in Example 2.14.1 results in a symmetric square wave with unit peak amplitude. Calculate the linear gain and THD. [*Hint:* Use the results of Example 2.5.1.]

ANSWER: (a) 637; (b) 23.4%.

2.15 THE FOURIER SPECTRUM AND EXAMPLES

The exponential Fourier series is composed of a summation of complex exponentials with the F_n representing the magnitudes and initial phase angles of the harmonically related rotating phasors. The resultant phasor is found by adding the individual phasors vectorially. For all but the simple cases of one or two terms, however, the addition of a series of phasors for each instant of time turns out to be an inconvenient way to describe a signal.

The thought now occurs to us: instead of summing and plotting phasors to describe a signal, why not plot the Fourier coefficients (the F_n) as a function of the angular rate (frequency)? A primary motivation behind this approach is that the resulting diagram will hold for all values of time and we can avoid the summing and plotting for each instant of time. Such a graph of the (complex) Fourier coefficients as a function of frequency for the signal $f(t)$ is called the *Fourier spectrum* of $f(t)$. It is quite common in engineering literature to simply call this the "spectrum" of $f(t)$ with the understanding that it refers to the exponential Fourier representation.

For a periodic signal, we have found that the Fourier coefficients correspond to the relative amplitude and phase weightings of a set of harmonically related phasors.

Therefore the Fourier spectrum exists only at $\omega = 0, \pm\omega_0, \pm 2\omega_0, \pm 3\omega_0, \ldots$; i.e., only at discrete values of ω. It is therefore a *discrete spectrum*, sometimes referred to as a *line spectrum*. We may represent this spectrum graphically by drawing vertical lines at the discrete points $\omega = 0, \pm\omega_0, \pm 2\omega_0, \ldots$, with their heights proportional to the amplitudes of the corresponding frequency components (assuming that the coefficients are real-valued). This *amplitude spectrum* then appears on a graph as a series of equally spaced vertical lines with heights proportional to the amplitudes of the respective frequency components.

In general, however, the F_n are complex-valued. To describe the coefficients then requires two graphs, the *magnitude spectrum* and *phase spectrum*. The following example illustrates these different ways of portraying spectra.

EXAMPLE 2.15.1

Sketch the amplitude spectrum and the magnitude and phase spectrum of the function used in Example 2.11.1: $f(t) = 2 \sin 100t$.

Solution The sketches are shown in Fig. 2.19. Note that the magnitude spectrum is symmetrical about $\omega = 0$. This is a general characteristic of magnitude spectra for real-valued $f(t)$, as can be seen from Eq. (2.46).

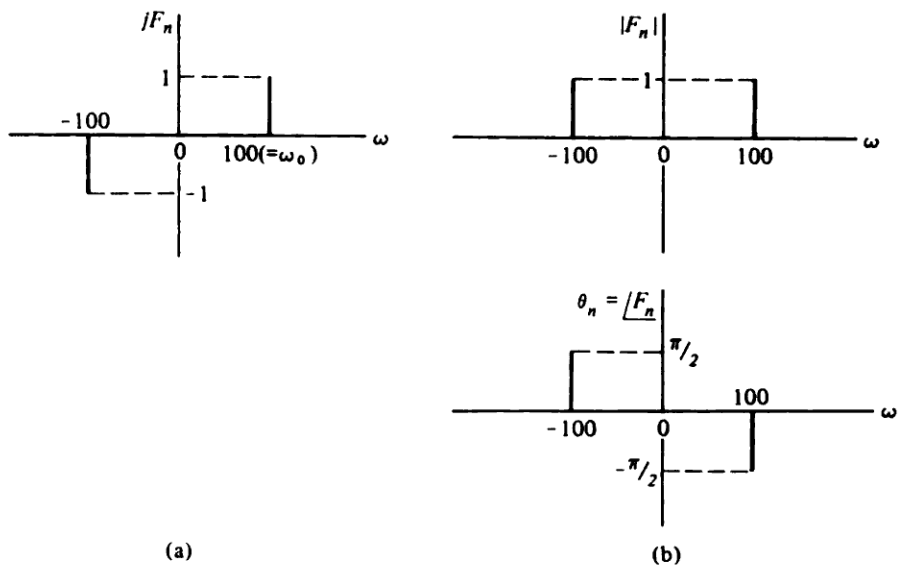


Figure 2.19 Line spectra corresponding to $f(t)$ in Example 2.15.1: (a) amplitude spectrum; (b) magnitude and phase spectra.

We could have used the trigonometric series to describe the spectrum. In this case, we would have one-sided amplitude and phase spectra (i.e., $n \geq 0$ only). The ampli-

tude and phase of each sinusoidal component can be found using Eqs. (2.58) and (2.59). An interesting interpretation of the trigonometric spectrum is illustrated in Fig. 2.20.

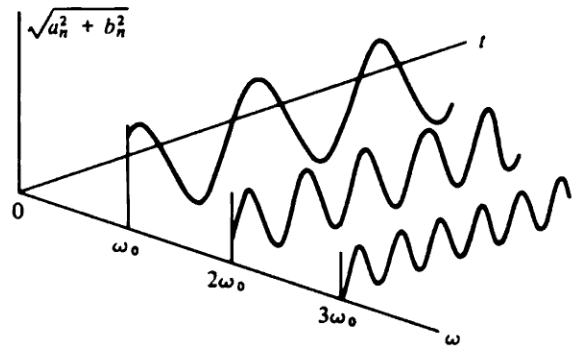


Figure 2.20 An interpretation of the trigonometric line spectrum.

Now let us find the Fourier spectrum for the periodic gate function shown in Fig. 2.21. The periodic gate function is a generalized case of the square-wave signal in which the width and height of rectangular pulses are variable and one level is at zero amplitude. The gate function and its corresponding spectrum play a major role in the analysis of many systems.

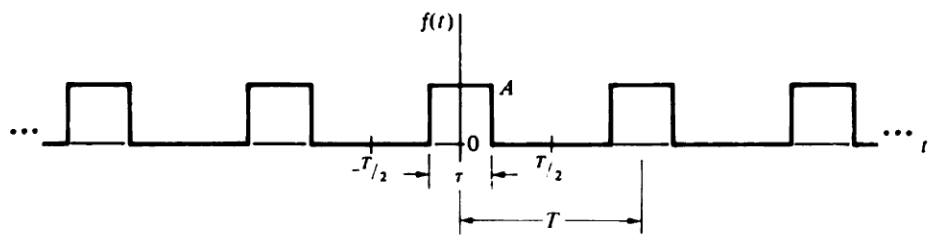


Figure 2.21 The periodic gate function.

Proceeding, the Fourier coefficients are

$$F_n = \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-jn\omega_0 t}\,dt$$

$$= \frac{1}{T}\int_{-\tau/2}^{\tau/2} Ae^{-jn\omega_0 t}\,dt$$

$$= \frac{-A}{jn\omega_0 T}\left(e^{-jn\omega_0\tau/2} - e^{jn\omega_0\tau/2}\right), \qquad n \neq 0,$$

$$= \frac{2A}{n\omega_0 T}\sin(n\omega_0\tau/2), \qquad n \neq 0,$$

$$F_n = \frac{A\tau}{T}\frac{\sin(n\omega_0\tau/2)}{(n\omega_0\tau/2)}, \qquad n \neq 0.$$

Defining a new normalized variable $x = n\omega_0\tau/2$, we have

$$F_n = \frac{A\tau}{T}\left[\frac{\sin x}{x}\right] \qquad x \neq 0.$$

The function within the brackets occurs so often that we shall assign it the convenient abbreviation "sine over argument"†

$$\mathrm{Sa}(x) \triangleq \frac{\sin x}{x}. \tag{2.82}$$

A graph of Eq. (2.82) appears as Fig. 2.22.

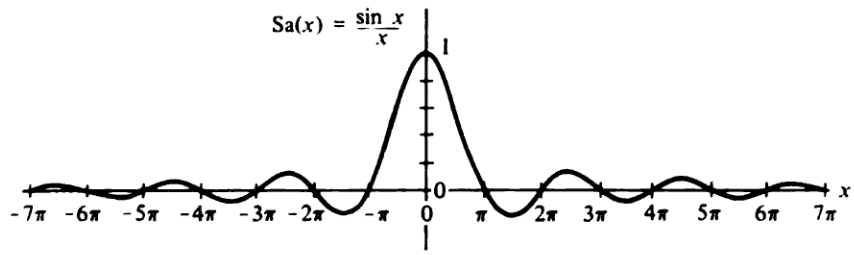


Figure 2.22 The function Sa(x).

From Fig. 2.22, observe that the amplitude of the function $\mathrm{Sa}(x)$ oscillates, decaying in either direction of x and approaching zero as $|x| \to \infty$. The maximum value of this function occurs as x approaches zero, for $[(\sin x)/x] \to 1$ as $x \to 0$. The latter can be seen by expanding $\sin x$ in a power series before taking the limit or by using l'Hospital's rule.

Because‡

$$F_0 = \frac{1}{T}\int_{-\tau/2}^{\tau/2} A\,dt = \frac{A\tau}{T},$$

the exponential Fourier representation of the periodic gate function is given by

$$f(t) = \frac{A\tau}{T}\sum_{n=-\infty}^{\infty} \mathrm{Sa}(n\pi\tau/T)e^{jn\omega_0 t}, \tag{2.83}$$

where

$$\omega_0 = 2\pi/T.$$

† This abbreviation is not used by all authors. Another common abbreviation is

$$\mathrm{sinc}\,(x) \triangleq \frac{\sin \pi x}{\pi x}.$$

‡ The coefficient F_0 represents the average value of the signal $f(t)$. Note that this term can be changed without affecting any of the other harmonic components in the exponential Fourier series.

Because the F_n are real, we need only display the amplitude spectrum. Aside from the scaling constant A, this spectrum is dependent only on the parameter choices τ and T.

It is instructive first to keep τ fixed and vary the period T. As T increases, two effects are noticed: (1) the amplitude of the spectrum decreases as $1/T$, and (2) the spacing between lines decreases as $2\pi/T$. These effects are shown in the drawings of Fig. 2.23. Note that the envelope of the spectra remains fixed except for the amplitude scale factor. As the period T increases, the spacing between components, $\Delta\omega = 2\pi/T$, becomes smaller and there are more and more frequency components in a given range of frequency. The amplitudes of these components decrease as T is increased.† Although the amplitude of the spectrum goes to zero, the *shape* of the spectrum does not change. We conclude that *the shape, or envelope, of the spectrum is dependent only upon the pulse shape and not on the repetition period T.*

It is also instructive to keep T fixed and vary τ (with the restriction that $\tau < T$). As τ increases, (1) the amplitude of the spectrum increases proportional to τ, and (2) the frequency content of the signal is compressed within an increasingly narrower range of frequencies. This trend is shown in Fig. 2.24. Thus there is *an inverse relationship between pulse width in time and the frequency "spread" of the spectrum.* A convenient measure of the frequency spread is the distance to the first zero crossing of the Sa(x) function.

Summarizing, the spectrum of a periodic function is a line spectrum with spectral components located at multiples of the basic repetition frequency. The spacing between frequency components is inversely proportional to the repetition period T. The frequency spread of the envelope of the spectrum is inversely proportional to the

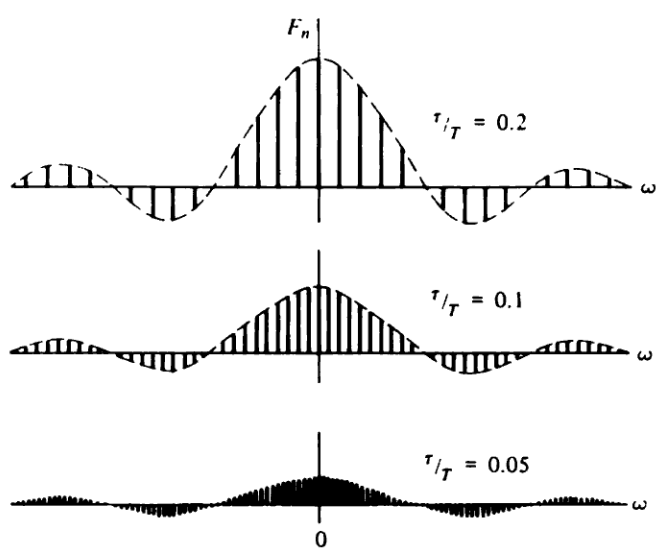


Figure 2.23 Amplitude spectra for various values of τ/T, τ fixed.

† In the limit as $T \to \infty$, only a single pulse remains because its nearest neighbors have been removed to infinity. In this limiting case, the frequency spacings approach zero and the spectrum becomes continuous (see Chapter 3).

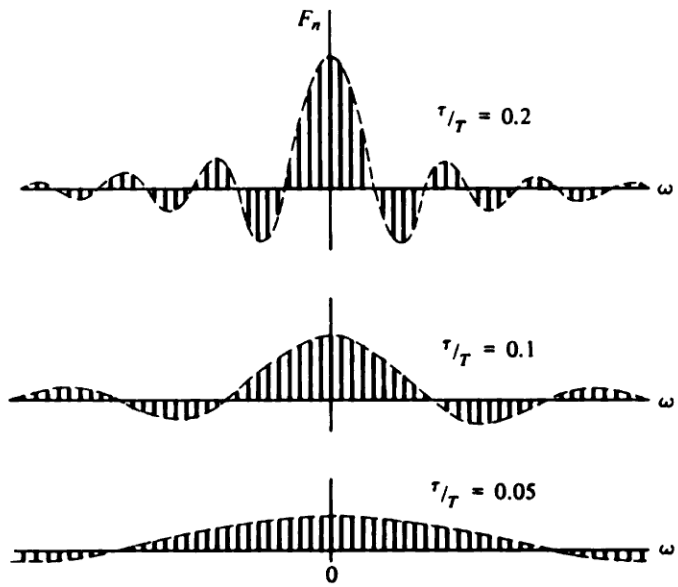


Figure 2.24 Amplitude spectra for various values of τ/T, T fixed.

pulse width τ. The shape of the spectrum depends only on the pulse shape and not on the repetition period T. Finally, even though these effects are illustrated here using the periodic gate function, the general conclusions hold for all periodic functions.

A device available for laboratory measurements of spectra of power signals is the scanning spectrum analyzer. Basically, the scanning spectrum analyzer is designed to accept only a very narrow spread (band) of frequencies, measure the power, and display the square root of the power as the vertical deflection on an oscilloscope. The placement of the frequencies accepted is movable electrically.

A simplified block diagram of a scanning spectrum analyzer is shown in Fig. 2.25. A sawtooth voltage derived from the oscilloscope sweep circuit causes the analyzer to sweep across the frequencies of interest in a linear fashion as the oscilloscope trace moves horizontally. Any nonzero spectral component present results in a vertical displacement of the oscilloscope trace as its frequency is passed. This can give a close approximation to the magnitude spectrum of the input waveform. The accuracy of the

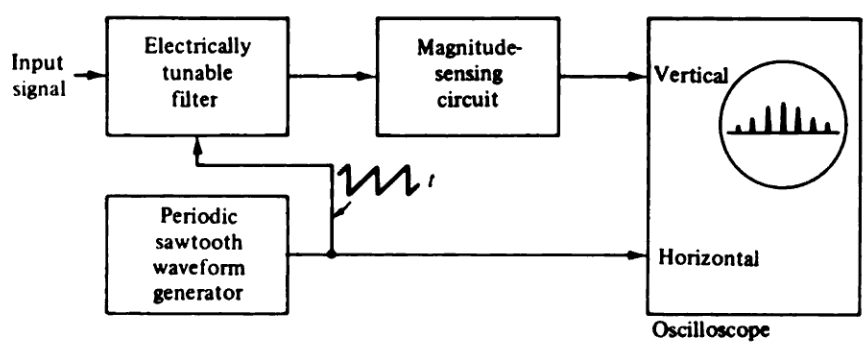


Figure 2.25 Block diagram of a scanning spectrum analyzer.

approximation is governed by how narrow a band of frequencies the analyzer uses to make the measurement and by how slowly it sweeps over the desired range of frequencies.

2.16 ★ NUMERICAL COMPUTATION OF FOURIER COEFFICIENTS

Calculation of the Fourier series coefficients may be approximated numerically. This becomes advantageous when an analytical expression for $f(t)$ is not known and $f(t)$ is available as numerical data points, or when the integration is difficult to perform or must be repeated often with parameter changes. We shall outline the approach here using only the most straightforward methods of numerical approximation.

Equations for the calculation of trigonometric Fourier series coefficients are given in Eqs. (2.54)–(2.56). Choosing the cosine term as an example, we have

$$a_n = \frac{2}{T}\int_0^T f(t)\cos n\omega_0 t\,dt\,, \tag{2.84}$$

where $\omega_0 = 2\pi/T$. We now assume that M equally spaced data points representing $f(t)$ over $(0, T)$ are either known or available to us, and that the interval between each such data point is Δt.[†] Approximating the integration in Eq. (2.84) with a summation of rectangular strips each Δt in width, we have[‡]

$$a_n \cong \frac{2}{T}\sum_{m=1}^{M} f(m\cdot\Delta t)\cos\left[n(2\pi/T)(m\cdot\Delta t)\right]\Delta t\,, \tag{2.85}$$

where

$$\Delta t = T/M\,. \tag{2.86}$$

Using Eq. (2.86) in Eq. (2.85), we obtain our desired result:

$$a_n \cong \frac{2}{M}\sum_{m=1}^{M} f(m\cdot\Delta t)\cos(2\pi mn/M)\,. \tag{2.87}$$

Thus once the M equally spaced data points have been determined, the computation is expressed in terms of the integers m (the integer time scale) and n (the integer frequency scale).

[†] If the data points are not equally spaced, a subroutine can be employed to make a "best fit" numerically to the available data and then interpolate to obtain the equally spaced data samples.

[‡] Equivalently, the summation may go from 0 to $(M - 1)$.

Computation of the other Fourier coefficients follows in the same manner so that

$$b_n \cong \frac{2}{M} \sum_{m=1}^{M} f(m \cdot \Delta t) \sin 2\pi mn/M), \tag{2.88}$$

$$a_0 \cong \frac{1}{M} \sum_{m=1}^{M} f(m \cdot \Delta t). \tag{2.89}$$

An example of such a calculation for a known function demonstrates the procedure.

EXAMPLE 2.16.1

Using 100 equally spaced sample points per period, compute the coefficients of the first 10 harmonic terms of the trigonometric Fourier series for the triangular waveform of Drill Problem 2.9.1.

Solution Over the interval $(0, 2\pi)$,

$$f(t) = \begin{cases} \pi/2 - t & 0 \le t < \pi \\ t - 3\pi/2 & \pi \le t < 2\pi \end{cases}$$

and thus we have

$$T = 2\pi,$$
$$M = 100,$$
$$\Delta t = T/M = \pi/50.$$

Using Eqs. (2.87)–(2.89), we find

$$a_n \cong \frac{2}{100} \sum_{m=1}^{100} f(m\pi/50) \cos(\pi mn/50), \text{ etc.}$$

A sample program and corresponding numerical solution for a typical programmable calculator is shown in Table 2.3.† [*Note:* The choice of variables in the program displayed is $a_n \to A$; $b_n \to B$; $n \to C$; $m \to X$.] One look at the solution printout reveals that most of the computation time is being spent to compute coefficients which are zero! Actually, we should have known this as a result of the symmetry in the waveform.

† This type of program is chosen here because it is easy to read and follow.

TABLE 2.3 Program and Solution Printout for Example 2.16.1

Program	Solution printout
1: 0→C	0
2: 0→A;0→B;1→X	-.000000
3: IF X>50;GTO 5	1
4: .02(π/2-πX/50)COS(πCX/50)→Z; GTO 6	1.273659
5: .02(πX/50-3π/2)COS(πCX/50)→Z	.000000
6: Z+A→A	2
7: IF X>50;GTO 9	.000000
8: .02(π/2-πX/50)SIN(πCX/50)→Z; GTO 10	.000000
9: .02(πX/50-3π/2)SIN(πCX/50)→Z	3
10: Z+B→B	.141891
11: IF X=100;GTO 13	.000000
12: X+1→X;GTO 3	4
13: IF C>0;GTO 15	.000000
14: FXD 0; PRT C; FXD 6; PRT A/2;GTO 16	.000000
15: FXD 0; PRT C; FXD 6; PRT A,B	5
16: IF C=10; GTO 18	.051351
17: C+1→C;GTO 2	.000000
18: END	6
	.000000
	.000000
	7
	.026407
	.000000
	8
	.000000
	.000000
	9
	.016145
	.000000
	10
	.000000
	.000000

It is interesting to compare the above results determined by numerical integration to the exact results determined in Drill Problem 2.9.1. A comparison is shown below.

n	*Approximate* a_n	*"Exact"* a_n	*Difference*
1	1.273659	1.273240	+0.000419
3	0.141891	0.141471	+0.000420
5	0.051351	0.050930	+0.000421
7	0.026407	0.025984	+0.000423
9	0.016145	0.015719	+0.000426

Note that agreement between the two is good, but not perfect. One who is acquainted with numerical integration procedures may be tempted to try to improve the numeri-

cal accuracy by more sophisticated integration procedures. However, this is not the main source of error here; we shall devote some attention to this shortly.

DRILL PROBLEM 2.16.1

(a) Write a computer program for Example 2.16.1.

(b) Run your program to investigate the effects on the accuracy of the coefficients for the following number of equally spaced sample points per period: 20; 40; 60; 80.

ANSWER FOR a_1: +0.010523; +0.002621; +0.001164; +0.000654.

When M equally spaced data points are used to represent a given waveform for computation of Fourier coefficients, that waveform is not only made periodic with period M but the data points themselves also exhibit a periodicity (with period $1/M$). The net result is that orthogonality between terms with differing harmonic numbers is guaranteed only if that difference is less than $M/2$. Taking the harmonic numbers (n) as the consecutive integers starting with zero, we have

$$n \leq M/2 . \tag{2.90}$$

In other words, *the highest harmonic coefficient which can be determined uniquely from M sampled data points is that one for which the harmonic number is less than or equal to one-half the number of sampled data points.*

An illustration of a typical line spectrum of a periodic function and a line spectrum of its numerical approximation is shown in Fig. 2.26. Note that there is some approximation error even for those coefficients for which $n < M/2$ but that this approximation error becomes proportionally very large for $n > M/2$. The error arising in those terms for which $n < M/2$ is called "aliasing error" and is usually the main source of error in the numerical approximation. Aliasing error is the subject of the next section. Using more sampled data points for a fixed T (i.e., increasing M and thus decreasing Δt) aids in decreasing the amount of aliasing error present.

2.17 ★ EFFECTS OF ALIAS TERMS

If all frequency components in a given waveform are within the harmonic range given by Eq. (2.90), the Fourier coefficients can be determined within the accuracy of the numerical integration procedures. However, the magnitudes of the frequency components in a given waveform do not always lie within a finite range but rather may decrease gradually as n becomes large. Because orthogonality is no longer guaranteed, the numerical evaluation of a coefficient at a given harmonic within the range

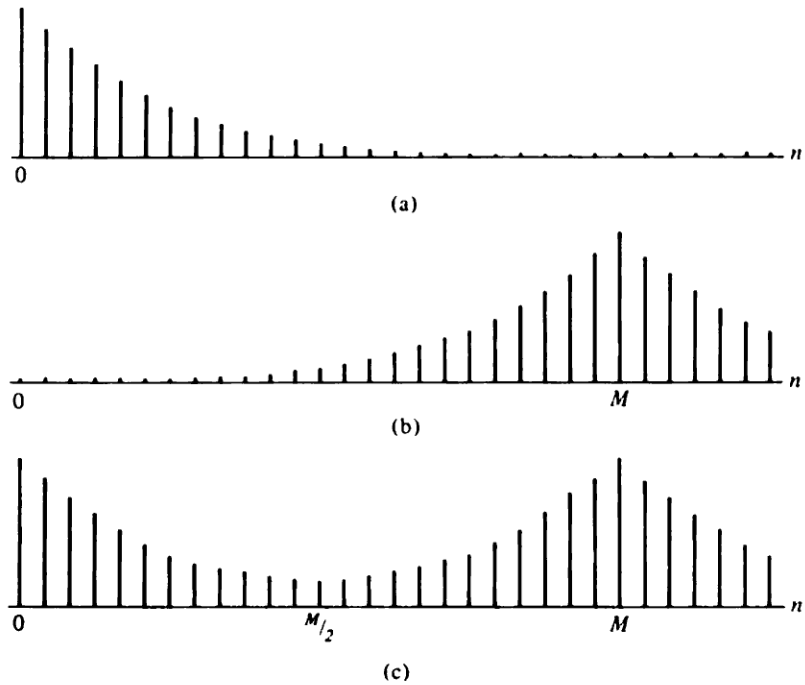


Figure 2.26 (a) The line spectrum of a periodic function; (b) spectral replica produced by sampling; (c) composite numerical approximation to the line spectrum.

given by Eq. (2.90) may actually include waveform components that otherwise would be assigned to higher coefficients that lie outside this range (see Fig. 2.26). This type of error is called *aliasing error*.

It turns out that only certain harmonic terms contribute (erroneously) to the result.† The net effect on the computation can be summarized as follows. Designating the true values of the Fourier coefficients by a'_n, b'_n, the coefficient values actually computed from the M equally spaced data points are given by

$$a_n = a'_n + \sum_{k=1}^{\infty} (a'_{kM+n} + a'_{kM-n}), \tag{2.91}$$

$$b_n = b'_n + \sum_{k=1}^{\infty} (b'_{kM+n} - b'_{kM-n}). \tag{2.92}$$

Thus if we restrict the range of computation to $n \leq M/2$, Eqs. (2.91) and (2.92) show that the numerical accuracy of the computed Fourier coefficients a_n, b_n is dependent on how fast the true coefficients decrease for $n \leq M/2$. The terms within the summations in Eqs. (2.91) and (2.92) are called alias terms and arise as a result of the sampled nature of the data. The presence of the alias terms is usually the main source of inaccuracy in the numerical computation of the Fourier coefficients. Their effect can be minimized for a given waveform with T fixed by decreasing Δt and thereby increasing the number of sample points, M.

† The proof of these statements is postponed until Chapter 3.

To emphasize the effect of the alias terms on numerical accuracy, let us attempt to estimate the accuracy of the first Fourier coefficient a_1 in Example 2.16.1. We define the error in the numerical approximation by

$$\epsilon_n = a_n - a'_n .$$

Using Eq. (2.91) and $M = 100$, the error in a_1 is

$$\epsilon_1 = \sum_{k=1}^{\infty} (a'_{100k+n} + a'_{100k-n}) .$$

From the results of Drill Problem 2.9.1, we know that the true Fourier coefficients for this waveform are

$$a'_n = \begin{cases} 4/(\pi n^2) & n \text{ odd} \\ 0 & n \text{ even} \end{cases} .$$

Therefore the aliasing error in the computation of the first Fourier coefficient is

$$\epsilon_1 = \frac{4}{\pi}\left[\frac{1}{(101)^2} + \frac{1}{(201)^2} + \frac{1}{(301)^2} + \frac{1}{(401)^2} + \cdots \right.$$
$$\left. + \frac{1}{(99)^2} + \frac{1}{(199)^2} + \frac{1}{(299)^2} + \frac{1}{(399)^2} + \cdots\right],$$

$$\epsilon_1 \cong \frac{8}{\pi}\frac{1}{(100)^2}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots\right).$$

However, we can also write†

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

so that

$$\epsilon_1 \cong \frac{8}{\pi}\frac{1}{(100)^2}\frac{\pi^2}{6} = 0.000419 .$$

A comparison with the results of Example 2.16.1 reveals that all of the approximation error (to six places) is due to aliasing, despite the fact that we used 100 data samples to compute the first harmonic coefficient! Thus the effects of aliasing are not always negligible. Also, as the harmonic number n is increased, some terms in the summation in Eq. (2.91) or Eq. (2.92) become more predominant so that the aliasing error generally increases with increasing n. This trend is evident in Example 2.16.1.

For a given number of data points, the effect of aliasing depends on the rate of convergence of the Fourier coefficients as n becomes large. In many cases the true coefficient behavior may not be known. However, the rate of convergence of the co-

† At points of finite discontinuity, the Fourier series converges to the arithmetic average of the values of the function to either side of the discontinuity. Thus this result can be obtained by evaluating the result of Example 2.9.1 at $t = 2$.

efficients can often be estimated either from some knowledge of the physical nature of the source of the data or from the characteristics of the waveform itself. We shall take up the latter topic again in a later section of this chapter.

What happens if the computation of a coefficient is attempted for $n > M/2$? In this case one of the alias terms in Eq. (2.91) or (2.92) will predominate. For instance, if the computation of a_{99} in Example 2.16.1 were attempted, Eq. (2.91) gives

$$a_{99} = a'_{99} + (a'_{199} + a'_{299} + \cdots) + (a'_1 + a'_{101} + \cdots).$$

The term a'_1 will, of course, predominate if the coefficients in the series are convergent. In fact, if one writes out these results for a few different values of n, it becomes evident that all the computed a_n are symmetric about $n = M/2$ and the b_n are antisymmetric about $n = M/2$. This is shown in the following example.

EXAMPLE 2.17.1

Numerically evaluate the coefficients of the first 9 harmonic terms of the trigonometric Fourier series for the waveform defined by

$$f(t) = \begin{cases} t & 0 < t < \pi \\ 0 & \pi < t < 2\pi \end{cases}.$$

Use 10 sample points per period. At points of finite discontinuity in the waveform, use the average of the values to either side of the discontinuity.

Solution $T = 2\pi$; $M = 10$; $\Delta t = 2\pi/10 = \pi/5$;

$$a_n \cong \frac{2}{10} \sum_{m=1}^{10} f(m\pi/5) \cos(2\pi mn/10), \text{ etc.}$$

A sample program and solution printout is shown in Table 2.4. Note the symmetry properties of the coefficients about the harmonic number $n = M/2 = 5$.

DRILL PROBLEM 2.17.1

Write a program for computation of the trigonometric Fourier coefficients for the waveform of Drill Problem 2.9.2. Use 10 sample points over the interval (0, 2); assign the average value of the waveform to the sample point at $t = 2$. (a) Determine the numerical error in b_1. (b) What is the relationship b_9 to b_1?

ANSWER: (a) +0.021083; (b) equal but reversed sign.

TABLE 2.4 Program and Solution Printout for Example 2.17.1

Program	Solution printout
1: 0→C	0
2: 0→A;0→B;1→X;1→Y	.785398
3: IF X=5;.5→Y	1
4: .2(πX/5)Y COS(πCX/5)→Z	-.657984
5: Z+A→A	.966883
6: .2(πX/5)Y SIN(πCX/5)→Z	2
7: Z+B→B	.000000
8: IF X=5;GTO 10	-.432403
9: X+1→X; GTO 3	3
10: FXD 0;PRT C	-.095999
11: IF C=0;FXD 6;PRT A/2; GTO 13	.228250
12: FXD 6;PRT A,B	4
13: IF C=10;GTO 15	.000000
14: C+1→C;GTO 2	-.102077
15: END	5
	-.062832
	.000000
	6
	.000000
	.102077
	7
	-.095999
	-.228250
	8
	.000000
	.432403
	9
	-.657984
	-.966883

Computation of coefficients in the exponential Fourier series follows in a similar manner through use of Euler's identities. Thus we can rewrite Eq. (2.36) as

$$F_n = \frac{1}{T}\int_0^T f(t)\cos n\omega_0 t\,dt - j\frac{1}{T}\int_0^T f(t)\sin n\omega_0 t\,dt \tag{2.93}$$

$$= \frac{a_n}{2} - j\frac{b_n}{2}, \tag{2.94}$$

and the numerical computation of the a_n, b_n has already been discussed.

While computation of coefficients for the exponential Fourier series is essentially the same as for the trigonometric series, interpretation of the results is, in fact, slightly different. Recall that there are both positive and negative frequencies in the exponential series, but only positive frequencies in the trigonometric series. For real-valued $f(t)$, the relation between the positive and negative frequency components in the exponential series is that of a complex conjugate; i.e., $F_{-n} = F_n^*$ [cf. Eq. (2.46)]. Suppose we numerically compute the coefficients for the exponential series using

Eqs. (2.87), (2.88), and (2.94). The computed coefficients for $0 \le n \le M/2$ yield the positive frequency components, just as they did for the trigonometric series. But those coefficients computed for $M/2 \le n \le M$ are the complex conjugates of those for $0 \le n \le M/2$ [cf. Eqs. (2.91), (2.92), and (2.94)]. Thus we conclude that *the positive frequency components in the numerical computation of the exponential Fourier series coefficients are given over the harmonic range* $0 \le n \le M/2$ *and the negative frequency components are over* $M/2 \le n \le M$. The numerical accuracy of the exponential series is affected by aliasing in the same way that it is for the trigonometric series.

This introduction will enable the reader to proceed to use the digital computer to numerically evaluate the Fourier series. A more theoretical basis for understanding the aliasing effect must be postponed until Chapter 3.

EXAMPLE 2.17.2

Compute and plot the magnitude spectrum of the exponential Fourier series for the periodic waveform shown in Fig. 2.27(a). Use 50 sample points per period and plot the spectrum over the range $0 \le n < 50$. At points of a finite discontinuity in the waveform, assume continuity to the left.

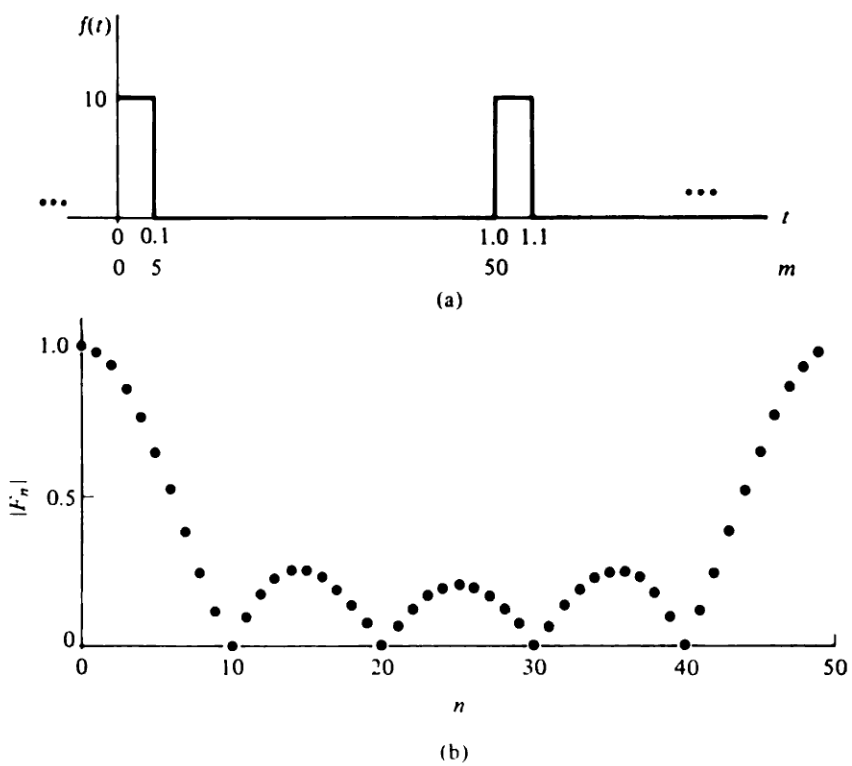


Figure 2.27 (a) A periodic rectangular function; (b) its computed magnitude spectrum using the discrete exponential Fourier series (50 sample points).

Solution See Fig. 2.27(b).

DRILL PROBLEM 2.17.2

Determine the common characteristic(s) which are true for the following coefficients in the exponential Fourier series ($k = 0, 1, 2, \ldots$): (a) $n = kM$; (b) $n = kM/2$.

ANSWER: (a) All represent the average value of the waveform; (b) All must be real-valued.

2.18 SINGULARITY FUNCTIONS

There is a particular class of functions which plays a role of such importance in signal analysis that we treat it separately here. Members of this class have simple mathematical forms but they are either not finite everywhere or they do not have finite derivatives of all orders everywhere. For this reason they are called *singularity functions*.

Singularity functions are mathematical idealizations and, strictly speaking, do not occur in physical systems. They are useful in signal analysis because they serve as good approximations to certain limiting conditions in physical systems. This, together with their simple mathematical forms and properties, makes it possible often to evaluate complicated expressions which might otherwise be impossible—or at least very difficult—to solve.

A singularity function of great importance to us is the unit impulse function (also called the Dirac delta or impulse function), $\delta(t)$. This function has the property exhibited by the following integral:

$$\int_a^b f(t)\,\delta(t - t_0)\,dt = \begin{cases} f(t_0) & a < t_0 < b \\ 0 & \text{elsewhere} \end{cases} \tag{2.95}$$

for any $f(t)$ continuous at $t = t_0$, t_0 finite. The impulse function selects or sifts out a particular value of the function $f(t)$, namely, the value at $t = t_0$, in the integration process. Equation (2.95) is often referred to as the *sifting property* of the impulse function. An example illustrating the use of the sifting property follows.

EXAMPLE 2.18.1

Evaluate the definite integral $\int_{-\infty}^{\infty} e^{\cos t}\,\delta(t - \pi)\,dt$.

Solution $\int_{-\infty}^{\infty} e^{\cos t}\,\delta(t - \pi)\,dt = e^{\cos \pi} = e^{-1} = 0.368.$

DRILL PROBLEM 2.18.1

Evaluate:

$$\text{(a)} \int_1^{\infty} e^{-x^2}\delta(x)\,dx, \qquad \text{(b)} \int_1^{100} \log t\,\delta(t-10)\,dt.$$

ANSWER: (a) 0; (b) 1.

As seen from Eq. (2.95), the impulse function is no ordinary function. However, we can treat $\delta(t)$ as a function formally obeying the rules of integration provided that all conclusions are based on Eq. (2.95) and not on any point properties of $\delta(t)$. Using Eq. (2.95) we can demonstrate formally the following properties of the impulse function.

Area (strength)

If $f(t) = 1$, Eq. (2.95) becomes

$$\int_a^b \delta(t - t_0)\,dt = 1, \qquad a < t_0 < b. \tag{2.96}$$

Therefore $\delta(t)$ has unit area.[†] In an entirely analogous manner, $A\delta(t)$ has an area of A units.

Amplitude

We have

$$\delta(t - t_0) = 0 \qquad \text{for all} \qquad t \neq t_0. \tag{2.97}$$

That this must be true follows directly from a consideration of Eq. (2.95); i.e., all values of $f(t)$ for $t \neq t_0$ are neglected (given zero weight) in the integration process. *The amplitude at the point $t = t_0$ is undefined.*

Graphic Representation

It is fairly obvious from the above properties that we are going to encounter some difficulty in graphing the impulse function. To avoid any attempt to display the amplitude at $t = t_0$, we shall draw an arrow at the point $t = t_0$ as an indicator of the impulse function. The *area* or *strength* of the impulse is designated by a quantity in parentheses beside the arrow or by the height of the arrow (or both), as shown in Fig. 2.28(a). An arrow pointing down indicates negative area.

[†] Area in this sense denotes "strength" or "weight" of the impulse function.

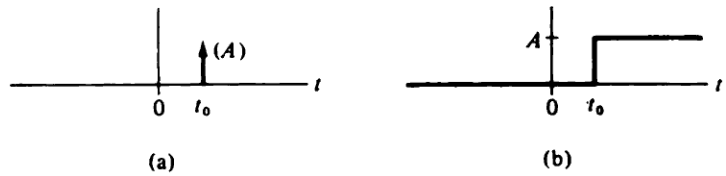


Figure 2.28 Graphic representations for (a) $A\,\delta(t - t_0)$; (b) $Au(t - t_0)$.

Symmetry

We define $\delta(t)$ to be an even function; that is,

$$\delta(t) = \delta(-t). \tag{2.98}$$

This follows from the fact that it makes no difference in Eq. (2.95) whether the point $t = t_0$ is approached from the left or from the right in the integration.

Time Scaling

A scaling in the argument of the impulse function can be converted as follows:

$$\delta(at) = \frac{1}{|a|}\delta(t). \tag{2.99}$$

This can be shown using the change of variable $x = at$; for $a > 0$,

$$\int_{-\infty}^{\infty} f(t)\,\delta[a(t - t_0)]\,dt = \int_{-\infty}^{\infty} f\left(\frac{x}{a}\right)\delta(x - at_0)\frac{dx}{a} = \frac{1}{a}f(t_0).$$

Repeating for $a < 0$ and letting the change of variable be $x = at$, we have

$$\int_{-\infty}^{\infty} f(t)\,\delta[a(t - t_0)]\,dt = \int_{\infty}^{-\infty} f\left(\frac{x}{a}\right)\delta(x - at_0)\frac{dx}{a} = -\frac{1}{a}f(t_0).$$

Combining these two, we can write

$$\int_{-\infty}^{\infty} f(t)\,\delta[a(t - t_0)]\,dt = \frac{1}{|a|}f(t_0).$$

Graphically, the introduction of this scale factor is necessary to retain unity area in the definition of the impulse function.

EXAMPLE 2.18.2

Evaluate the definite integral

$$\int_{-\infty}^{\infty} t^2 e^{-\sin t}\cos 2t\;\delta(2t - 2\pi)\,dt.$$

Solution Using the time scaling property and then the sifting property in succession, we have

$$\int_{-\infty}^{\infty} t^2 e^{-\sin t} \cos 2t\, \delta(2t - 2\pi)\, dt = \frac{1}{2} \int_{-\infty}^{\infty} t^2 e^{-\sin t} \cos 2t\, \delta(t - \pi)\, dt = \frac{1}{2}\pi^2.$$

DRILL PROBLEM 2.18.2

Evaluate the definite integral

$$\int_{-\infty}^{\infty} \delta(1 - \pi t) \cos (1/t)\, dt.$$

ANSWER: $-1/\pi$.

Multiplication by a Time Function

Sometimes we wish to use $\delta(t)$ without the defining integration. Such operations are formal only and result in the following relation:

$$f(t)\,\delta(t - t_0) = f(t_0)\,\delta(t - t_0), \qquad f(t) \text{ continuous at } t_0. \tag{2.100}$$

This can be shown easily by integrating both sides.

Relation to the Unit Step Function

The unit step function is that function defined by

$$u(t - t_0) \triangleq \begin{cases} 1 & t > t_0 \\ 0 & t < t_0 \end{cases}. \tag{2.101}$$

A graphic representation of $Au(t)$ is shown in Fig. 2.28(b). Using Eq. (2.95) and letting $a = -\infty$, $b = t$, and $f(t) = 1$,

$$\int_a^b f(\tau)\,\delta(\tau - t_0)\, d\tau = \int_{-\infty}^{t} \delta(\tau - t_0)\, d\tau = \begin{Bmatrix} 1 & t > t_0 \\ 0 & t < t_0 \end{Bmatrix} = u(t - t_0),$$

or

$$\int_{-\infty}^{t} \delta(\tau - t_0)\, d\tau = u(t - t_0). \tag{2.102}$$

Therefore the integral of the unit impulse function is the unit step function. The converse can also be shown by formally differentiating both sides of Eq. (2.102):

$$\frac{d}{dt} \int_{-\infty}^{t} \delta(\tau - t_0)\, d\tau = \frac{d}{dt} u(t - t_0),$$

$$\delta(t - t_0) = \frac{d}{dt} u(t - t_0).$$

The derivative of the unit step function at $t = t_0$ is the unit impulse occurring at $t = t_0$.

EXAMPLE 2.18.3

Compute and graph the derivative of the rectangular pulse shown in Fig. 2.29(a).

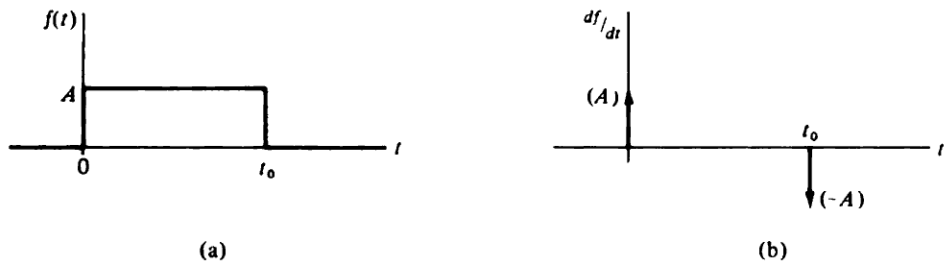


Figure 2.29 A rectangular pulse and its derivative.

Solution Writing an equation for the pulse in terms of step functions, we have

$$f(t) = Au(t) - Au(t - t_0),$$

$$\frac{df}{dt} = A\,\delta(t) - A\,\delta(t - t_0).$$

We can check our result by integrating

$$\int_{-\infty}^{t} \frac{df}{d\tau}\,d\tau = \int_{-\infty}^{t} [A\,\delta(\tau) - A\,\delta(\tau - t_0)]\,d\tau = Au(t) - Au(t - t_0),$$

confirming our result. The corresponding graphical solution is shown in Fig. 2.29(b).

DRILL PROBLEM 2.18.3

It is given that $f(t) = u(t) - u(t - 3) - K\,\delta(t - 4)$. Determine the required numerical value of K such that $\int_{-\infty}^{\infty} f(t)\,dt = 0$.

ANSWER: 3.

The operations indicated involving the impulse function all arose formally from its integral definition. We have observed that the impulse function is not really a true function in a mathematical sense. The impulse function has, however, been justified mathematically using a theory of "generalized functions."† In this approach, the im-

† M. J. Lighthill, *Fourier Analysis and Generalized Functions*, Cambridge University Press, 1959.

pulse function is defined as the limit of a sequence of regular well-behaved functions having the required property that the area remains a constant (unity) as the width is reduced. All operations are performed on this sequence. Finally the limit of this sequence is taken to define the impulse function as the width is reduced toward zero.

The defining sequence of pulses is not unique and many pulse shapes can be chosen. In fact, the shape of the particular pulse is relatively unimportant as long as the sequence satisfies the conditions that (1) the sequence formed describes a function which becomes infinitely high and infinitesimally narrow in such a way that (2) the enclosed area is a constant (unity). For example, the following sequences all satisfy these conditions (see Fig. 2.30):

a) Rectangular pulse: $$\delta(t) = \lim_{\tau \to 0} \frac{1}{\tau}\left[u\left(t + \frac{\tau}{2}\right) - u\left(t - \frac{\tau}{2}\right)\right]. \tag{2.103}$$

b) Triangular pulse: $$\delta(t) = \lim_{\tau \to 0} \frac{1}{\tau}\left[1 - \frac{|t|}{\tau}\right], \quad |t| < \tau. \tag{2.104}$$

c) Two-sided exponential: $$\delta(t) = \lim_{\tau \to 0} \frac{1}{\tau} e^{-|2t|/\tau}. \tag{2.105}$$

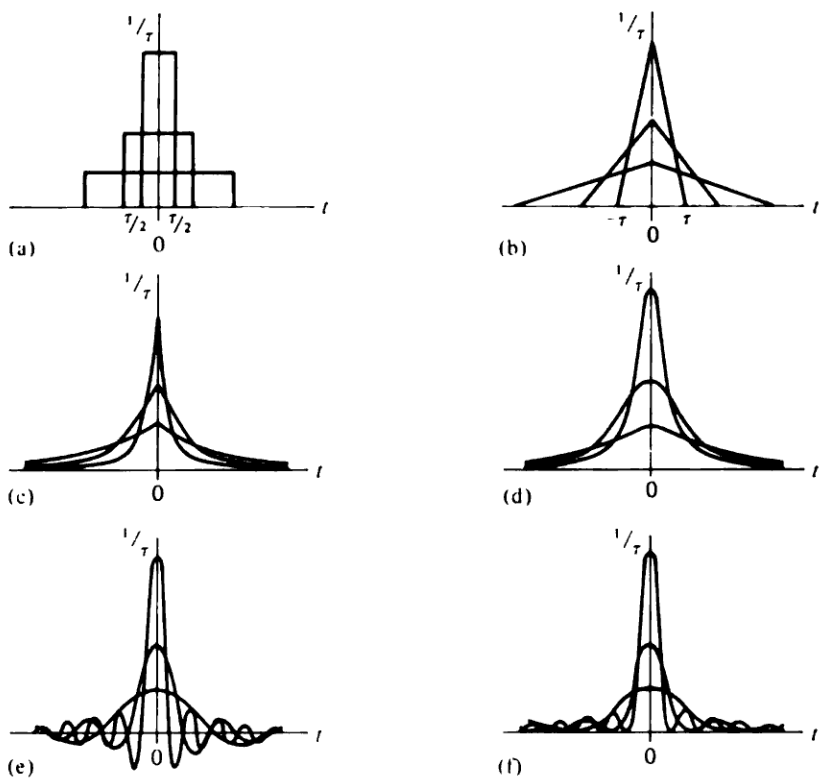


Figure 2.30 Function sequence definition of the impulse function: (a) rectangular pulse; (b) triangular pulse; (c) two-sided exponential; (d) gaussian pulse; (e) Sa(t) function; (f) Sa^2(t) function.

d) Gaussian pulse: $\delta(t) = \lim_{\tau \to 0} \frac{1}{\tau} e^{-\pi(t/\tau)^2}$. (2.106)

e) $\mathrm{Sa}(t)$ function: $\delta(t) = \lim_{\tau \to 0} \frac{1}{\tau} \mathrm{Sa}(\pi t/\tau)$. (2.107)

f) $\mathrm{Sa}^2(t)$ function: $\delta(t) = \lim_{\tau \to 0} \frac{1}{\tau} \mathrm{Sa}^2(\pi t/\tau)$. (2.108)

In Eq. (2.107), it is helpful to note that $1/\pi \int_{-\infty}^{\infty} \mathrm{Sa}(x)\,dx = 1$. This sequence can also be rewritten in slightly different form by letting $\pi/\tau = k$ so that

$$\delta(t) = \lim_{k \to \infty} \frac{k}{\pi} \mathrm{Sa}(kt). \tag{2.109}$$

In a similar manner, Eq. (2.108) can be rewritten as

$$\delta(t) = \lim_{k \to \infty} \frac{k}{\pi} \mathrm{Sa}^2(kt). \tag{2.110}$$

2.19 IMPULSE RESPONSE

To describe the response of a system in the time domain, we shall use reasoning analogous to that used in the frequency domain. The time function whose occurrence is the most definite of the elementary functions at our disposal (i.e., the one having the least time spread) is the impulse function. If we test the system with a unit impulse, the resulting output will describe the behavior of a system over all time in response to the unit impulse function occurring at one point in time. The resulting output is called the *impulse response* of the system.

This can be expressed mathematically in the following manner. If the input to the system is $f(t)$ and the corresponding output is $g(t)$, then, from Eq. (2.11), we have

$$g(t) = \mathcal{T}\{f(t)\}. \tag{2.111}$$

Specifically, we choose $f(t) = \delta(t)$, set all initial conditions equal to zero, and define the resulting output as the impulse response:

$$h(t) \triangleq \mathcal{T}\{\delta(t)\}. \tag{2.112}$$

For simple systems, Eq. (2.112) may be applied directly. For more complicated systems, a sequence such as one of those shown in Fig. 2.30 may be used for the input. After obtaining a solution for the corresponding output, the impulse response may be found by taking the limit.

EXAMPLE 2.19.1

Determine the impulse response of the system shown in Fig. 2.31(a).

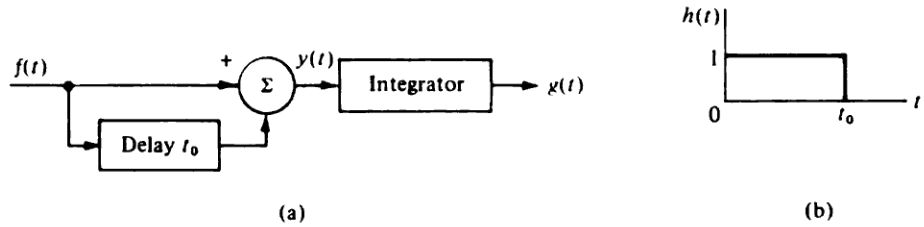


Figure 2.31 A holding circuit and its impulse response function.

Solution To find the impulse response, let $f(t) = \delta(t)$. Then (see Fig. 2.31a)

$$y(t) = \delta(t) - \delta(t - t_0),$$

and

$$h(t) = \int_{-\infty}^{t} [\delta(\tau) - \delta(\tau - t_0)]\,d\tau = u(t) - u(t - t_0).$$

This is an example of a "holding circuit" commonly used in pulse systems. Its impulse response is sketched in Fig. 2.31(b).

DRILL PROBLEM 2.19.1

Determine the impulse response of the system shown in Fig. 2.32 for (a) $K = 1$ and (b) $K < 1$. This type of system is useful in the detection of periodic signals if the period is known.

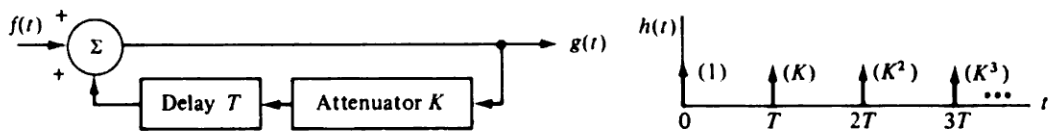


Figure 2.32 A recirculating-type filter and its impulse response.

ANSWER: (a) $h(t) = \sum_{n=0}^{\infty} \delta(t - nT)$; (b) $h(t) = \sum_{n=0}^{\infty} (K)^n \delta(t - nT)$.

2.20 ★ CONVERGENCE OF THE FOURIER SERIES

Finally, with our added perspectives, we can now state some rather general results for the Fourier series. If a signal $f(t)$ meets the conditions, which are known as the "Dirichlet conditions," over the interval $(0, T)$:

1. $f(t)$ has only a finite number of maxima and minima in the interval T;
2. $f(t)$ has only a finite number of finite discontinuities (i.e., $f(t)$ is "piecewise continuous") in the interval T;
3. $f(t)$ satisfies the inequality

$$\int_0^T |f(t)|\,dt < \infty;$$

then a Fourier series representation can be written which converges to $f(t)$ at all points of continuity. These conditions are sufficient but not necessary, and we shall encounter some mathematical functions which do not satisfy the Dirichlet conditions and yet have a Fourier series representation.

The Dirichlet conditions are satisfied for signals which occur in the laboratory because we insist that the somewhat weaker condition

$$\int_0^T |f(t)|^2\,dt < \infty$$

always holds. Stated another way, a Fourier series can always be used to represent any signal which has finite energy over $(0, T)$.

For any function satisfying the Dirichlet conditions, the magnitude of the coefficients decreases as n becomes large at least as fast as n^{-1}; that is,[†]

$$|F_n| \le \frac{M_N}{n} \quad \text{for } n > N, \tag{2.113}$$

where M is some positive number that is independent of n. If the given function has a derivative satisfying the Dirichlet conditions, the Fourier coefficients decrease at least as fast as n^{-2}. Likewise, if the given function has a second derivative satisfying the Dirichlet conditions, the Fourier coefficients decrease at least as fast as n^{-3}. In general, if the kth derivative satisfies the Dirichlet conditions, then we know that the magnitude of the Fourier coefficients will decrease at least as fast as $n^{-(k+1)}$ as n becomes very large.

The unit impulse function does not satisfy the Dirichlet conditions. However, the integral of a periodic sequence of impulse functions satisfies the Dirichlet conditions as long as the net area per period is zero. Also, the unit impulse function arises whenever a sufficient number of derivatives are taken of a piece-wise continuous function. Therefore the order of the derivative of a given function that first exhibits impulse functions represents that negative order of n which bounds the rate of decrease of the Fourier coefficients. For example, the magnitude of the Fourier coefficients for a square wave must decrease at least as fast as $1/n$ for large n, those for a

[†] The analogous relations for the trigonometric Fourier series are

$$|a_n| \le \frac{M_N}{n}, \; |b_n| \le \frac{M_N}{n} \quad \text{for } n > N.$$

triangular waveform as $1/n^2$, those for a waveform composed of parabolic sections as $1/n^3$, etc. Reasons why this is true will become evident when derivative properties of the Fourier transform are discussed in Chapter 3. For the present, these observations can serve as useful checks in our calculations of Fourier coefficients.

The Fourier series representation converges in mean-square (i.e., the mean-square error approaches zero uniformly and absolutely as $N \to \infty$) to $f(t)$ for every $f(t)$ satisfying the Dirichlet conditions.† For many functions this convergence is rapid enough that a truncation of the series after only a few terms gives a fairly good approximation to the original function.

At points of finite discontinuity in the function, the Fourier series representation converges to the arithmetic average value of the function to either side of the point of finite discontinuity. As the number of terms (N) in the Fourier series is increased, the mean-square error between the representation and the given function decreases and the approximation to the given function improves everywhere except in the immediate vicinity of a finite discontinuity. However, in the neighborhood of points of discontinuity in $f(t)$ the Fourier series representation fails to converge even though the mean-square error in the representation approaches zero. Hence even though the number of terms is increased, the representation exhibits overshoot near points of discontinuity of $f(t)$. This behavior is known as the Gibbs phenomenon, after its discoverer, and is shown in Fig. 2.33 for a periodic sawtooth waveform. Note that the

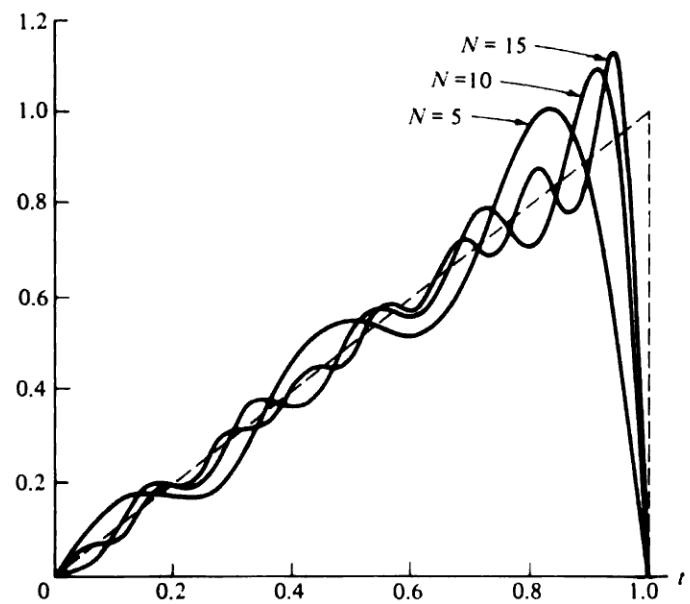


Figure 2.33 Gibbs phenomenon for a periodic sawtooth waveform.

† Let $f^N(t) = \sum_{n=1}^{N} f_n(t) = \lim_{N\to\infty} f^N(t)$ over $(0, T)$. If for each $\epsilon > 0$ there exists some integer M such that for every t in $(0, T)$, $|f^N(t) - f(t)| < \epsilon$, $M \le N$, we say that $f^N(t)$ converges *uniformly* to $f(t)$ in $(0, T)$. If $\sum_{n=1}^{N} |f_n(t)|$ converges uniformly, then $f^N(t)$ converges *absolutely and uniformly* to $f(t)$ in $(0, T)$.

overshoot peak moves closer to the point of discontinuity as more terms in the series are added. Even though the mean-square error approaches zero in the limit for large n, this amplitude overshoot approaches a value approximately 9% greater than the amplitude of the sawtooth waveform. Therefore "convergence in mean-square" implies convergence in amplitude in the neighborhood of points of continuity of $f(t)$ only.

2.21 SUMMARY

In a systems approach, an input signal and the resulting response are used to characterize the behavior of a given system. For our purposes a signal is defined to be a single-valued function of time. This function may be complex-valued for mathematical convenience but our observations in practice are real-valued.

Signals may be classified into basic categories by their finite energy or power. Other useful categories include those of periodic or aperiodic signals and random or deterministic signals.

Mathematically, a system is a rule for mapping inputs into outputs. If the system is linear, superposition applies. If the system response is not dependent on absolute values of time, but only on time differences, the system is time-invariant.

A useful analogy may be made between vectors and signals. Use of orthogonality allows the representation of a given vector or signal by a linearly independent set of known vectors or signals. Only the coefficients need to be determined for a given case. Many different sets of orthogonal functions or vectors may be chosen to represent a given signal or vector.

Operations of particular interest are those which can be described by linear differential equations with constant coefficients. The Fourier series proves very useful because the exponential form repeats under differentiation. The complex exponential and trigonometric functions yield coefficients in terms of frequency which retain a correct measure of energy and power. Parseval's theorem equates the energy and power in terms of frequency to that expressed in terms of time.

The Fourier series representation over an interval converges in an integral-square (energy) sense at points of continuity in the signal. Signals with a finite energy rate (power) may be handled by dividing by the interval. This is useful in representing periodic signals if the interval is equated to the period. Such representations converge in a mean-square sense at points of continuity in the signal.

A plot of the Fourier coefficients as a function of frequency is called the spectrum of the signal. The spectrum of a periodic signal consists of a set of equally spaced lines whose heights indicate the relative weights of each frequency component. A signal meeting certain conditions, known as the Dirichlet conditions, can always be described uniquely in terms of both time and frequency.

Singularity functions are simple mathematical functions which may not be finite everywhere or do not have finite derivatives of all orders. The unit impulse function is of particular importance in this class of functions and can be defined in terms of its

integral properties. The unit impulse function allows us to handle functions that exist only at a point and yet exert a dominant influence on the integration process. The unit step function is the indefinite integral of the unit impulse function.

A linear time-invariant system is determined uniquely by its transfer function or by its impulse response. The transfer function is a complex-valued function that can be found at a frequency by applying a complex exponential of fixed frequency to the input and dividing the steady-state output by this input. The impulse response can be found by applying a unit impulse to the input and recording the output waveform. By use of linearity, it may also be found by applying a unit step and differentiating the output waveform.

The output of a linear time-invariant system has only those frequencies that are present at the input. The magnitude and phase of each frequency term is modified by the system transfer function. Systems that are not linear or time-invariant can generate new frequency (harmonic) content. The Fourier spectrum is a very sensitive indicator of harmonic content.

Computation of Fourier coefficients from M numerical data points limits the computation to $M/2$ harmonic terms. Higher harmonic terms are reflected about $M/2$ with conjugate symmetry and affect the numerical accuracy of the computed coefficients. This is called aliasing. Effects of aliasing are minimized by a rapid decrease in Fourier coefficient magnitudes with increasing frequency and by choosing more sample points per interval. A bound on the rate of decrease of Fourier coefficients with increasing frequency can often be estimated from a knowledge of the signal waveform.

Selected References for Further Reading

1. A. Papoulis. *Circuits and Systems: A Modern Approach*. New York: Holt, Rinehart and Winston, 1980.
   Introduces the Fourier series with added emphasis on the discrete series and aliasing.
2. R. E. Ziemer, W. H. Tranter, and D.R. Fannin. *Signals and Systems*. New York: Macmillan, 1983.
   This text has a good discussion of the Fourier series, both the exponential and trigonometric forms, in Chapter 3.
3. B. P. Lathi. *Signals, Systems and Communication*. New York: John Wiley, 1965.
   Chapter 3 treats the analogy between signal and vector representations using a minimization of mean-square error.
4. R. A. Gabel and R. A. Roberts. *Signals and Linear Systems*. Third ed. New York: John Wiley & Sons, 1987.
   Chapter 5 has a good discussion of minimization of mean-square error that leads to the Fourier series representation.
5. A. V. Oppenheim, A. S. Willsky, and I. T. Young. *Signals and Systems*. Englewood Cliffs, N.J.: Prentice Hall, 1983.
   This text has a parallel treatment of discrete and continuous Fourier series, with a good discussion of convergence.
6. C. M. Close. *The Analysis of Linear Circuits*. New York: Harcourt, Brace & World, 1966.
   Chapter 9 has a good discussion of the Fourier series with examples and applications in linear circuits and the use of the frequency transfer function for periodic inputs.
7. H. S. Carslaw. *Theory of Fourier's Series and Integrals*. London: Cambridge University Press, 1930. (Reprinted by Dover Publications, New York, 1952.)

A classic treatment, at a more advanced level, of the trigonometric Fourier series and its applications to problems in heat transfer.

8. E. A. Guillemin. *The Mathematics of Circuit Analysis*. New York: John Wiley, 1949.
A more detailed and advanced treatment of the Fourier series, particularly on the topics of convergence and the Gibbs phenomenon.

PROBLEMS

2.2.1 Classify the following signals as energy signals or power signals, and find the normalized energy or normalized power of each. (All signals are defined over $-\infty < t < \infty$.)

a) 4
b) $\cos t + \cos 2t$
c) $\exp(-2|t|)$
d) $\exp(j2\pi t)$

2.2.2 Determine if each of the following signals is periodic. If a signal is periodic, determine its period.

a) $\cos(\sqrt{3}\,t)$
b) $\exp(j2\pi t - \pi/4)$
c) $\sin^2(2t)$
d) $\sin(2t^2)$

2.3.1 Let $f(t)$ be the input to a given system and $g(t)$ be the corresponding output. The input-output relationships of several systems are given below. Classify the systems whether they are linear, time-invariant, causal.

a) $g(t) = \dfrac{d}{dt}[f(t)]$
b) $g(t) = f(t) + f(-t)$
c) $g(t) = f(t/2)$
d) $g(t) = \exp[2f(t)]$

★ **2.4.1** Three vectors, expressed in the Cartesian coordinate system, are $\boldsymbol{\phi}_1 = \mathbf{x}_1 + \mathbf{x}_2$; $\boldsymbol{\phi}_2 = \mathbf{x}_1 - \mathbf{x}_2$; $\boldsymbol{\phi}_3 = \mathbf{x}_3$.

a) Show that $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, $\boldsymbol{\phi}_3$ are mutually orthogonal.
b) Represent the vectors **C** and **D** in Drill Problem 2.4.1 in terms of this set of orthogonal vectors.

★ **2.4.2** For the vectors $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, $\boldsymbol{\phi}_3$ in the preceding problem, determine the necessary conditions for the vector $\mathbf{Y} = a\mathbf{x}_1 + b\mathbf{x}_2 + c\mathbf{x}_3$ to be

a) Represented only by $\boldsymbol{\phi}_1$.
b) Orthogonal to $\boldsymbol{\phi}_1$.
c) Orthogonal to $\boldsymbol{\phi}_2$, $\boldsymbol{\phi}_3$ and have the same length as $\boldsymbol{\phi}_1$.

★ **2.5.1** Retaining only one term in the representation of $f(t)$ in Eq. (2.29), the remaining error over (t_1, t_2) is

$$\epsilon_1 = f(t) - f_1\phi_1(t).$$

a) Show that this error is orthogonal to $\phi_1(t)$ over (t_1, t_2).
b) Show that the error energy is the difference in energies between $f(t)$ and $f_1\phi_1(t)$ over (t_1, t_2).

★ **2.5.2** Find the first three terms in the generalized Fourier representation of the signal

$$f(t) = \begin{cases} \sin 2\pi t & 0 \le t < 1 \\ 0 & 1 \le t < 2 \end{cases}$$

over the interval (0, 2).

★ **2.5.3** A given set of functions, $\phi_n(t)$, is shown in Fig. P–2.5.3.†

a) Show that these functions form an orthogonal set over the interval (0, 1). Is the set an orthonormal set?

b) Represent the given signal $f(t) = 2t$ over the interval (0, 1) using this set of orthogonal functions.

c) Sketch $f(t)$ and the representation of $f(t)$ on the same graph and compare.

d) Compute the energy in each term of the series and the error energy remaining after each term is included.

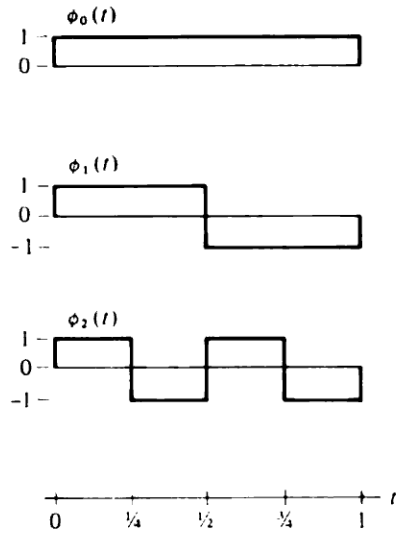


Figure P–2.5.3.

★ **2.5.4** A given set of three functions is‡

$$\phi_0(t) = 1; \qquad \phi_1(t) = t; \qquad \phi_2(t) = \frac{3}{2}t^2 - \frac{1}{2}.$$

a) Show that these functions are mutually orthogonal over the interval (−1, 1).

b) Represent the given signal

$$f(t) = \begin{cases} t & t > 0 \\ 0 & t < 0 \end{cases}$$

over the interval (−1, 1) using this set of functions.

c) Sketch $f(t)$ and the representation of $f(t)$ on the same graph and compare.

† These are selected functions from a larger set of orthogonal functions called Walsh functions. Designating members of the set by $w_n(t)$, where n is the number of zero crossings in the open interval (0, 1), we find the above set is formed by $w_0(t)$, $w_1(t)$, $w_3(t)$.

‡ These functions are the first three members of a set of orthogonal functions called Legendre polynomials.

2.7.1 The complex exponential Fourier series of the signal $f(t) = a \exp(-a|t|)$ over a certain symmetric interval $(-T/2, T/2)$ is

$$f(t) = \sum_{n=-\infty}^{\infty} \frac{a^2(1 - e^{-a} \cos \pi n)}{a^2 + n^2\pi^2} e^{jn\pi t}.$$

a) Determine the value of T.
b) What is the average value of $f(t)$? (Use two ways.)
c) The component of $f(t)$ at a certain frequency can be expressed as $A \cos 3\pi t$. Determine the numerical value of the constant A.
d) Write the first three nonzero terms in a series expansion for the value of $f(t)$ at $t = 0$.

2.7.2 Find the exponential Fourier series representation for the signal $f(t) = A \cos \omega_0 t$ and verify your result using Euler's identity, given by Eq. (2.43).

2.7.3 Find the exponential Fourier series representation of the raised-cosine pulse signal defined by

$$f(t) = \begin{cases} 1 + \cos 2\pi t & |t| < \frac{1}{2} \\ 0 & \text{elsewhere} \end{cases}$$

over the interval $(-1, 1)$.

2.7.4 The expression for the coefficients in the exponential Fourier series for a given $f(t)$ over (t_1, t_2) is given by Eq. (2.36). A new function, $g(t)$, is formed from $f(t)$ using the operations below. The series coefficients for $g(t)$ are designated by G_n. Determine the relations for G_n in terms of F_n for each of the following:

a) $g(t) = f(at)$;
b) $g(t) = f(t - t_0)$;
c) $g(t) = e^{j\omega_0 t} f(t)$.

2.8.1 Show that the following relations are true for the complex variable $z = x + jy$ and its conjugate $z^* = x - jy$:

a) $z + z^* = 2\ \mathcal{R}e\{z\}$
b) $z - z^* = 2j\ \mathcal{I}m\{z\}$
c) $(z_1 \pm z_2)^* = z_1^* \pm z_2^*$
d) $(z_1 z_2)^* = z_1^* z_2^*$
e) $\mathcal{R}e\{z_1 z_2\} = \mathcal{R}e\{z_1\}\ \mathcal{R}e\{z_2\} - \mathcal{I}m\{z_1\}\ \mathcal{I}m\{z_2\}$
f) $\mathcal{I}m\{z_1 z_2\} = \mathcal{R}e\{z_1\}\ \mathcal{I}m\{z_2\} + \mathcal{I}m\{z_1\}\ \mathcal{R}e\{z_2\}$

2.8.2 a) Show that if a signal has the even-symmetry property: $f(-t) = f(t)$ then the complex exponential Fourier coefficients are given by

$$F_n = \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} f(t) \cos n\omega_0 t\, dt.$$

b) Use this expression to find the F_n for

$$f(t) = \begin{cases} 2A|t|/T & |t| \le T/4 \\ 0 & |t| > T/4 \end{cases}$$

over $(-T/2, T/2)$.

c) What is the expression for the F_n if instead the signal has the odd-symmetry property: $f(-t) = -f(t)$?

2.8.3 In single-phase induction motors, the stator magnetic field is of the form: $A \cos \omega_0 t$, and can be visualized as two counter-rotating phasors of equal magnitude. For motor starting applications, it is necessary to develop a nonzero quadrature component (for some circular rotation). Develop an expression for the angle between magnetic fields if one of the phasors is delayed in time corresponding to an angle ψ. (This principle is used in the "shaded pole" induction motor.)

2.9.1 a) Determine the trigonometric Fourier series representation for the signal

$$f(t) = \begin{cases} 2 & -\frac{1}{2} < t < \frac{1}{2} \\ 0 & \text{elsewhere} \end{cases}$$

over the interval $(-1, 1)$.

b) Compare your results with those obtained in Example 2.5.1.

2.9.2 Represent the signal

$$f(t) = \begin{cases} 2 & -1 < t < 0 \\ 1 & 0 < t < 1 \\ 0 & \text{elsewhere} \end{cases}$$

over the interval $(-2, 2)$.

a) Use the exponential Fourier series.

b) Use the trigonometric Fourier series.

c) Compare your results using Eqs. (2.49)–(2.51).

2.10.1 One method used in the design of direct current (dc) power supplies is to rectify the input waveform. If the rectifier input is $f(t)$ and its output is $g(t)$, then the output of a full-wave rectifier is given by $g(t) = |f(t)|$. In this case let the input to the full-wave rectifier be: $f(t) = A \cos \omega_0 t$.

a) Sketch the input and output waveforms.

b) Determine the dc component in $f(t)$ and in $g(t)$.

c) Find the coefficients of the trigonometric Fourier series of $f(t)$ and $g(t)$ at the frequencies ω_0 and $2\omega_0$.

2.10.2 Repeat Problem 2.10.1 for the half-wave rectifier defined by

$$g(t) = \begin{cases} f(t) & f(t) > 0, \\ 0 & f(t) \leq 0. \end{cases}$$

2.10.3 The average normalized power in the error between two periodic waveforms is called the *mean-square error*. Compute and plot the mean-square error in the representation of a square wave (with zero average value) with its trigonometric Fourier series over the harmonic range $0 \leq n \leq 10$.

2.10.4 Show that if a periodic signal satisfies the half-wave symmetry condition $f(t) = -f(t \pm T/2)$, then all even harmonic amplitudes in the Fourier series vanish, with the possible exception of the $n = 0$ term.

2.11.1 Show that Parseval's theorem for the trigonometric Fourier series is

$$\frac{1}{T}\int_0^T [f(t)]^2\,dt = a_0^2 + \frac{1}{2}\sum_{n=1}^{\infty}(a_n^2 + b_n^2).$$

2.11.2 Apply Parseval's theorem to the symmetric square wave in Drill Problem 2.10.1 to show that

$$\sum_{\substack{n=1 \\ n\,\text{odd}}}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{8}.$$

√† **2.11.3** Compute and plot the percent of the average power contained within the first zero crossings in frequency, for the periodic gate signal shown in Fig. 2.21, as a function of the ratio T/τ. (Let T/τ be an integer, $2 \le T/\tau \le 5$.)

√ **2.11.4** Determine the minimum number of harmonic terms that must be retained in the periodic gate signal shown in Fig. 2.21 to include a) 90%; b) 95%; c) 99% of the average power in the signal for the choice: $T/\tau = 4$. (Do you expect it to change for higher integer values of T/τ?)

2.12.1 a) Determine the frequency transfer function of the system shown in Fig. 2.8.
b) Sketch the magnitude and phase as a function of frequency, and label the points at which the magnitude is down to $1/\sqrt{2}$ its maximum value.
c) Compare the results for this system with that of Example 2.12.1.

2.12.2 The differential equation relating the output voltage, $g(t)$, to the input voltage, $f(t)$, of the system shown in Fig. P–2.12.2 is

$$L^2C\frac{d^3g}{dt^3} + RLC\frac{d^2g}{dt^2} + 2L\frac{dg}{dt} + Rg(t) = Rf(t).$$

a) Find the frequency transfer function of the system.
b) Determine the magnitude response and the phase shift response of the system.
c) Repeat (b) for the special case in which $\omega = 1/\sqrt{LC}$.

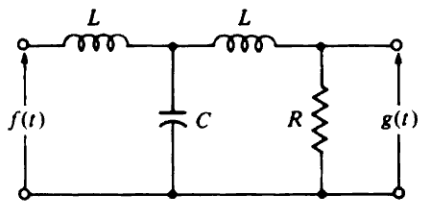


Figure P–2.12.2

2.13.1 The periodic signal in Problem 2.7.1 is applied to the input of a system whose frequency transfer function is

$$H(\omega) = \begin{cases} 1/(1 + j\omega/\pi) & |\omega| < 3\pi/2, \\ 0 & \text{elsewhere.} \end{cases}$$

a) Determine an expression for the output, $g(t)$, of the system.
b) Find the average power in the system output.

† Problems intended to be solved using numerical methods are designated with √ in the margin.

2.13.2 Determine the output, $g(t)$, of the RC lowpass filter in Example 2.12.1 for $f(t) = 10 \cos 4\pi t$ if

a) $RC = 1/\pi$;
b) $RC = 1/(4\pi)$;
c) $RC = 1/(16\pi)$.

2.14.1 A periodic signal $f(t)$ is shown in Fig. P–2.14.1. (For $a = 1$, this signal is a periodic square wave.)

a) Find the trigonometric Fourier series of $f(t)$ for arbitrary values of a.
b) Determine the effect on the amplitude of the second harmonic of $f(t)$ if there is a very small departure from perfect square-wave symmetry. To do this, let $a = 1 - \epsilon$, where $\epsilon \ll 1$, and find the second harmonic dependence on ϵ. The parameter a is sometimes called the signal symmetry factor (where $0 \leq a \leq 1$).
c) Repeat (b) for the third harmonic of $f(t)$.

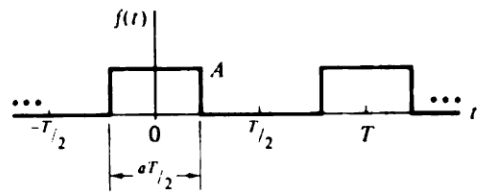


Figure P–2.14.1.

2.14.2 A potential problem in push-pull amplifiers is that a voltage pedestal ϵA may be introduced into the sinusoidal waveform, as shown in Fig. P–2.14.2.

a) Develop an expression for the resulting total harmonic distortion (THD).
b) Compute the THD for $\epsilon = 0.10$.

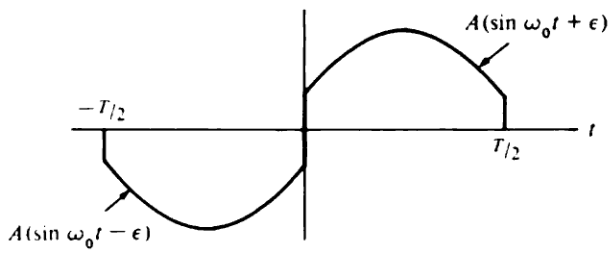


Figure P–2.14.2.

2.14.3 A proposed method for generation of harmonic content is the triangular waveform with fixed slope shown in Fig. P–2.14.3, where $0 < \tau < T/2$.

a) Determine the amplitude of a given harmonic n.
b) Find the optimum value of τ for generation of a given harmonic n.
c) For this optimum value of τ, compute the average power in a given harmonic, normalized to the averge power in the triangular waveform.

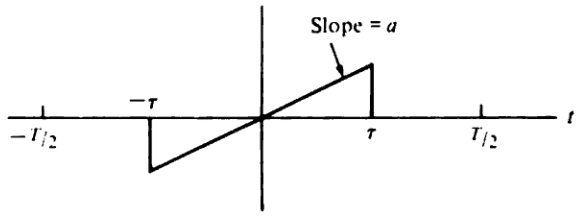


Figure P–2.14.3.

2.15.1 The line spectrum, in volts, of a certain periodic function $f(t)$ is shown in Fig. P–2.15.1.

a) What is the value of the period, T?
b) What is the value of $f(0)$?
c) What is the value of $f(\frac{1}{2})$?
d) A second function, $g(t)$, is related to $f(t)$ by $g(t) = f(t/2)$. Sketch the line spectrum of $g(t)$.

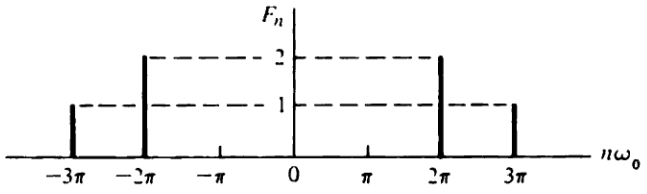


Figure P–2.15.1.

2.15.2 Sketch the Fourier line spectrum of the exponential and trigonometric series for the first three harmonic terms (i.e., $|n| \leq 3$) of the waveform in Problem 2.7.1 for the case $a = 1$.

2.15.3 The output, $g(t)$, of a half-wave rectifier for input $f(t)$ can be written as

$$g(t) = \begin{cases} f(t) & \text{if } f(t) > 0, \\ 0 & \text{if } f(t) \leq 0. \end{cases}$$

a) Find the trigonometric Fourier series for the output if the input is $f(t) = \sin 2\pi t$.
b) Sketch the Fourier magnitude line spectrum for both $f(t)$ and $g(t)$.

2.15.4 The following facts are known about a given waveform $f(t)$:

i) It has zero average value.
ii) It is real-valued and has even symmetry in t.
iii) Coefficients for harmonic terms $|n| = 1$ and $|n| \geq 5$ are zero.
iv) The following sample points are known: $f(0) = 4$; $f(T/4) = -1$; $f(T/2) = 2$.

a) Sketch the Fourier line spectrum of the exponential and trigonometric series for $f(t)$ if $T = 1$.
b) Calculate the percentage of the total average power in the waveform for $|n| > 2$.

√★ **2.16.1** The trapezoidal waveform $f(t)$ shown in Fig. P–2.16.1 is proposed for generation of harmonic content. Using numerical methods, compute and plot the magnitude line spectrum of the trigonometric Fourier series for this waveform. Use 100 points per period and compute the first 20 harmonics for $t_1 = 1$, 2, and 5 sample points, and compare the results.

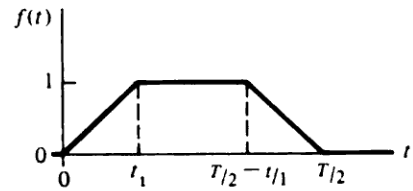


Figure P–2.16.1.

√★ **2.16.2** Compute and plot the magnitude line spectrum of the trigonometric Fourier series for the waveform shown in Fig. P–2.16.2, using the directions of Problem 2.16.1 and $A = 1$, $t_1 = T/4$. At points of finite discontinuity, assume continuity to the left.

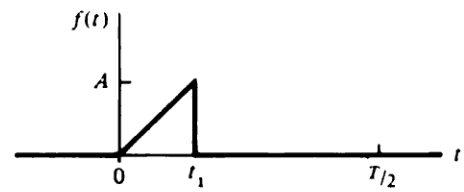


Figure P–2.16.2.

√★ **2.16.3** Compute and plot the magnitude of the third harmonic of the waveform shown in Fig. P–2.16.2 as a function of the parameter t_1 $(0 < t_1 \leq T/2)$. Use the directions of Problem 2.16.1, and normalize your plot so that the area under the waveform squared remains a constant with changes in t_1.

√★ **2.17.1** Compute the exponential Fourier series coefficients for the first 10 positive harmonic frequencies and for the first ten negative harmonic frequencies for the waveform $f(t)$ described by

$$f(t) = 5 \exp(-t), \qquad 0 \leq t \leq 2\pi .$$

Use 20 sample points over the period $T = 2\pi$ and follow the directions of Problem 2.16.2 at the finite discontinuities.

√★ **2.17.2** Compute and plot the magnitude line spectrum of the exponential Fourier series for the waveform

$$f(t) = \begin{cases} 4 \cos \pi t & |t| < 0.5 \\ 0 & |t| > 0.5 \end{cases}$$

when repeated periodically with period $T = 2$. Use 20 sample points per period, and plot over the integer frequency increments: $0, \pm 1, \pm 2, \ldots, \pm 20$.

2.18.1 Show that a periodic train of unit impulse functions, spaced T seconds apart, may be represented by the trigonometric Fourier series:

$$\frac{1}{T} + \frac{2}{T} \sum_{n=1}^{\infty} \cos n\omega_0(t - \tau), \qquad \omega_0 = 2\pi/T,$$

if one of the impulses is at $t = \tau$.

2.18.2 Evaluate the following integrals:

a) $\int_{-\infty}^{\infty} \delta(t-2)e^{-t}\cos \pi(t-1)\,dt$

b) $\int_{-\infty}^{t} \delta(\tau-3)e^{-\tau}\,d\tau$

c) $\int_{-\infty}^{\infty} (t^2+4)\,\delta(4-t)\,dt$

d) $\int_{-\infty}^{\infty} (t^2+4)\,\delta(4-2t)\,dt$

e) $\int_{0}^{3}\int_{-\infty}^{\infty} (t^2+4)\,\delta(t-\tau)\,d\tau\,dt$

2.18.3 Evaluate the following integrals:

a) $\int_{-\infty}^{\infty} tu(t)u(2-t)\,dt$

b) $\int_{-\infty}^{\infty} t[u(t)-u(t-2)]\,dt$

c) $\int_{-\infty}^{\infty} e^{-(t-t_1)}u(t-t_1)\,dt$

d) $\int_{-\infty}^{\infty} \delta(t-t_0)e^{-(t-t_1)}u(t-t_1)\,dt$

e) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(t+\tau)}u(t+\tau)\,\delta(t-\tau)\,d\tau\,dt$

2.18.4 Show that the following properties hold for the derivative of an impulse function defined under the operation of integration.

a) $\int_{-\infty}^{\infty} f(t)\,\delta'(t)\,dt = -f'(0)\,, \qquad \text{where } f'(t) = \dfrac{df}{dt}$

b) $\int_{-\infty}^{\infty} f(t)\,\delta^{(n)}(t)\,dt = (-1)^n f^{(n)}(0)$

c) $t\delta'(t) = -\delta(t)$

2.19.1 A rectangular pulse of width $\Delta\tau$ and height $1/\Delta\tau$ is applied to the input of a RC low-pass filter (cf. Fig. 2.16) at $t = 0$.

a) Write an expression for the output signal.

b) Take the limit as $\Delta \to 0$ and show that the impulse response of the filter is

$$h(t) = \frac{1}{RC}e^{-t/RC}u(t)\,.$$

c) Determine the dimensions of the voltage transfer function, $H(\omega)$, and the impulse response, $h(t)$.

2.19.2 The impulse response of a given system is: $h(t) = ae^{-at}u(t)$.

a) Determine the unit step response of the system.

b) For a known input, the output $g(t)$ can be expressed in terms of the input $f(t)$ and the system impulse response $h(t)$ as†

$$g(t) = \int_{-\infty}^{\infty} h(\tau)f(t-\tau)\,d\tau .$$

Evaluate this integral to determine $g(t)$ for the input $f(t) = \cos \omega_0 t$, and express your answer in terms of $\sin \omega_0 t$ and $\cos \omega_0 t$. (You may wish to consider use of exponential notation and then take a real part after the integration.)

★ **2.20.1** Estimate the minimum rate of convergence (for large n) of the Fourier coefficients for the periodic waveforms shown in Fig. P–2.20.1.

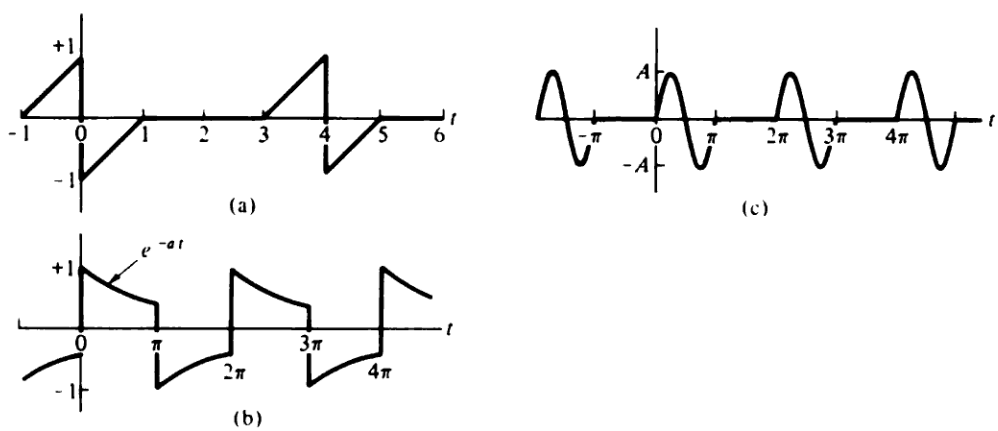


Figure P–2.20.1.

★ **2.20.2** Estimate the minimum rate of convergence (for large n) of the Fourier coefficients for the waveforms described in

a) Problem 2.14.1;

b) Problem 2.14.2;

c) Problem 2.14.3.

† This integral, known as the convolution integral, is discussed in Chapter 3.

DESIGN PROBLEMS

Design Problem

2.1 Waveform Design for Harmonic Generation

An important concept in the design of power amplifiers is that of instantaneous power, defined as the product of voltage times current. For a 1-ohm resistive load, this is given by the (instantaneous) value of the waveform magnitude squared.

In this problem we consider some aspects in the waveform design of a frequency multiplier using power amplification and frequency-selective filtering. Specifically, we wish to obtain the maximum magnitude of a fifth-harmonic component from a periodic waveform that you may design, with the restriction that the peak instantaneous power of the waveform may not exceed three times the average power in the waveform. Exercise good judgment in your waveform design, then use the computer to verify the correctness of your choice(s). Compare your results with the magnitude of a square wave with zero average value and with the same average power as your waveform.

CHAPTER

3 The Fourier Transform and Applications

In Chapter 2 we found how to represent a given function in terms of a complex exponential (or trigonometric) series over a finite interval. For the particular case in which the function is periodic, this representation can be extended over the entire interval $(-\infty, \infty)$.

Now let us consider another limiting case: the frequency domain description of one signal pulse. Our approach will be to consider a periodic signal and represent it in terms of complex exponentials with a period approaching infinity to yield, at least formally, a representation for the aperiodic signal. This representation will completely describe the given aperiodic signal *for all values of time*.

3.1 REPRESENTATION OF AN APERIODIC FUNCTION OVER THE ENTIRE REAL LINE

In Chapter 2 we discussed the Fourier exponential series representation of a given time function. It was observed that as the period T is made larger, the fundamental frequency becomes smaller and the frequency spectrum becomes more dense while the amplitude of each frequency component decreases. The *shape* of the spectrum, however, remains unchanged with varying T.

Suppose we are given an aperiodic function $f(t)$, as shown in Fig. 3.1(a). We wish to represent this function as a sum of exponential functions over the entire interval $(-\infty, \infty)$. For this purpose, we construct a new periodic function $f_T(t)$ with period T so that the function $f(t)$ is forced to repeat itself completely every T seconds.

The original function can be obtained back again by letting $T \to \infty$, that is,

$$\lim_{T \to \infty} f_T(t) = f(t). \tag{3.1}$$

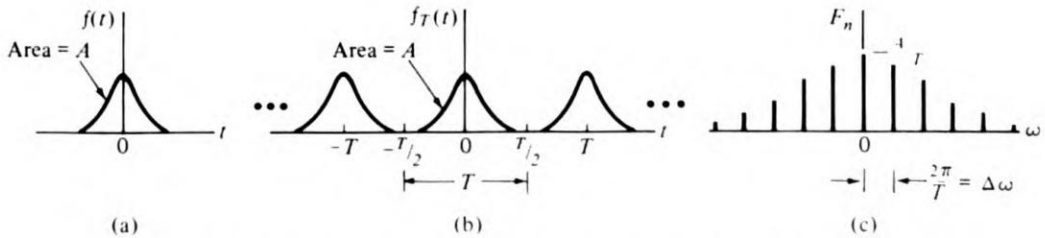


Figure 3.1 Generation of a periodic function and its line spectrum.

In other words, taking the limit as the period approaches infinity effectively removes all the nearest neighbors of the pulse $f(t)$ to infinity. To make this method effective, we adopt the following procedure: (1) the desired pulse is located at the origin, and (2) the limit as $T \to \infty$ is taken symmetrically, measured from the origin.

The new function $f_T(t)$ is a periodic function and consequently can be represented by an exponential Fourier series. This exponential series is written as

$$f_T(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t}, \tag{3.2}$$

where

$$F_n = \frac{1}{T}\int_{-T/2}^{T/2} f_T(t) e^{-jn\omega_0 t}\, dt \tag{3.3}$$

and

$$\omega_0 = 2\pi/T. \tag{3.4}$$

Before taking any limiting operation, however, we have to adjust things a little so that the magnitude components of the F_n don't all go to zero as the period is increased. We make the following changes:

$$\omega_n \triangleq n\omega_0, \tag{3.5}$$

$$F(\omega_n) \triangleq TF_n. \tag{3.6}$$

When we use these definitions, Eqs. (3.2) and (3.3) become

$$f_T(t) = \sum_{n=-\infty}^{\infty} \frac{1}{T} F(\omega_n) e^{j\omega_n t}, \tag{3.7}$$

$$F(\omega_n) = \int_{-T/2}^{T/2} f_T(t) e^{-j\omega_n t}\, dt. \tag{3.8}$$

The spacing between adjacent lines in the line spectrum of $f_T(t)$ is (see Fig. 3.1c)

$$\Delta\omega = 2\pi/T. \tag{3.9}$$

Using this relation for T in Eq. (3.7), we get the alternate form

$$f_T(t) = \sum_{n=-\infty}^{\infty} F(\omega_n) e^{j\omega_n t} \frac{\Delta\omega}{2\pi}. \tag{3.10}$$

Now as T becomes very large, $\Delta\omega$ becomes smaller and the spectrum becomes denser (i.e., the lines become very close and the frequency components within a given frequency range increase). In the limit, as $T \to \infty$, the discrete lines in the spectrum of $f_T(t)$ merge and the frequency spectrum becomes continuous. Mathematically, the infinite sum in Eq. (3.10) becomes the ordinary Riemann integral so that

$$\lim_{T\to\infty} f_T(t) = \lim_{T\to\infty} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(\omega_n) e^{j\omega_n t} \Delta\omega$$

becomes

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\,d\omega\,. \tag{3.11}$$

In a similar manner, Eq. (3.8) becomes

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}\,dt\,. \tag{3.12}$$

The results expressed in Eqs. (3.11) and (3.12) are commonly referred to as the *Fourier transform pair*. Equation (3.12) is known as the direct or forward Fourier transform of $f(t)$ (more commonly, just the *Fourier transform*). Equation (3.11) is known as the *inverse Fourier transform*. Symbolically, this suggests the following operator notation:

$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}\,dt\,, \tag{3.13}$$

$$f(t) = \mathcal{F}^{-1}\{F(\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\,d\omega\,. \tag{3.14}$$

It is also useful to note that the complex exponential Fourier series coefficients can be evaluated in terms of the Fourier transform by combining Eqs. (3.5) and (3.6) to give

$$F_n = \frac{1}{T}F(\omega)\bigg|_{\omega=n\omega_0}\,, \tag{3.15}$$

provided that $F(\omega)$ is finite at $\omega = n\omega_0$[†] and that $f(\pm T/2) \approx 0$.

DRILL PROBLEM 3.1.1

Find the Fourier transform of $f(t) = \exp(-at)\,u(t)$, $a > 0$.[‡]

ANSWER: $1/(a + j\omega)$.

The Fourier transform can be visualized as the representation of a given signal in terms of an infinite sum of complex exponentials each weighted by $F(\omega)\,df$ [cf. Eq. (3.11)]. The concept of a phasor representation of the complex exponential and the addition of phasors using the rules for vector addition, as described in Section 2.8, aids in the visualization. To illustrate this, consider a signal $f(t)$ for which $F(\omega)$ is a constant for all ω. At $t = 0$, all of the complex exponential terms formed by

[†] A sufficient condition for this to occur is that the energy in $f(t)$ is finite; see Section 3.3.

[‡] For those acquainted with the Laplace transform, the Fourier transform can be found from the Laplace transform by setting $s = j\omega$ if *both* of the following conditions are true: (1) $f(t) = 0$ for $t < 0$; (2) $f(t)$ has finite energy, that is, $\int_{-\infty}^{\infty} |f(t)|^2\,dt < \infty$. [See, for example, Reference 1 at the end of this chapter.]

$[F(\omega)\,df]e^{j\omega t}$ have identical magnitudes and zero phase angle. Because there is an infinite number of these, the real part of the resultant is infinitely large. At values of time for which $t \neq 0$, however, the complex exponential terms have identical magnitudes but phases that are proportional to frequency. The net result is that they add up to zero. Therefore we conclude that this signal $f(t)$ can be described by a function that is infinite at $t = 0$ and is zero for $t \neq 0$, suggesting description by an *impulse* function. Reference to Fourier transform pair 8 in Table 3.1 (and illustrated in Fig. 3.3) will show this conclusion is correct.

Reconstruction of $f(t)$ from a set of complex exponentials is useful for visualization purposes. However, these results are qualitative and must be visualized at various values of time for us to draw general conclusions. A convenient manner to present the same information more precisely and independent of time—but without the intuitive aspects—is to make a graph of $F(\omega)$ versus ω. This is investigated in the next section.

3.2 THE SPECTRAL DENSITY FUNCTION

Equation (3.14) represents $f(t)$ as a continuous sum of exponential functions with frequencies lying in the interval $(-\infty, \infty)$. The relative amplitude of the components at any frequency ω is proportional to $F(\omega)$. If the signal $f(t)$ represents a voltage, $F(\omega)$ has the dimensions of voltage multiplied by time. Because frequency has the dimensions of inverse time, we can consider $F(\omega)$ as a voltage-density spectrum or, more generally speaking, it is known as the *spectral-density function* of $f(t)$.

The *area* under the spectral-density function $F(\omega)$ has the dimensions of voltage (current). Each point on the $F(\omega)$ curve contributes nothing to the representation of $f(t)$; it is the area that contributes. On the other hand, each point does indicate the *relative weighting* of each frequency component. The contribution of a given frequency band to the representation of $f(t)$ may be found by integrating to find the desired area.

In contrast to this, a periodic waveform has all its amplitude components at discrete frequencies. At each of these discrete frequencies there is some definite contribution; to either side there is none. It follows, then, that to portray the amplitude components of a periodic waveform on a spectral-density graph requires a representation with area equal to the respective amplitude components yet occupying zero frequency width. We recognize that this can be done formally by representing each amplitude component of the periodic function by an impulse function. The area (weight) of the impulse is equal to the amplitude component and the position of the impulse is determined by the particular discrete frequency.

Summarizing, a signal of finite energy can be described by a continuous spectral-density function. This spectral-density function is found by taking the Fourier transform of the signal. A periodic signal of finite average power can be described either by a set of lines on a spectral graph or by a set of impulse functions on a spectral-

density graph. Each impulse on the latter graph has an area corresponding to the height of each line, respectively, on the former graph.

For convenience in drawing spectral-density graphs of periodic functions, the height of each impulse is often made proportional to its area. Though technically not correct, this convention is very convenient because it makes the spectral-density plots resemble the spectral plots.

The above representation works out fine for spectral functions expressed in terms of inverse time (frequency). We choose to use radian frequency, ω. By the scaling property of the impulse function [see Eq. (2.99)], we can write

$$\delta(f) = 2\pi\,\delta(\omega),$$

and therefore we multiply each impulse function by 2π if it is expressed in radian frequency.

Let us find the Fourier transform (spectral density) of a gate function having unit amplitude, width τ, and centered on the origin. This special function is defined by†

$$\text{rect}\,(t/\tau) \triangleq \begin{cases} 1 & |t| < \tau/2 \\ 0 & |t| > \tau/2 \end{cases}, \tag{3.16}$$

where τ designates the width. A unit gate function (i.e., $\tau = 1$) is shown in Fig. 3.2(a).

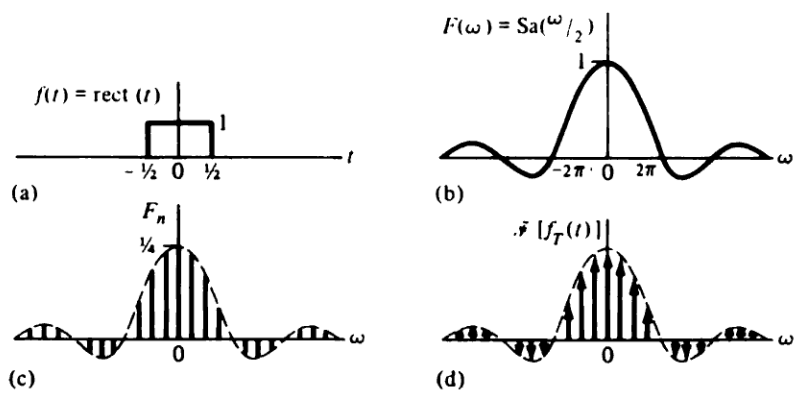


Figure 3.2 (a) The unit gate function and (b) its Fourier transform. (c) The line spectrum and (d) spectral density of the periodic gate function.

Using Eq. (3.12), we have

$$F(\omega) = \int_{-\infty}^{\infty} \text{rect}\,(t/\tau)e^{-j\omega t}\,dt = \int_{-\tau/2}^{\tau/2} e^{-j\omega t}\,dt$$

$$= (e^{j\omega\tau/2} - e^{-j\omega\tau/2})/(j\omega) = \tau\frac{\sin\,(\omega\tau/2)}{(\omega\tau/2)}$$

† Some authors use the symbols $G(t/\tau)$ or $\Pi(t/\tau)$ to designate the gate function. It can also be expressed in terms of unit step functions as $[u(t + \tau/2) - u(t - \tau/2)]$.

or, using Eq. (2.82),

$$\mathcal{F}\{\text{rect}\,(t/\tau)\} = \tau\ \text{Sa}\,(\omega\tau/2)\,. \tag{3.17}$$

It is helpful to note that because $\lim_{x\to 0} \text{Sa}\,(x) = 1$, the amplitude coefficient of $\text{Sa}\,(\omega\tau/2)$ in Eq. (3.17) is equal to the area beneath rect (t/τ). The result in Eq. (3.17) for $\tau = 1$ is shown in Fig. 3.2(b).

EXAMPLE 3.2.1

Determine the coefficients of the Fourier exponential series (the F_n) for the unit gate function if it is repeated every 4 seconds; plot on both a spectral and a spectral-density graph.

Solution Using Eq. (3.15), we have

$$F_n = \frac{1}{T}F(\omega)\bigg|_{\omega=n\omega_0} = \frac{1}{4}\,\text{Sa}\,(n\pi/4)\,.$$

The line spectrum and the spectral density are shown in Fig. 3.2(c) and (d).

DRILL PROBLEM 3.2.1

Find the Fourier transform of the two-sided exponential $f(t) = \exp\,(-a|t|)$, $a > 0$.

ANSWER: $2a/(a^2 + \omega^2)$.

3.3 ★ EXISTENCE OF THE FOURIER TRANSFORM

Sufficient conditions for the existence of the Fourier series, known as the Dirichlet conditions, were given in Section 2.20. These conditions, appropriately modified, also are sufficient for the existence of the Fourier transform:

1. $f(t)$ has only a finite number of maxima and minima in any finite time interval;
2. $f(t)$ has only a finite number of finite discontinuities in any finite time interval;
3. $f(t)$ is absolutely integrable; that is, $\int_{-\infty}^{\infty}|f(t)|\,dt < \infty$.

A somewhat weaker condition is $\int_{-\infty}^{\infty}|f(t)|^2\,dt < \infty$. Because this condition corresponds to our definition of an energy signal, we can state that the Fourier transform can be used to uniquely represent any energy signal.

3.4 PARSEVAL'S THEOREM FOR ENERGY SIGNALS

The energy delivered to a 1-ohm resistor is [Eq. (2.5)]:

$$E = \int_{-\infty}^{\infty} |f(t)|^2 \, dt = \int_{-\infty}^{\infty} f(t) f^*(t) \, dt . \tag{3.18}$$

We would like to express the energy in terms of the frequency components of $f(t)$. Using Eq. (3.11) in Eq. (3.18), we get

$$E = \int_{-\infty}^{\infty} f(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} F^*(\omega) e^{-j\omega t} \, d\omega \right] dt .$$

Interchanging the order of integration on t and ω,[†]

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} F^*(\omega) \left[\int_{-\infty}^{\infty} f(t) e^{-j\omega t} \, dt \right] d\omega . \tag{3.19}$$

But from Eq. (3.12), this is

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} F^*(\omega) F(\omega) \, d\omega . \tag{3.20}$$

Combining Eqs. (3.18) and (3.20), we obtain what is known as *Parseval's theorem for energy signals:*[‡]

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 \, d\omega . \tag{3.21}$$

From this, we can find the energy of a signal in either the time domain or the frequency domain.

In the MKS system of units, energy is measured in joules. From Eq. (3.21) we observe that the quantity $|F(\omega)|^2$ is an energy density having the units of joules per Hz (note that $\omega/2\pi$ is measured in Hz).

DRILL PROBLEM 3.4.1

Calculate the energy in the signal of Drill Problem 3.2.1 (across a 1-ohm resistance) in both the time domain and the frequency domain.

ANSWER: $1/a$ joules.

[†] This assumes that the integrals are absolutely convergent, i.e., that they are convergent when the integrands are replaced by their absolute values (a sufficient condition).

[‡] This relation is also known as Plancherel's theorem or Rayleigh's theorem.

3.5 SOME FOURIER TRANSFORMS INVOLVING IMPULSE FUNCTIONS

The procedure for finding the Fourier transform of signals of finite energy using Eq. (3.13) is straightforward. In contrast, this is not always true for signals of infinite energy. Of particular interest to us are those cases that involve the unit impulse function. Our discussions will be brief; the interested reader is referred to the references for more detailed discussions.

3.5.1 The Impulse Function

Using the integral properties of the impulse function, the Fourier transform of a unit impulse, $\delta(t)$, is

$$\mathcal{F}\{\delta(t)\} = \int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}\,dt = e^{j0} = 1\,. \tag{3.22}$$

If the impulse is time-shifted, we have

$$\mathcal{F}\{\delta(t - t_0)\} = \int_{-\infty}^{\infty} \delta(t - t_0)e^{-j\omega t}\,dt = e^{-j\omega t_0}. \tag{3.23}$$

It is evident from Eqs. (3.22) and (3.23) that an impulse function has a uniform magnitude spectrum over the entire frequency interval $(-\infty, \infty)$. This type of spectrum is called "white" in an analogy to white light. The phase spectrum of the time-shifted impulse is linear with a slope that is proportional to the time shift.

3.5.2 The Eternal Complex Exponential

We could expect that the spectral density of $e^{\pm j\omega_0 t}$ will be concentrated at $\pm\omega_0$. That this is the case is demonstrated below.

$$\begin{aligned}\mathcal{F}^{-1}\{\delta(\omega \mp \omega_0)\} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta(\omega \mp \omega_0)e^{j\omega t}\,d\omega \\ &= \frac{1}{2\pi}e^{\pm j\omega_0 t}.\end{aligned} \tag{3.24}$$

Taking the Fourier transform of both sides of Eq. (3.24), we have

$$\mathcal{F}\mathcal{F}^{-1}\{\delta(\omega \mp \omega_0)\} = \frac{1}{2\pi}\mathcal{F}\{e^{\pm j\omega_0 t}\}$$

or, by interchanging sides,

$$\mathcal{F}\{e^{\pm j\omega_0 t}\} = 2\pi\mathcal{F}\mathcal{F}^{-1}\{\delta(\omega \mp \omega_0)\},$$

$$\mathcal{F}\{e^{\pm j\omega_0 t}\} = 2\pi\delta(\omega \mp \omega_0). \tag{3.25}$$

This result should hardly surprise us because the complex exponential describes a phasor whose angular rate is ω_0. The spectral description of such a phasor is a line at this angular rate (which is greater than $\omega = 0$ if rotating in a positive direction, less than $\omega = 0$ if in a negative direction). The spectral-density description is an impulse at these angular rates and the factor of 2π arises because we are using radian frequency.†

In a more general sense, Eq. (3.25) may be used to define the unit impulse function. Note that for $\omega_0 = 0$, Eq. (3.25) simplifies to

$$\mathcal{F}\{1\} = 2\pi\delta(\omega). \tag{3.26}$$

It follows that any signal with a nonzero average value over the infinite interval $(-\infty, \infty)$ has an impulse in its spectral-density function at $\omega = 0$.

3.5.3 Eternal Sinusoidal Signals

The sinusoidal signals $\cos \omega_0 t$ and $\sin \omega_0 t$ can be written in terms of the complex exponentials using Euler's identities [see Eqs. (2.43) and (2.44)]. Their Fourier transforms can be found directly from Eq. (3.25):

$$\begin{aligned}\mathcal{F}\{\cos \omega_0 t\} &= \mathcal{F}\{\tfrac{1}{2}e^{j\omega_0 t} + \tfrac{1}{2}e^{-j\omega_0 t}\} \\ &= \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0),\end{aligned} \tag{3.27}$$

$$\begin{aligned}\mathcal{F}\{\sin \omega_0 t\} &= \mathcal{F}\left\{\frac{1}{2j}e^{j\omega_0 t} - \frac{1}{2j}e^{-j\omega_0 t}\right\} \\ &= [\pi\delta(\omega - \omega_0) - \pi\delta(\omega + \omega_0)]/j.\end{aligned} \tag{3.28}$$

Graphs of the trigonometric functions $\cos \omega_0 t$ and $\sin \omega_0 t$ together with their spectral-density functions are shown in Fig. 3.3. Note that it is convenient to multiply Eq. (3.28) by the factor j so that the result is real-valued for the graph.

3.5.4 The Signum Function And The Unit Step

The signum function, sgn (t), is that function which changes sign when its argument is zero:

$$\operatorname{sgn}(t) = \frac{|t|}{t} = \begin{cases} 1 & t > 0 \\ 0 & t = 0. \\ -1 & t < 0 \end{cases} \tag{3.29}$$

† If written in terms of frequency in Hz, the factor of 2π disappears in Eqs. (3.25) and (3.26) because $\delta(\omega) = \delta(2\pi f) = [1/(2\pi)]\delta(f)$.

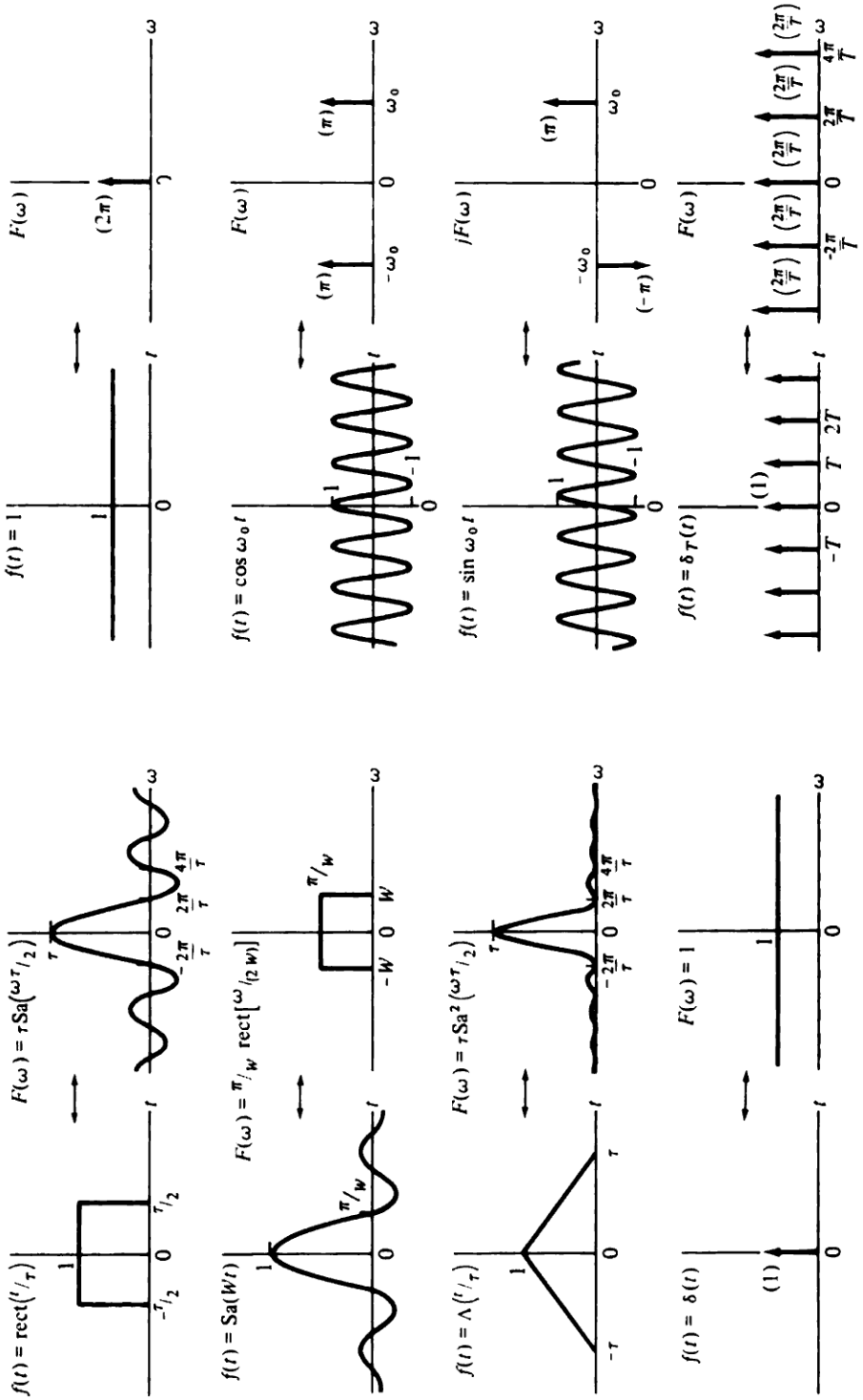


Figure 3.3 Some functions of time and their spectral-density functions.

The signum function has an average value of zero and is piecewise continuous, but not absolutely integrable. In order to make it absolutely integrable we multiply sgn (t) by $e^{-a|t|}$ and then take the limit as $a \to 0$:

$$\mathcal{F}\{\text{sgn}\,(t)\} = \mathcal{F}\left\{\lim_{a\to 0}\,[e^{-a|t|}\,\text{sgn}\,(t)]\right\}.$$

Interchanging the operations of taking the limit and integrating,

$$\begin{aligned}\mathcal{F}\{\text{sgn}\,(t)\} &= \lim_{a\to 0}\left\{\int_{-\infty}^{\infty} e^{-a|t|}\,\text{sgn}\,(t)e^{-j\omega t}\,dt\right\}\\ &= \lim_{a\to 0}\left\{-\int_{-\infty}^{0} e^{(a-j\omega)t}\,dt + \int_{0}^{\infty} e^{-(a+j\omega)t}\,dt\right\}. \qquad (3.30)\end{aligned}$$

Proceeding, we get

$$\begin{aligned}\mathcal{F}\{\text{sgn}\,(t)\} &= \lim_{a\to 0}\left\{\frac{-2j\omega}{a^2+\omega^2}\right\}\\ &= \frac{2}{j\omega}. \qquad (3.31)\end{aligned}$$

The unit step function, expressed in terms of its average value and the signum function, is

$$u(t) = \tfrac{1}{2} + \tfrac{1}{2}\,\text{sgn}\,(t). \qquad (3.32)$$

The Fourier transform is

$$\mathcal{F}\{u(t)\} = \tfrac{1}{2}\mathcal{F}\{1\} + \tfrac{1}{2}\mathcal{F}\{\text{sgn}\,(t)\},$$

which, using Eqs. (3.26) and (3.31), becomes

$$\mathcal{F}\{u(t)\} = \pi\,\delta(\omega) + 1/j\omega. \qquad (3.33)$$

Therefore the spectral-density function of the unit step contains an impulse at $\omega = 0$ corresponding to the average value of $\frac{1}{2}$ in the step function. It also has all the high frequency components of the signum function reduced by one-half.

Note that Eq. (3.30) is the integral of a complex function taken symmetrically about the origin. Taking the integral of a complex function in a nonsymmetrical manner may lead to errors. For example, if one attempted to evaluate the Fourier transform of the unit step directly,

$$\lim_{a\to 0}\int_{0}^{\infty} e^{-at}e^{-j\omega t}\,dt = 1/j\omega,$$

which is not in agreement with the correct result in Eq. (3.33).

3.5.5 Periodic Functions

In Section 3.2 we discussed the characteristics of the spectral-density function of a periodic signal in a qualitative manner. Now we are prepared to formalize the relationship in a more quantitative way.

We can express a function $f(t)$ that is periodic with period T by its exponential Fourier series

$$f_T(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t} \quad \text{where } \omega_0 = 2\pi/T .$$

Taking the Fourier transform, we find

$$\mathcal{F}\{f_T(t)\} = \mathcal{F}\left\{\sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t}\right\} .$$

If we assume the operations of integration and summation can be interchanged,

$$\mathcal{F}\{f_T(t)\} = \sum_{n=-\infty}^{\infty} F_n \, \mathcal{F}\{e^{jn\omega_0 t}\} .$$

Using Eq. (3.25), we get

$$\mathcal{F}\{f_T(t)\} = 2\pi \sum_{n=-\infty}^{\infty} F_n \, \delta(\omega - n\omega_0) . \tag{3.34}$$

Thus the Fourier transform (spectral density) of a periodic signal consists of a set of impulses located at the harmonic frequencies of the signal. The area (weight) of each impulse is 2π times the value of its corresponding coefficient in the exponential Fourier series. This result permits us to handle both periodic and nonperiodic functions in one unified treatment.

EXAMPLE 3.5.1

Find the spectral-density function of an even periodic square wave whose average value is zero, whose period is two seconds, and whose peak-to-peak amplitude is A.

Solution We shall use known Fourier transform relations to obtain coefficients for the Fourier series and then find the Fourier transform of the series. For a rectangular pulse of unit width and height A, Eq. (3.17) gives

$$F(\omega) = A \ \mathrm{Sa}\,(\omega/2) .$$

Because $\omega_0 = \pi$ here, Eq. (3.15) becomes

$$F_n = \frac{1}{T} F(\omega)\big|_{\omega = n\omega_0} = (A/2)\ \mathrm{Sa}\,(n\pi/2) .$$

Writing the series and noting that the average value is zero, we get

$$f_T(t) = (A/2) \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \mathrm{Sa}\,(n\pi/2) e^{jn\pi t} .$$

Using Eq. (3.34), the Fourier transform of this function is

$$\mathcal{F}\{f_T(t)\} = \pi A \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \mathrm{Sa}\,(n\pi/2)\,\delta(\omega - \pi n)\,.$$

The spectral-density function for this example is shown in Fig. 3.4. Note that the absence of an impulse at $\omega = 0$ is an indicator that the average value of the signal is zero.

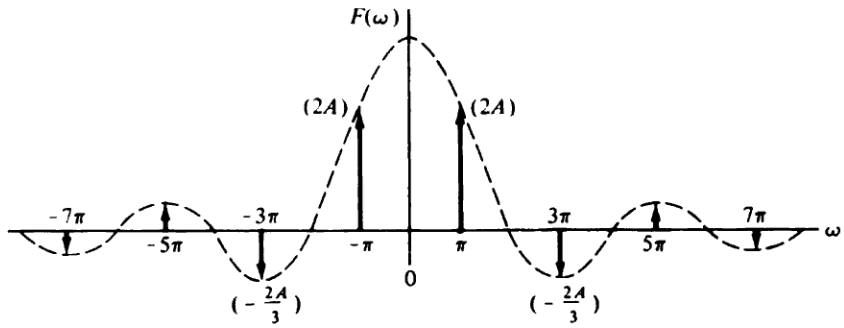


Figure 3.4 The spectral-density function of a periodic square wave with zero average value.

A periodic function of considerable importance in later chapters is that of a periodic sequence of unit impulse functions (sometimes referred to as a "comb" function). For convenience, we write such a sequence with period T as

$$\delta_T(t) \triangleq \sum_{n=-\infty}^{\infty} \delta(t - nT)\,. \tag{3.35}$$

Because this is a periodic function, we can express it in terms of a Fourier series by choosing $\omega_0 = 2\pi/T$,

$$\delta_T(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t},$$

where

$$F_n = \frac{1}{T}\int_{-T/2}^{T/2} \delta(t) e^{-jn\omega_0 t}\,dt = \frac{1}{T},$$

so that

$$\delta_T(t) = \frac{1}{T}\sum_{n=-\infty}^{\infty} e^{jn\omega_0 t}.$$

Thus the impulse train (comb) function contains components of the harmonic frequencies $\omega = 0, \pm\omega_0, \pm 2\omega_0, \ldots$ with equal weighting. Using Eq. (3.34), the Fourier

transform of the unit impulse train function is

$$\mathcal{F}\{\delta_T(t)\} = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_0)$$

$$= \omega_0 \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_0),$$

$$\mathcal{F}\{\delta_T(t)\} = \omega_0 \delta_{\omega_0}(\omega). \tag{3.36}$$

The Fourier transform of a periodic impulse train in the time domain gives an impulse train that is periodic in the frequency domain. The impulse train and its Fourier transform are shown in Fig. 3.3. A listing of some selected Fourier transform pairs appears in Table 3.1. (Not all entries in the table have been covered here, however.)

TABLE 3.1 Some Selected Fourier Transform Pairs

	$f(t)$	$F(\omega) = \mathcal{F}\{f(t)\}$
1.	$e^{-at}u(t)$	$1/(a + j\omega)$
2.	$te^{-at}u(t)$	$1/(a + j\omega)^2$
3.	$e^{-a\lvert t\rvert}$	$2a/(a^2 + \omega^2)$
4.	$e^{-t^2/(2\sigma^2)}$	$\sigma\sqrt{2\pi}\, e^{-\sigma^2\omega^2/2}$
5.	$\operatorname{sgn}(t)$	$2/(j\omega)$
6.	$j/(\pi t)$	$\operatorname{sgn}(\omega)$
7.	$u(t)$	$\pi\delta(\omega) + 1/(j\omega)$
8.	$\delta(t)$	1
9.	1	$2\pi\delta(\omega)$
10.	$e^{\pm j\omega_0 t}$	$2\pi\delta(\omega \mp \omega_0)$
11.	$\cos \omega_0 t$	$\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$
12.	$\sin \omega_0 t$	$-j\pi[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]$
13.	$\operatorname{rect}(t/\tau)$	$\tau \operatorname{Sa}(\omega\tau/2)$
14.	$\frac{W}{2\pi} \operatorname{Sa}(Wt/2)$	$\operatorname{rect}(\omega/W)$
15.	$\frac{W}{\pi} \operatorname{Sa}(Wt)$	$\operatorname{rect}(\omega/(2W))$
16.	$\Lambda(t/\tau)$	$\tau[\operatorname{Sa}(\omega\tau/2)]^2$
17.	$\frac{W}{2\pi}[\operatorname{Sa}(Wt/2)]^2$	$\Lambda(\omega/W)$
18.	$\cos(\pi t/\tau) \operatorname{rect}(t/\tau)$	$\frac{2\tau}{\pi} \frac{\cos(\omega\tau/2)}{1 - (\omega\tau/\pi)^2}$
19.	$\frac{2W}{\pi^2} \frac{\cos(Wt)}{1 - (2Wt/\pi)^2}$	$\cos[\pi\omega/(2W)] \operatorname{rect}[\omega/(2W)]$
20.	$\delta_T(t)$	$\omega_0 \delta_{\omega_0}(\omega)$, where $\omega_0 = 2\pi/T$

3.6 PROPERTIES OF THE FOURIER TRANSFORM

The Fourier transform is a method of expressing a given function of time (or any other appropriate coordinate, for that matter) in terms of a continuous set of exponential components of frequency. The resulting spectral-density function gives the relative weighting of each frequency component.

We shall use the convention that a time function (signal) and its Fourier transform are indicated by lowercase and uppercase letters respectively.† A convenient shorthand notation to show the correspondence between the two domains is to use a double arrow: $f(t) \leftrightarrow F(\omega)$. The Fourier transform relationship between time and frequency is assumed to be unique; that is, $\mathcal{F}^{-1}\mathcal{F}[f(t)] = f(t)$ for all t.

What happens in one domain when an elementary operation is performed on the function in the other domain? One way to find out is to go through the transform relation, perform the desired operation, and then go through the inverse transform relation. However, this procedure is clumsy and often quite difficult. The purpose of this section is to introduce some properties of the Fourier transform in a general form.

3.6.1 Linearity (Superposition)

The Fourier transform is a linear operation based on the properties of integration and therefore superposition applies. Thus for any arbitrary constants a_1 and a_2,

$$\mathcal{F}\{a_1 f_1(t) + a_2 f_2(t)\} = a_1 F_1(\omega) + a_2 F_2(\omega)\,, \tag{3.37}$$

or

$$a_1 f_1(t) + a_2 f_2(t) \leftrightarrow a_1 F_1(\omega) + a_2 F_2(\omega)\,.$$

This follows from the integral definition of the Fourier transform. Although the proof is trivial, the consequences of this property to the study of linear systems are of major importance.

Some caution must be exercised in adding spectral densities. Recall that in general $F(\omega)$ is complex-valued and therefore it is important that this addition be complex. A common error is that of adding magnitudes only without regard for phase.

3.6.2 Complex Conjugate

For any complex-valued signal, we have

$$\mathcal{F}\{f^*(t)\} = F^*(-\omega)\,. \tag{3.38}$$

† A departure from this convention is allowed when we wish to form special symbols using more than one letter; e.g., rect (t), rect (ω), Sa (t), Sa (ω), etc.

Proof

$$\begin{aligned}\mathcal{F}\{f^*(t)\} &= \int_{-\infty}^{\infty} f^*(t)e^{-j\omega t}\,dt \\ &= \left[\int_{-\infty}^{\infty} f(t)e^{j\omega t}\,dt\right]^* \\ &= F^*(-\omega)\end{aligned}$$

An important consequence of this property is that if the signal $f(t)$ is real-valued, then $f^*(t) = f(t)$ and $F^*(-\omega) = F(\omega)$.

3.6.3 Symmetry

Any signal can be expressed as a sum of an even function, $f_e(t)$, and an odd function, $f_o(t)$ [cf. Eqs. (2.62) and (2.63)]. This gives rise to the following Fourier transform properties:

$$\mathcal{F}\{f_e(t)\} = F_e(\omega) \qquad \text{(and real)}, \tag{3.39}$$

$$\mathcal{F}\{f_o(t)\} = F_o(\omega) \qquad \text{(and imaginary)}. \tag{3.40}$$

Proof For the first part, we have

$$\begin{aligned}\mathcal{F}\{f_e(t)\} &= \int_{-\infty}^{\infty} f_e(t)e^{-j\omega t}\,dt \\ &= \int_{-\infty}^{\infty} f_e(t)\cos\omega t\,dt - j\int_{-\infty}^{\infty} f_e(t)\sin\omega t\,dt \\ &= 2\int_{0}^{\infty} f_e(t)\cos\omega t\,dt\,.\end{aligned}$$

Because $\cos\omega t = \cos[(-\omega)t]$, this expression is even in ω. Proof of the second part follows in a similar manner.

For a causal signal (i.e., one that is zero for $t < 0$) we must have very specific combinations of even and odd functions. It follows that the real and imaginary parts of the spectrum of a causal signal must be related. In other words, both the real and imaginary parts of the spectrum of a causal signal cannot be specified arbitrarily.

3.6.4 Duality

A duality exists between the time domain and the frequency domain. This is exhibited in the Fourier transform relations [cf. Eqs. (3.13) and (3.14)] and can be stated explicitly in the following manner.

If

$$\mathcal{F}\{f(t)\} = F(\omega),$$

then

$$\mathcal{F}\{F(t)\} = 2\pi f(-\omega). \tag{3.41}$$

A proof of Eq. (3.41) involves an interchange in t and ω in the Fourier transform integrals. Its use is shown in the following example.

EXAMPLE 3.6.1

If it is given that $\mathcal{F}\{\text{rect}\,(t)\} = \text{Sa}\,(\omega/2)$, determine $\mathcal{F}\{\text{Sa}\,(t/2)\}$.

Solution Let $F(\omega) = \text{Sa}\,(\omega/2)$; then $F(t) = \text{Sa}\,(t/2)$. Using Eq. (3.41), we can write

$$\begin{aligned}\mathcal{F}\{F(t)\} &= 2\pi\ \text{rect}\,(-\omega)\\ &= 2\pi\ \text{rect}\,(\omega).\end{aligned}$$

An illustration of this symmetry is shown in Fig. 3.5.

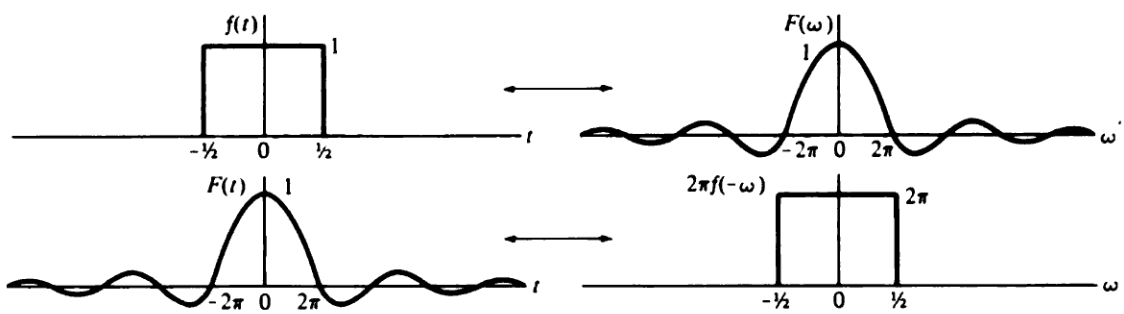


Figure 3.5 Duality of the Fourier transformation.

3.6.5 Coordinate Scaling (Reciprocal Spreading)

The expansion or compression of a time waveform affects the spectral density of the waveform. For a real-valued scaling constant α and any pulse signal $f(t)$,

$$\mathcal{F}\{f(\alpha t)\} = \frac{1}{|\alpha|}F\left(\frac{\omega}{\alpha}\right). \tag{3.42}$$

Proof

$$\mathcal{F}\{f(\alpha t)\} = \int_{-\infty}^{\infty} f(\alpha t)e^{-j\omega t}\,dt$$

We consider positive and negative values of α separately. For α positive, and changing the variable of integration to $x = \alpha t$, we have

$$\begin{aligned}\mathcal{F}\{f(\alpha t)\} &= \int_{-\infty}^{\infty} f(x)e^{-j\omega x/\alpha}\,dx/\alpha \\ &= \frac{1}{\alpha}F\left(\frac{\omega}{\alpha}\right) \quad \text{for } \alpha > 0\,.\end{aligned}$$

When α is negative, the limits on the integral are reversed when the variable of integration is changed so that

$$\mathcal{F}\{f(\alpha t)\} = -\frac{1}{\alpha}F\left(\frac{\omega}{\alpha}\right) \quad \text{for } \alpha < 0\,.$$

These two cases can be combined into the more compact form

$$\mathcal{F}\{f(\alpha t)\} = \frac{1}{|\alpha|}F\left(\frac{\omega}{\alpha}\right).$$

If α is positive and greater than unity, $f(\alpha t)$ is a compressed version of $f(t)$ and its spectral density is expanded in frequency by $1/\alpha$. The magnitude of the spectral density also changes—an effect necessary to maintain an energy balance between the two domains. If α is positive but less than unity, $f(\alpha t)$ is an expanded version of $f(t)$ and its spectral density is compressed. When α is negative, $f(\alpha t)$ is reversed in time compared to $f(t)$ and is expanded or compressed depending whether $|\alpha|$ is less than or greater than unity. (It may be helpful to note that the factor $1/|\alpha|$ is called the "magnification" in optics.) The effect of a time scale change on a pulse signal and its spectral density is shown in Fig. 3.6.

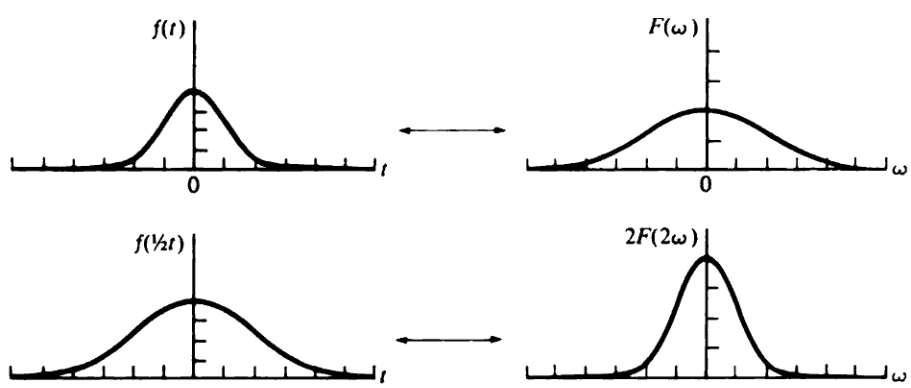


Figure 3.6 Effects of time scale changes on the signal spectral density.

DRILL PROBLEM 3.6.1

A given signal is recorded on magnetic tape at $7\frac{1}{2}$ inches per second (ips) and played back at $3\frac{3}{4}$ ips. If we assume a flat frequency response, what is the magnitude spectral density of the signal on playback if the original spectral density was of the form (a) $\exp(-|\omega|)$; (b) $1/\sqrt{\omega^2 + 1}$; (c) $\delta(\omega - \omega_0) + \delta(\omega + \omega_0)$?

ANSWER: (a) $2\exp(-2|\omega|)$; (b) $1/\sqrt{\omega^2 + 1/4}$; (c) $\delta(\omega - \omega_0/2) + \delta(\omega + \omega_0/2)$.

3.6.6 Time Shifting (Delay)

Another geometric operation is a translation of the time origin, causing the signal to be delayed (or advanced) in time by some time t_0. The corresponding effect on the signal spectral density is

$$\mathcal{F}\{f(t - t_0)\} = F(\omega)e^{-j\omega t_0}. \tag{3.43}$$

Proof

$$\mathcal{F}\{f(t - t_0)\} = \int_{-\infty}^{\infty} f(t - t_0)e^{-j\omega t}\,dt$$

Changing the variable of integration, let $x = t - t_0$,

$$\begin{aligned}\mathcal{F}\{f(t - t_0)\} &= \int_{-\infty}^{\infty} f(x)e^{-j\omega(x + t_0)}\,dx \\ &= e^{-j\omega t_0}\int_{-\infty}^{\infty} f(x)e^{-j\omega x}\,dx \\ &= e^{-j\omega t_0}F(\omega).\end{aligned}$$

Thus if a signal $f(t)$ is delayed in time by t_0, its magnitude spectral density remains unchanged and a negative phase $(-\omega t_0)$ is added to each frequency component. If t_0 is negative, the time function is advanced in time and the phase spectral density added has a positive slope.

EXAMPLE 3.6.2

Determine the Fourier transform of the two rectangular pulses shown in Fig. 3.7.

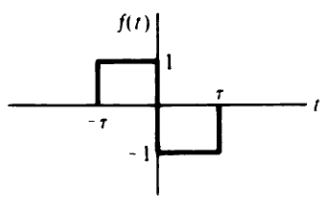


Figure 3.7 Rectangular pulses in Example 3.6.2.

Solution $f(t) = \text{rect}\,[(t + \tau/2)/\tau] - \text{rect}\,[(t - \tau/2)/\tau]$

Using the coordinate scaling property, we get

$$\mathcal{F}\{\text{rect}\,(t/\tau)\} = \tau\,\text{Sa}\,(\omega\tau/2)\,.$$

Using the delay property gives us

$$\mathcal{F}\{f(t)\} = \tau\,\text{Sa}\,(\omega\tau/2)[e^{j\omega\tau/2} - e^{-j\omega\tau/2}] = j(4/\omega)\sin^2(\omega\tau/2)\,.$$

DRILL PROBLEM 3.6.2

The spectral density of a given signal $f(t)$ is $F(\omega)$. The signal undergoes a distortion so that the new spectral density $F_1(\omega)$ is $F_1(\omega) = F(\omega)[1 + 2\alpha\cos\omega T]$, where α and T are parameters. Determine $f_1(t)$ in terms of $f(t)$.

ANSWER: $f_1(t) = f(t) + \alpha f(t - T) + \alpha f(t + T)$.

3.6.7 Frequency Shifting (Modulation)

The dual of the delay property is the *frequency-translation* property,

$$\mathcal{F}\{f(t)e^{j\omega_0 t}\} = F(\omega - \omega_0)\,. \tag{3.44}$$

Proof

$$\begin{aligned}\mathcal{F}\{f(t)e^{j\omega_0 t}\} &= \int_{-\infty}^{\infty} f(t)e^{j\omega_0 t}e^{-j\omega t}\,dt \\ &= \int_{-\infty}^{\infty} f(t)e^{-j(\omega-\omega_0)t}\,dt \\ &= F(\omega - \omega_0)\end{aligned}$$

Therefore multiplying a time function by $e^{j\omega_0 t}$ causes its spectral density to be translated in frequency by ω_0 rad/sec.

DRILL PROBLEM 3.6.3

Determine the complex-valued constant k required to make the following Fourier transform pair relation hold:

$$\mathcal{F}\{\exp\,(j\pi t)f(t - \tfrac{1}{2})\} = k\exp\,(-j\omega/2)F(\omega - \pi)\,.$$

ANSWER: $k = j$.

For real-valued $f(t)$, it is now a relatively simple matter to find the Fourier transform of $\mathcal{R}e\{f(t)e^{j\omega_0 t}\}$:

$$f(t)\cos\omega_0 t = \tfrac{1}{2}f(t)[e^{j\omega_0 t} + e^{-j\omega_0 t}].$$

Using the frequency translation property, we get

$$\mathcal{F}\{f(t)\cos\omega_0 t\} = \tfrac{1}{2}[F(\omega+\omega_0) + F(\omega-\omega_0)]. \tag{3.45}$$

This process of multiplying a signal by a harmonic function (i.e., a sine or cosine) to translate the spectral density is known as *amplitude modulation*. Equation (3.45) is known as the *modulation property* of the Fourier transform. Note that in the process of modulation one-half of the spectral density is moved up in frequency and one-half is moved down in frequency.

EXAMPLE 3.6.3

Find the spectral density of the pulse waveform,

$$A\ \mathrm{rect}\,(t/\tau)\cos\omega_0 t.$$

Solution From Table 3.1,

$$\mathcal{F}\{A\ \mathrm{rect}\,(t/\tau)\} = A\tau\ \mathrm{Sa}\,(\omega\tau/2).$$

Use of the modulation property gives

$$\mathcal{F}\{A\ \mathrm{rect}\,(t/\tau)\cos\omega_0 t\} = \tfrac{1}{2}A\tau\{\mathrm{Sa}\,[(\omega+\omega_0)\tau/2] + \mathrm{Sa}\,[(\omega-\omega_0)\tau/2]\}.$$

The pulse and its spectral density are shown in Fig. 3.8. Note that the modulation property gives us a convenient way to handle sinusoids of finite length.

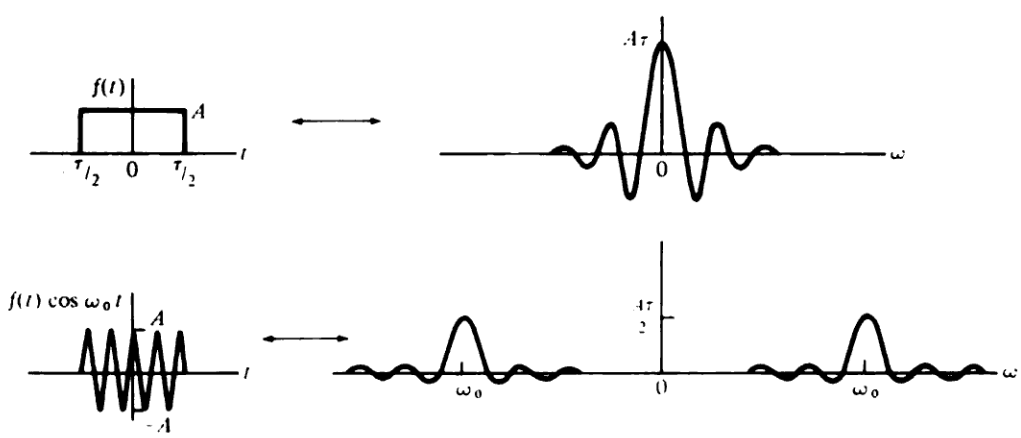


Figure 3.8 Effects of modulation on the signal spectral density.

DRILL PROBLEM 3.6.4

Based on the knowledge that

$$\mathcal{F}\{\exp(-at)u(t)\} = 1/(a + j\omega),$$

determine the Fourier transform of $\exp(-at)\cos\omega_0 t\, u(t)$ using the modulation property.

ANSWER: $(a + j\omega)/[(a^2 + \omega_0^2 - \omega^2) + j2a\omega]$.

3.6.8 Differentiation and Integration

If df/dt is absolutely integrable, then

$$\frac{d}{dt}f(t) \leftrightarrow j\omega F(\omega). \tag{3.46}$$

Proof Using the inverse Fourier transform, we have

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\,d\omega,$$

$$\begin{aligned}\frac{d}{dt}f(t) &= \frac{1}{2\pi}\frac{d}{dt}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\,d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\partial}{\partial t}[F(\omega)e^{j\omega t}]\,d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} j\omega F(\omega)e^{j\omega t}\,d\omega.\end{aligned}$$

Taking the Fourier transform of both sides,

$$\mathcal{F}\left\{\frac{d}{dt}f(t)\right\} = j\omega F(\omega).$$

Therefore time differentiation enhances the high-frequency components of a signal.

The corresponding integration property is

$$\mathcal{F}\left\{\int_{-\infty}^{t} f(\tau)\,d\tau\right\} = \frac{1}{j\omega}F(\omega) + \pi F(0)\,\delta(\omega), \tag{3.47}$$

where

$$F(0) = \int_{-\infty}^{\infty} f(t)\,dt. \tag{3.48}$$

Integration in time suppresses the high-frequency components of a signal. This conclusion agrees with the time-domain viewpoint that integration smooths out the time fluctuations in a signal.

If $F(0) \neq 0$, the proof of Eq. (3.47) must rely on a generalized function approach because the integrals may not converge. What is happening, of course, is that an ideal integrator senses any net positive or negative area in a pulse and holds it for an infinite time interval, thus creating a nonzero average value, as given in Eq. (3.48).

To illustrate the integration property, let us first define the *triangular function:*

$$\Lambda(t/\tau) = \begin{cases} 1 - |t|/\tau & |t| < \tau \\ 0 & |t| > \tau \end{cases}. \tag{3.49}$$

Equation (3.49) may be obtained by integrating the signal in Example 3.6.2 and dividing by the constant τ. Using Eq. (3.47) and the result of Example 3.6.2, we have

$$\begin{aligned} \mathcal{F}\{\Lambda(t/\tau)\} &= \frac{1}{j\omega}\frac{1}{\tau}[j(4/\omega)\sin^2(\omega\tau/2)] \\ &= \tau\,\mathrm{Sa}^2(\omega\tau/2). \end{aligned} \tag{3.50}$$

This useful transform pair is also illustrated in Fig. 3.3.

Differentiation in the time domain is equivalent to multiplication by the algebraic factor $(j\omega)$ in the frequency domain. On the other hand, integration in the time domain is equivalent to division by $(j\omega)$ in the frequency domain. A combination of these two properties often proves convenient in determining the Fourier transform of piecewise continuous functions. The procedure is illustrated below. [Note that Eq. (3.48) is always zero following a derivative operation because a derivative deletes any nonzero average value.]

EXAMPLE 3.6.4

Determine the Fourier transform of the trapezoidal pulse shown in Fig. 3.9(a).

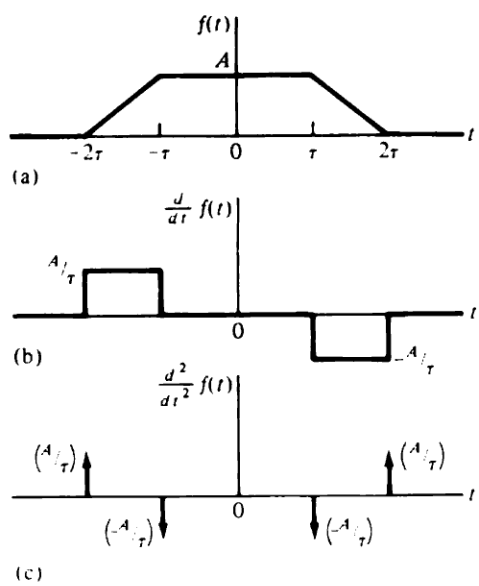


Figure 3.9 Trapezoidal pulse and its derivatives.

Solution Taking two successive derivatives, as shown in Figs. 3.9(b) and (c), we can write by inspection

$$(j\omega)^2 F(\omega) = \frac{A}{\tau}(e^{j2\omega\tau} - e^{j\omega\tau} - e^{-j\omega\tau} + e^{-j2\omega\tau}).$$

Simplifying, we get

$$(j\omega)^2 F(\omega) = \frac{A}{\tau}(e^{j\omega\tau/2} - e^{-j\omega\tau/2})^2(e^{j\omega\tau} + 1 + e^{-j\omega\tau}),$$

or

$$F(\omega) = A\tau\,\mathrm{Sa}^2\,(\omega\tau/2)[1 + 2\cos\omega\tau].$$

DRILL PROBLEM 3.6.5

Determine the Fourier transform of the triangular pulse

$$f(t) = 2t[u(t) - u(t-1)] - (t-3)[u(t-1) - u(t-3)].$$

ANSWER: $F(\omega) = [2 + \exp(-j3\omega) - 3\exp(-j\omega)]/(j\omega)^2$.

3.6.9 Time Convolution

As discussed in Chapter 2, one method of characterizing a system is by its frequency transfer function; a second method is by its impulse response. We now wish to relate these methods using the principle of convolution.

For the test signal $f(t) = \delta(t - \tau)$, the system impulse response is defined as [cf. Eq. (2.11)]

$$\mathcal{T}\{\delta(t-\tau)\} = h(t,\tau), \tag{3.51}$$

where τ is the delay or "age" variable. If the system is *time-invariant*, $h(t,\tau)$ takes the special form

$$\mathcal{T}\{\delta(t-\tau)\} = h(t-\tau). \tag{3.52}$$

The input signal $f(t)$ may be expressed in terms of impulse functions by [cf. Eqs. (2.95) and (2.98)]

$$f(t) = \int_{-\infty}^{\infty} f(\tau)\,\delta(\tau - t)\,d\tau = \int_{-\infty}^{\infty} f(\tau)\,\delta(t-\tau)\,d\tau.$$

If the system is linear, the impulse responses corresponding to each value of the age variable τ may be added up to form the output response,

$$g(t) = \mathcal{T}\left\{\int_{-\infty}^{\infty} f(\tau)\,\delta(t-\tau)\,d\tau\right\}.$$

From integration theory, we can rewrite this as

$$g(t) = \mathcal{T}\left\{\lim_{\Delta\tau\to 0}\sum_{n=-\infty}^{\infty} f(\tau_n)\,\delta(t-\tau_n)\,\Delta\tau\right\}.$$

Using the principle of superposition, we move the system operator inside the summation. Also, the $f(\tau_n)$ are the weights (areas) of the impulse functions and are constants for each impulse. Therefore we have

$$g(t) = \lim_{\Delta\tau\to 0}\sum_{n=-\infty}^{\infty} f(\tau_n)\mathcal{T}\{\delta(t-\tau_n)\}\,\Delta\tau.$$

Returning to the integral form and using Eq. (3.51), we have

$$g(t) = \int_{-\infty}^{\infty} f(\tau)h(t,\tau)\,d\tau. \tag{3.53}$$

This is a key result in signal analysis for it links the input to the output by means of an integral operation and holds for any linear system. In the particular case in which the system is also time-invariant, Eq. (3.53) reduces to the important form,

$$g(t) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)\,d\tau. \tag{3.54}$$

This result is known as the *convolution integral*. We find it convenient to use the following shorthand notation for the convolution integral:

$$\int_{-\infty}^{\infty} f(\tau)h(t-\tau)\,d\tau \triangleq f(t) \circledast h(t).$$

An important property of the Fourier transform is that it reduces the convolution integral operation to an algebraic product. This is stated mathematically in the following way. If

$$f(t) \circledast h(t) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)\,d\tau$$

and

$$\mathcal{F}\{f(t)\} = F(\omega), \qquad \mathcal{F}\{h(t)\} = H(\omega);$$

then

$$\mathcal{F}\{f(t) \circledast h(t)\} = F(\omega)H(\omega). \tag{3.55}$$

Proof

$$\mathcal{F}\{f(t) \circledast h(t)\} = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(\tau)h(t-\tau)\,d\tau\right]e^{-j\omega t}\,dt$$

Changing the order of integration and integrating with respect to t first yields

$$\mathcal{F}\{f(t) \circledast h(t)\} = \int_{-\infty}^{\infty} f(\tau)\left[\int_{-\infty}^{\infty} h(t-\tau)e^{-j\omega t}\,dt\right]d\tau\,.$$

Using the time-shift (delay) property, we have

$$\mathcal{F}\{h(t-\tau)\} = e^{-j\tau\omega}H(\omega)\,.$$

Then proceeding with the remaining integration over τ, we have

$$\begin{aligned}\mathcal{F}\{f(t) \circledast h(t)\} &= \int_{-\infty}^{\infty} f(\tau)H(\omega)e^{-j\omega\tau}\,d\tau \\ &= H(\omega)\int_{-\infty}^{\infty} f(\tau)e^{-j\omega\tau}\,d\tau \\ &= H(\omega)F(\omega)\,.\end{aligned}$$

Thus convolution in the time domain corresponds to multiplication in the frequency domain.

DRILL PROBLEM 3.6.6

Determine the Fourier transform of the signal $g(t) = \exp(-t)u(t) \circledast \exp(-2t)u(t)$ by (a) performing the convolution first and then taking the Fourier transform of the result; and (b) taking the Fourier transform of each term and then multiplying.

ANSWER: $1/(2 - \omega^2 + j3\omega)$.

3.6.10 Frequency Convolution

A dual to the preceding property is the following:

If

$$\mathcal{F}\{f_1(t)\} = F_1(\omega), \qquad \mathcal{F}\{f_2(t)\} = F_2(\omega)\,;$$

then[†]

$$\mathcal{F}\{f_1(t)f_2(t)\} = \frac{1}{2\pi}[F_1(\omega) \circledast F_2(\omega)]\,, \tag{3.56}$$

[†] The factor of $(1/2\pi)$ arises here because we are using radian frequency, ω. It is helpful to remember that frequency, f, and time, t, are reciprocals dimensionally. Radian frequency is 2π times reciprocal time, thus suggesting a division by 2π to keep the correct units.

where

$$F_1(\omega) \circledast F_2(\omega) = \int_{-\infty}^{\infty} F_1(u)F_2(\omega - u)\,du\,.$$

This property can be proved in the same way as that for the time convolution. Thus *the multiplication of two functions in the time domain is equivalent to the convolution of their spectral densities in the frequency domain.*

DRILL PROBLEM 3.6.7

Evaluate: rect (ω) exp $(-j\omega) \circledast$ exp $(-j\omega)$ at $\omega = \pi$.

ANSWER: -1.

TABLE 3.2 Some Fourier Transforms Corresponding to Given Mathematical Operations

Operation	$f(t)$	$\leftrightarrow$	$F(\omega)$
Linearity (superposition)	$a_1f_1(t) + a_2f_2(t)$		$a_1F_1(\omega) + a_2F_2(\omega)$
Complex conjugate	$f^*(t)$		$F^*(-\omega)$
Scaling	$f(\alpha t)$		$\frac{1}{\lvert\alpha\rvert}F\left(\frac{\omega}{\alpha}\right)$
Delay	$f(t - t_0)$		$e^{-j\omega t_0}F(\omega)$
Frequency translation	$e^{j\omega_0 t}f(t)$		$F(\omega - \omega_0)$
Amplitude modulation	$f(t)\cos\omega_0 t$		$\frac{1}{2}F(\omega + \omega_0) + \frac{1}{2}F(\omega - \omega_0)$
Time convolution	$\int_{-\infty}^{\infty} f_1(\tau)f_2(t - \tau)\,d\tau$		$F_1(\omega)F_2(\omega)$
Frequency convolution	$f_1(t)f_2(t)$		$\frac{1}{2\pi}\int_{-\infty}^{\infty} F_1(u)F_2(\omega - u)\,du$
Duality: time-frequency	$F(t)$		$2\pi f(-\omega)$
Symmetry: even-odd	$f_e(t)$		$F_e(\omega)$ [real]
	$f_o(t)$		$F_o(\omega)$ [imaginary]
Time differentiation	$\frac{d}{dt}f(t)$		$j\omega F(\omega)$
Time integration	$\int_{-\infty}^{t} f(\tau)\,d\tau$		$\frac{1}{j\omega}F(\omega) + \pi F(0)\,\delta(\omega)$, where $F(0) = \int_{-\infty}^{\infty} f(t)\,dt$

3.6.11 Remarks On Properties Of The Fourier Transform

The Fourier transform pair was obtained by a limiting operation on the complex exponential Fourier series. Assuming that an interchange of operations is valid, it follows that all of the properties discussed above for the Fourier transform also hold for the exponential Fourier series. The proofs are almost identical, the major change being the substitution of the discrete frequency variable $n\omega_0$ for the continuous variable ω.

A table of Fourier transform properties is given in Table 3.2. Note the symmetry between the time domain and the frequency domain.

Throughout this discussion we have been dealing with the Fourier transform pair, $f(t) \leftrightarrow F(\omega)$, where t represents time and ω represents radian frequency. The use of Fourier analysis, however, is not restricted to time-frequency variable pairs. In general, any pair of variables x, y may be used to form a Fourier transform pair $f(x) \leftrightarrow F(y)$ as long as their product (x, y) is dimensionless. The Fourier transform can also be extended to handle more than one dimension by choosing appropriate variable pairs in each dimension.

3.7 SOME CONVOLUTION RELATIONSHIPS

The convolution integral, as expressed in Eq. (3.54), holds as long as the system is linear and time-invariant. In addition, if the system is causal (i.e., physically realizable), then $h(t) = 0$ for all $t < 0$ and there is no contribution to the integration in Eq. (3.54) for $(t - \tau) < 0$:

$$g(t) = \int_{-\infty}^{t} f(\tau)h(t - \tau)\,d\tau; \qquad h(t) \text{ causal.} \tag{3.57}$$

Often it turns out that the input, $f(t)$, also satisfies the condition $f(t) = 0$ for $t < 0$, and Eq. (3.57) further simplifies to

$$g(t) = \int_{0}^{t} f(\tau)h(t - \tau)\,d\tau; \qquad f(t),\ h(t) \text{ causal.} \tag{3.58}$$

3.7.1 Properties Of Convolution

Some useful properties of convolution are listed below.

Commutative Law

$$f_1(t) \circledast f_2(t) = f_2(t) \circledast f_1(t) \tag{3.59}$$

Distributive Law

$$f_1(t) \circledast [f_2(t) + f_3(t)] = f_1(t) \circledast f_2(t) + f_1(t) \circledast f_3(t) \tag{3.60}$$

Associative Law

$$f_1(t) \circledast [f_2(t) \circledast f_3(t)] = [f_1(t) \circledast f_2(t)] \circledast f_3(t) \tag{3.61}$$

Proofs of these properties follow readily from the integral definitions and possible changes in the order of integration.

3.7.2 Convolution Involving Singularity Functions

The unit step response is the indefinite integral of the unit impulse response. This can be shown as follows:

$$u(t) \circledast h(t) = \int_{-\infty}^{\infty} u(\tau)h(t-\tau)\,d\tau = \int_{0}^{\infty} h(t-\tau)\,d\tau .$$

Changing the variable of integration, let $x = t - \tau$; then

$$u(t) \circledast h(t) = \int_{-\infty}^{t} h(x)\,dx . \tag{3.62}$$

This result gives a method for determining the impulse response of a system in the laboratory. Although technically the unit step function exists forever, most systems have an impulse response that lasts only for a relatively short duration of time. If we use a low-frequency square-wave generator whose repetition rate is much longer than the duration of the impulse response of the system, the system for all practical purposes sees a step. After recording the step response of the system, one can take the derivative of the output graph as a function of time. This will be the impulse response of the system.

Convolution with the unit impulse function follows using the integral properties of the impulse function:

$$\begin{aligned} f(t) \circledast \delta(t - t_0) &= \int_{-\infty}^{\infty} f(\tau)\,\delta(t - t_0 - \tau)\,d\tau \\ &= f(t - t_0) . \end{aligned} \tag{3.63}$$

Therefore the convolution of a function with the unit impulse reproduces that function exactly except that it is delayed (or advanced) by the delay (or advance) of the impulse.

EXAMPLE 3.7.1

Find $f(t) \circledast h(t)$ for the $f(t)$, $h(t)$ shown in Fig. 3.10(a), (b).

Solution From Fig. 3.10, $f(t) = A \sin \pi t\, u(t)$, $h(t) = \delta(t) - \delta(t-2)$,

$$g(t) = f(t) \circledast h(t) = \int_{-\infty}^{\infty} [A \sin \pi\tau\, u(\tau)][\delta(t-\tau) - \delta(t-2-\tau)]\, d\tau$$

$$= [A \sin \pi t]u(t) - [A \sin \pi(t-2)]u(t-2)$$

$$g(t) = \begin{cases} 0 & t < 0 \\ A \sin \pi t & 0 < t < 2 \\ 0 & t > 2 \end{cases}.$$

A sketch of $g(t)$ is shown in Fig. 3.10(c). This type of system is sometimes used in radar and navigation to determine the time-of-arrival of partially known waveforms. Note that the required $h(t)$ is that one used in Example 2.19.1 without the integrator.

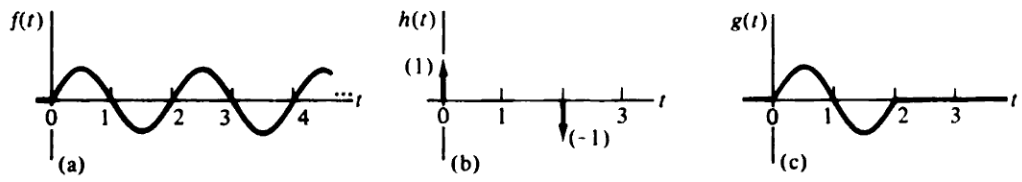


Figure 3.10 The delay-differencing system discussed in Example 3.7.1.

An extension to Eq. (3.63) is that the convolution of an impulse function with an impulse function results in another impulse function. Stated mathematically, we have

$$A\,\delta(t-t_0) \circledast B\,\delta(t-t_1) = AB \int_{-\infty}^{\infty} \delta(\tau - t_0)\,\delta(t - t_1 - \tau)\, d\tau$$

$$= AB\,\delta(t - t_0 - t_1). \tag{3.64}$$

A proof of this result follows from an application of the method of generalized functions (cf. Section 2.18); we shall not pursue this topic here.

3.8 GRAPHIC INTERPRETATION OF CONVOLUTION

The graphic interpretation of convolution permits one to grasp visually the results of the more abstract mathematical operations. Suppose we wish to find the convolution of two given functions $f_1(t)$ and $f_2(t)$. The operations to be performed are based on the convolution integral:

$$f_1(t) \circledast f_2(t) = \int_{-\infty}^{\infty} f_1(\tau) f_2(t-\tau)\, d\tau.$$

First of all we list the required operations, step by step:

1. Replace t by τ in $f_1(t)$, giving $f_1(\tau)$.
2. Replace t by $(-\tau)$ in $f_2(t)$. This folds the function $f_2(\tau)$ about the vertical axis passing through the origin of the τ axis.
3. Translate the entire frame of reference of $f_2(-\tau)$ by an amount t.† (As far as the integration is concerned, t is merely a parameter.) Thus the amount of translation, t, is the difference between the moving frame of reference and the fixed frame of reference. The origin in the moving frame is at $\tau = t$; the origin in the fixed frame is at $\tau = 0$. The function in the moving frame represents $f_2(t - \tau)$; the function in the fixed frame represents $f_1(\tau)$.
4. At any given relative shift between the frames of reference, for example, t_0, we must find the area under the product of the two functions, that is,
$$\int_{-\infty}^{\infty} f_1(\tau) f_2(t_0 - \tau)\, d\tau = [f_1(t) \circledast f_2(t)]_{t=t_0}.$$
5. This procedure is to be repeated for different values of $t = t_0$ by successively progressing the movable frame and finding the values of the convolution integral at those values of t. For continuous functions, this can be accomplished by a straightforward integration. For piecewise continuous functions, the product will be piecewise continuous and we must integrate over each continuous section.
6. If the amount of shift of the movable frame is along the negative τ axis (i.e., to the left), t is negative. If the shift is along the positive τ axis (i.e., to the right), t is positive.

EXAMPLE 3.8.1

Find the convolution of the rectangular pulse $f_1(t)$ and the triangular pulse $f_2(t)$ shown in Fig. 3.11.

Solution The various steps described above are shown in Fig. 3.11.

DRILL PROBLEM 3.8.1

Find the convolution rect $(t/\alpha) \circledast$ rect (t/α).

ANSWER: $\alpha\Lambda(t/\alpha)$.

† This translation is analogous to the relative motion between rigid frames of reference in relativistic mechanics. In the theory of relativity it is the velocity between coordinate systems that is important, whereas in convolution the important thing is the distance and direction between coordinate systems.

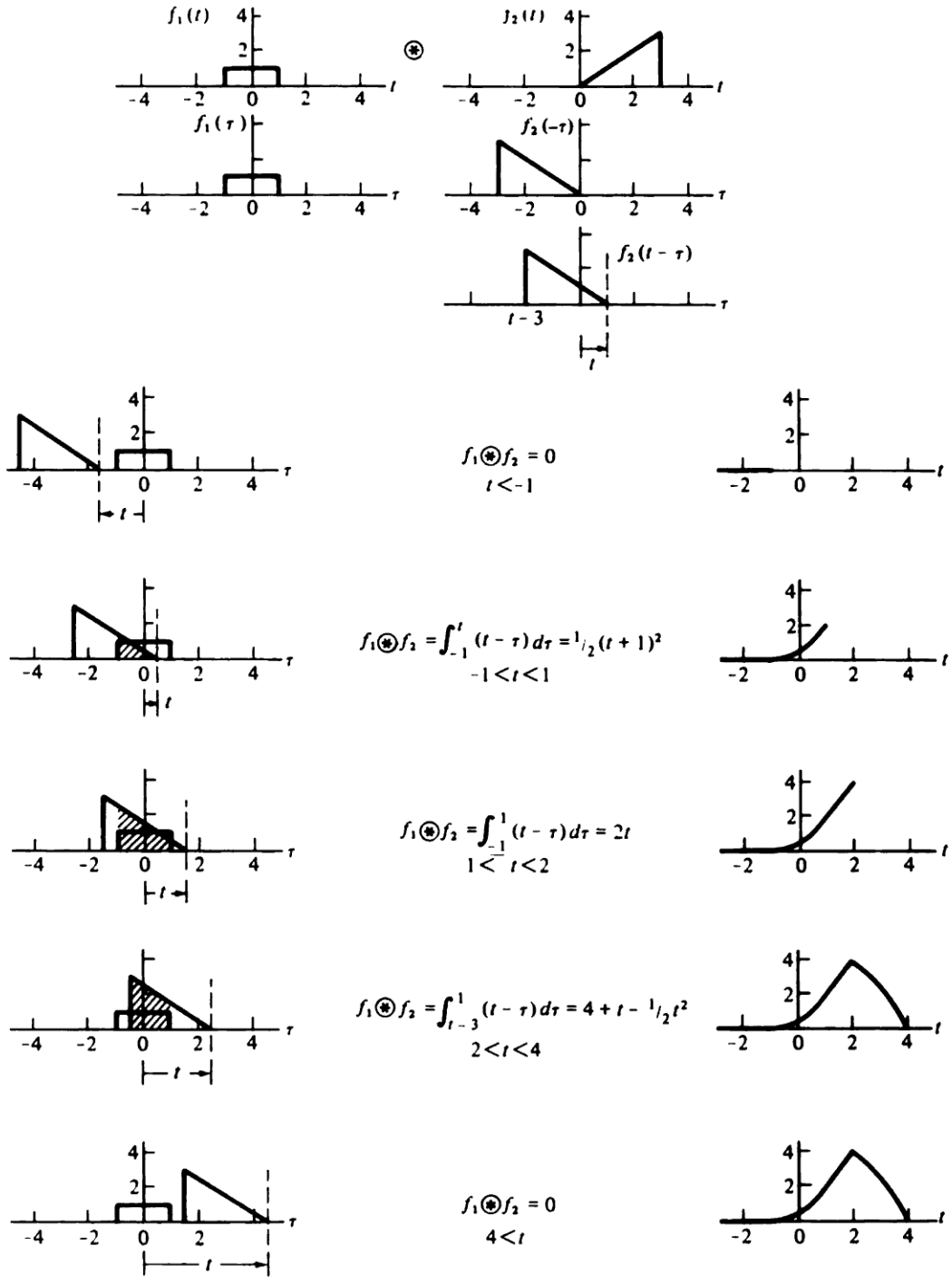


Figure 3.11 The convolution of a rectangular and a triangular pulse.

DRILL PROBLEM 3.8.2

A group of five narrow rectangular pulses, each of amplitude A and spaced by an interval T, is applied to the system of Drill Problem 2.19.1. Using convolution, determine the peak signal output. (This type of system is useful in the detection of periodic signals if only the period is known.)

ANSWER: $A(1 + K + K^2 + K^3 + K^4)$.

The computation procedure for the convolution integral is easily programmed on a digital computer. Thus if both $f(t)$ and $g(t)$ are zero for $t < 0$, we can write

$$f(t) \circledast g(t) = \int_0^t f(\tau)g(t - \tau)\,d\tau$$

$$\cong \alpha \sum_{m=1}^{n} f(m\alpha)g(n\alpha - m\alpha)\,,$$

where m, n are integers. Other cases follow in a similar manner.

3.9 FILTER CHARACTERISTICS OF LINEAR SYSTEMS

The application of an input signal $f(t)$ to a linear time-invariant system results in a corresponding output $g(t)$, as shown in Fig. 3.12. The spectral densities of the input and output are given by $F(\omega)$ and $G(\omega)$, respectively. The impulse response and the system frequency transfer function are designated by $h(t)$ and $H(\omega)$, respectively. The following relations hold:

$$g(t) = f(t) \circledast h(t)\,, \tag{3.65a}$$

$$G(\omega) = F(\omega)H(\omega)\,. \tag{3.65b}$$

Therefore the frequency transfer function is the Fourier transform of the impulse response of a linear time-invariant system. In the time domain, the system modifies the shape of the input signal. In frequency, the system modifies the spectral density of the input signal. These effects are related by the Fourier transformation.

A linear time-invariant system acts as a filter on the various frequency components applied to the system. Some frequency components may be amplified, some attenu-

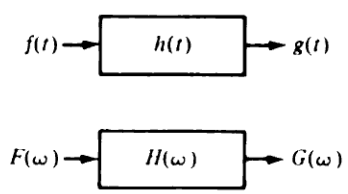


Figure 3.12 Representation of a system in time and frequency.

ated, and some remain unaffected. Each frequency component may be shifted in phase as it passes through the system. Rewriting Eq. (3.65b) to separate these two effects,

$$|G(\omega)|e^{j\theta_g(\omega)} = |F(\omega)|e^{j\theta_f(\omega)}|H(\omega)|e^{j\theta_h(\omega)},$$

$$|G(\omega)| = |F(\omega)|\,|H(\omega)|\,, \tag{3.66}$$

$$\theta_g(\omega) = \theta_f(\omega) + \theta_h(\omega)\,. \tag{3.67}$$

In other words, the magnitude response is given by the product of the magnitude of the signal spectral density and the magnitude of the system transfer function. The phase response is given by the sum of the individual phase responses. This can be extended to the case of several systems connected in cascade as long as each system does not alter the transfer function of any other system.

EXAMPLE 3.9.1

Determine the magnitude response of a RC low-pass filter to a gate function of unit amplitude and width τ; assume that $\tau = 4RC$.

Solution

$$|F(\omega)| = \tau\ \mathrm{Sa}\ (\omega\tau/2)\,,$$

$$|H(\omega)| = \left|\frac{1}{1 + j\omega RC}\right| = \frac{1}{\sqrt{1 + (\omega\tau/4)^2}}\,,$$

$$|G(\omega)| = |F(\omega)|\,|H(\omega)| = \frac{\tau}{\sqrt{1 + (\omega\tau/4)^2}}|\mathrm{Sa}\ (\omega\tau/2)|\,.$$

The magnitude plots are shown in Fig. 3.13. This system attenuates the high frequencies contained in the input spectral density and allows the lower frequencies to pass with relatively little attenuation. The unequal transmission of all frequency components results in a somewhat distorted replica of the input signal, as portrayed in Fig. 3.13(e).

DRILL PROBLEM 3.9.1

A given system delays the input signal by t_0 units of time and then subtracts the delayed version from the original signal. Calculate the frequency transfer function of this system.

ANSWER: $|H(\omega)| = |\sin(\omega\ t_0/2)|$; $\theta(\omega) = \tan^{-1}\,[(\sin\,\omega t_0)/(1 - \cos\,\omega t_0)]$.

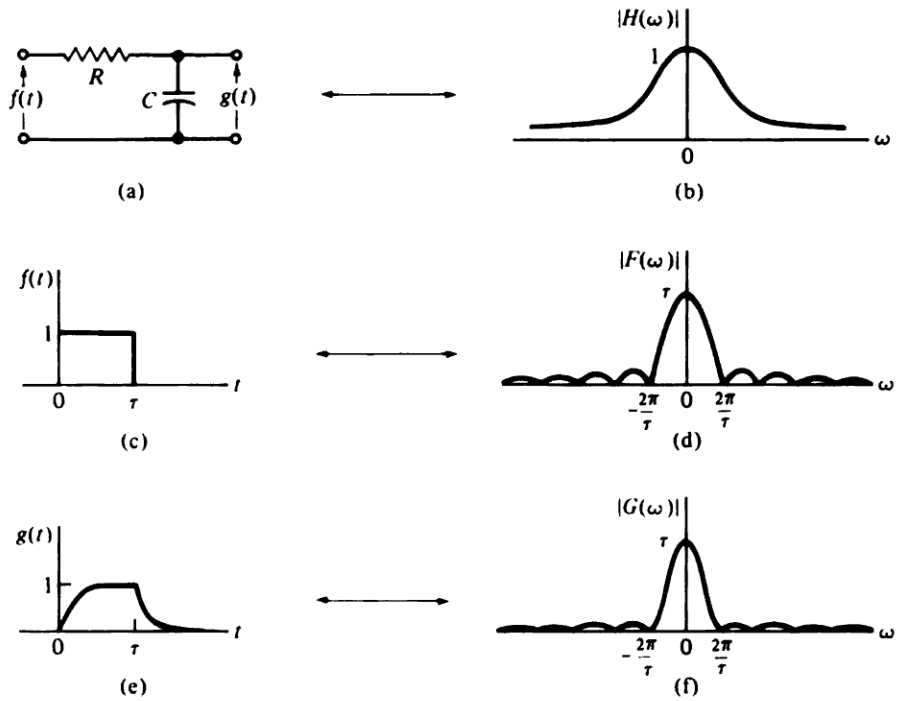


Figure 3.13 Magnitude response of a low-pass filter.

3.10 TRANSVERSAL FILTERS

The synthesis of a linear time-invariant system involves the design of a system to attain a given impulse response and frequency transfer function. One way in which this can be accomplished for the general case is by noting that the impulse response of a linear time-invariant system can be approximated by a delay line that is tapped at various points and weighted by a set of gain factors. To see this, we write the convolution integral for a causal system whose impulse response is $h(t)$ and whose signal input $f(t)$ is also causal:

$$g(t) = \int_0^t f(\tau)h(t-\tau)\,d\tau = \int_0^t f(t-\tau)h(\tau)\,d\tau. \tag{3.68}$$

The integration in Eq. (3.68) can be approximated by

$$g(t) \approx \sum_{k=0}^{t/\Delta\tau} f(t - k\,\Delta\tau)h(k\,\Delta\tau)\,\Delta\tau. \tag{3.69}$$

This result can be realized by using a delay line with taps at the delays $k\,\Delta\tau$; the output of each tap is multiplied by the preset weight $h(k\,\Delta\tau)\,\Delta\tau$. Such a system is shown in Fig. 3.14.

A filter constructed using a tapped delay line, tap weights, and adder in the configuration shown in Fig. 3.14 is called a *transversal* filter. The theoretical constraints

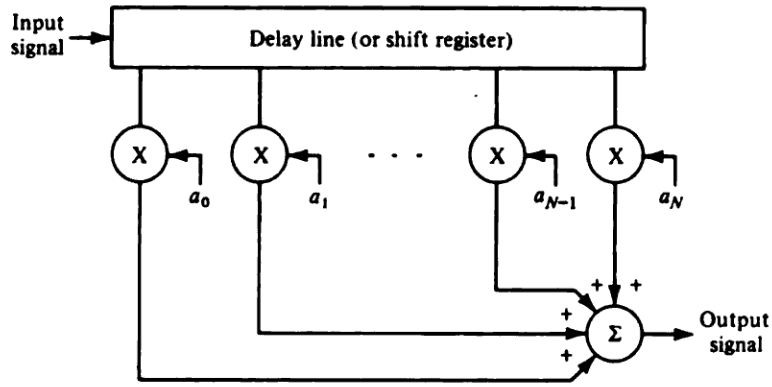


Figure 3.14 A tapped-delay-line (transversal) filter.

are more flexible on transversal filters than on lumped-RLC filters, and they are of increasing importance as new and more effective ways of implementing them are developed.

Both digital and analog methods can be used to implement transversal filters. In the digital implementation, delay is provided by a shift register, and the basic increment of delay is equal to the clock period. A computer can be used to implement the filter (in which case it is usually referred to as a digital filter). Another method used is the charge-coupled device (CCD), which eliminates the need for digitizing the signal.

The basic circuit realization of the CCD is a row of field-effect transistors (FET's) with drains and sources connected in series, and the drains capacitively coupled to the gates. This is illustrated in Fig. 3.15. Two clock lines are used to furnish alternating gate-control signals. The CCD samples the input signal and stores the sample as a charge on the first capacitor. This charge is then passed from stage to stage under the control of the clock pulses that turn on alternate FET's. This action transfers the charge to the next capacitor until eventually the charge reaches the end of the row of FET's. Taps at various delays may be weighted and combined to synthesize a transversal filter. CCD's operate at clock frequencies up to 10 MHz and can provide large delays (up to 1 sec). The CCD is sometimes referred to descriptively as a "bucket brigade" device.

A popular way to synthesize transversal filters using analog methods is that of the surface-acoustic-wave (SAW) filter. The essential components of a transversal SAW filter are two transducers attached to opposite ends of a polished piezoelectric

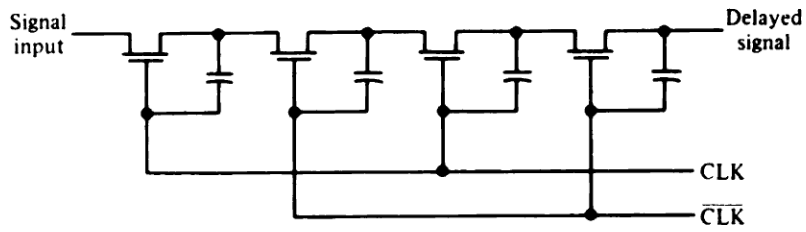


Figure 3.15 A charge-coupled-device (CCD) delay line.

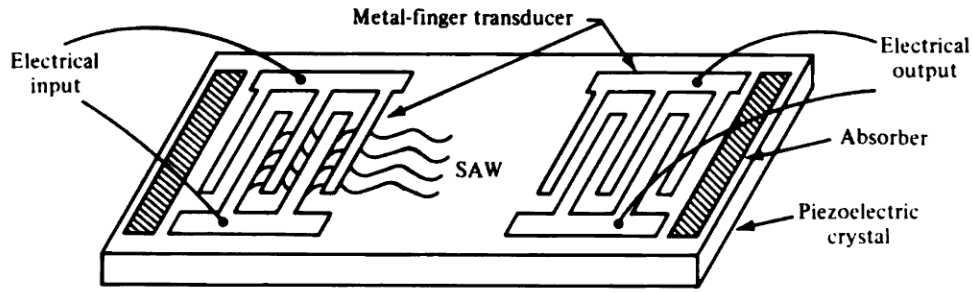


Figure 3.16 A surface-acoustic-wave (SAW) device.

substrate, as shown in Fig. 3.16. When an electric signal is applied to the input transducer, it physically distorts the piezoelectric surface, creating acoustic traveling waves. As these waves reach the output transducer, their mechanical energy is transformed back into an electrical signal. The substrate used is either quartz or lithium niobate, and the velocity of propagation on the substrate is about 3000 m/sec. Size of substrate, convenience of fabrication, and acoustic attenuation limit operating frequencies between 2 MHz and 2 GHz. In general, frequencies for standard production devices go from about 20 MHz to 500 MHz.

Each SAW transducer is a set of interleaved fingers of thin metallic film, as shown in Fig. 3.16. The purpose of the transducer is not only to convert signals from electric to acoustic energy but also to filter them. The impulse response of the transducer is directly linked to its geometry; relative finger positions determine the phase, and the overlap between adjacent fingers determines the amplitude weighting.[†] Thus the tapped delay line is formed by the finger spacings and the tap weights are determined by the overlap between adjacent fingers. The overall frequency transfer function of the SAW filter is given by the product of the Fourier transform of the impulse response of each transducer.

The SAW filters are used widely for high-quality custom filter applications. Once a SAW filter has been designed and constructed, of course, its characteristics cannot be adjusted. Research and development continues on other methods to build transversal filters.

3.11 BANDWIDTH OF A SYSTEM

The constancy of the magnitude $|H(\omega)|$ in a system is usually specified by a parameter called its *bandwidth*. The bandwidth W of a system is defined as the interval of *positive* frequencies over which the magnitude $|H(\omega)|$ remains within a given numeri-

[†] J. D. Maines and E. G. S. Paige, "Surface-Acoustic-Wave Devices for Signal Processing Applications," *Proceedings of the IEEE* [Special Issue on Surface Acoustic Wave Devices and Applications], 64 (May 1976): 639–671.

cal factor.† Although the exact criterion may vary, a popular numerical factor is −3 dB (that is, $1/\sqrt{2}$ in voltage or $\frac{1}{2}$ in power). Using this criterion, we refer to the bandwidth as the "−3 dB bandwidth" or the "half-power bandwidth" of the system.

According to this definition, the bandwidth of the system whose $|H(\omega)|$ is graphed in Fig. 3.17(a) is $W = \omega_1$ radians per second. The bandwidth of the system whose $|H(\omega)|$ plot is shown in Fig. 3.17(b) is $W = (\omega_2 - \omega_1)$ radians per second.

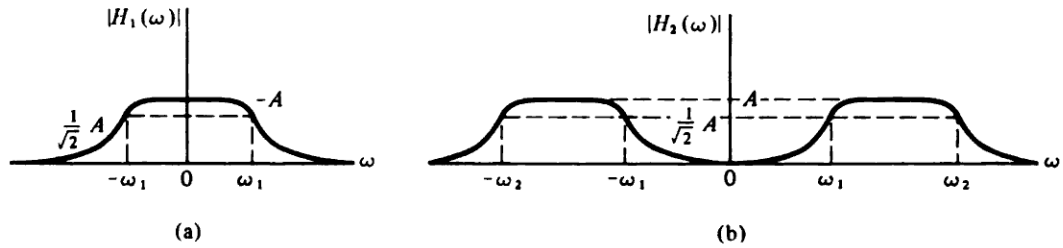


Figure 3.17 The bandwidth of a system as measured to its −3 dB points.

3.12 REQUIREMENTS FOR DISTORTIONLESS TRANSMISSION

What general requirements must be met for a linear time-invariant system to behave like an ideal transmission system? To answer this question, suppose that a given signal $f(t)$ is to be transmitted through such a system and we insist that the resulting output is to look just like the input. We allow this replica to have a different magnitude and some time delay as long as the signal shape is unaltered. Thus if $f(t)$ is the input signal, the required output is

$$g(t) = Kf(t - t_0). \tag{3.70}$$

Taking the Fourier transform of both sides of Eq. (3.70) and using the time-shift property, we get

$$G(\omega) = Ke^{-j\omega t_0}F(\omega). \tag{3.71}$$

Comparing Eqs. (3.65b) and (3.71), we see that the required system transfer function for distortionless transmission is

$$H(\omega) = Ke^{-j\omega t_0}. \tag{3.72}$$

Therefore, to achieve distortionless transmission, the overall system response must have a constant magnitude response and its phase shift must be linear with frequency. It is not enough that the system attenuate (or amplify) all frequency components

† The units of frequency can be expressed either in terms of radian frequency (radians per second) or in terms of cycles per second (Hz). As much as possible, we use W for radian frequency and B for frequency in Hz (hence $W = 2\pi B$). Use of the decibel (dB) is reviewed in Appendix B.

equally. All these frequency components must also arrive with identical time delay in order to add up correctly. This demands a phase shift that is proportional to frequency.† Note that a factor of $e^{j2\pi n}$, for any integer n, may be included in Eq. (3.71) without any net effect.

In practice, a signal may be distorted in passing through some parts of a system. Phase or amplitude correction (equalization) networks may be introduced elsewhere in the system to correct for this distortion. It is the overall input-output characteristic of the system which determines its behavior.

3.13 TIME RESPONSE OF FILTERS

As a result of physical limitations, one cannot build a system with an infinite bandwidth. Despite this difficulty, the ideas developed for distortionless transmission are quite helpful in an understanding of filters. As an approximation, we choose a truncated distortionless transmission system and call this an ideal filter. More specifically, an ideal filter passes without distortion all frequency components between its lower cutoff frequency, which we shall designate as ω_ℓ, and its upper cutoff frequency, ω_u. Outside this range, the ideal filter is assumed to have zero magnitude response. The frequency range defined by $\omega_\ell < |\omega| < \omega_u$ is called the "passband" of the filter. The effective width of the passband is specified by the bandwidth $W = (\omega_u - \omega_\ell)$ radians per second (the bandwidth in Hz is designated by B, where $W = 2\pi B$). The ideal filter is allowed a gain factor K and a time delay t_0 within its passband.

The ideal filter described above is called an ideal bandpass filter (BPF). An ideal low-pass filter (LPF) has $\omega_\ell = 0$ and $\omega_u = W$. The ideal high-pass filter (HPF) has $\omega_u \to \infty$. Sketches of the ideal low-pass and bandpass filters are shown in Fig. 3.18.

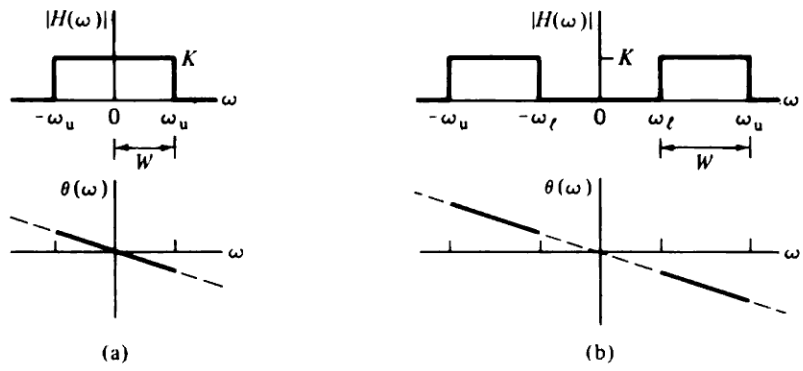


Figure 3.18 The ideal filter: (a) low-pass; (b) bandpass.

† For a fixed frequency, phase shift is proportional to time delay. Therefore the phase shift must vary linearly with frequency to yield a fixed time delay.

We now consider the ideal low-pass filter in some detail. Letting the attenuation constant K be unity for convenience, the transfer function of such a filter is

$$H(\omega) = |H(\omega)|e^{j\theta(\omega)} = \text{rect}\left(\frac{\omega}{2W}\right)e^{-j\omega t_0}. \tag{3.73}$$

The impulse response $h(t)$ of the ideal low-pass filter is (using transform pair number 15 in Table 3.1 and the time-shifting property)

$$h(t) = \mathcal{F}^{-1}\{H(\omega)\} = \frac{W}{\pi}\,\text{Sa}\,[W(t - t_0)]. \tag{3.74}$$

This impulse response is plotted in Fig. 3.19. From this plot, we observe that the width of the impulse response and the bandwidth are inversely related (as a consequence of the scaling property of the Fourier transform). Also note that the response of the LPF appears for $t < 0$, indicating that the ideal low-pass filter is not causal (physically realizable). Similar reasoning shows that the ideal bandpass filter is also noncausal.

Why are the ideal filters noncausal? They have finite bandwidths and linear phase characteristics. The source of the problem with the ideal filter is actually more the infinite loss outside the passband rather than the infinitely steep cutoff, although the latter is also impossible practically.† Of course, the very steep cutoff and the infinite loss outside the passband are reasons that the ideal filter is so desirable. Therefore we

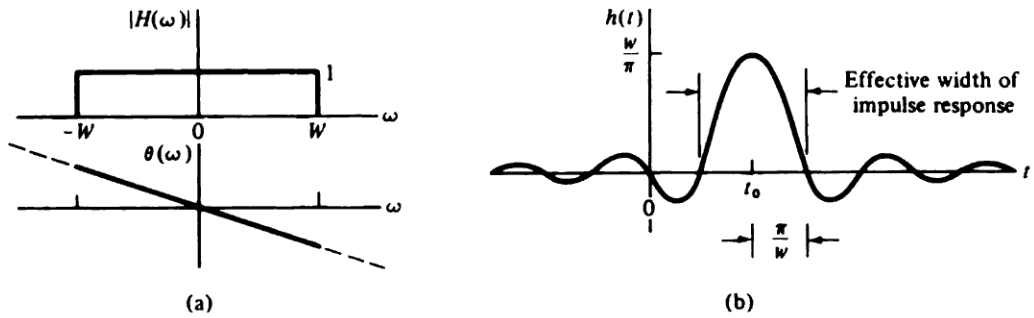


Figure 3.19 The ideal low-pass filter: (a) its system transfer function and (b) its impulse response.

† It can be shown that, for all $H(\omega)$ satisfying $\int_{-\infty}^{\infty}|H(\omega)|^2\,d\omega < \infty$, a necessary and sufficient condition on the magnitude spectrum for a filter to be physically realizable is that

$$\int_{-\infty}^{\infty}\left|\frac{\ln|H(\omega)|}{1+\omega^2}\right|d\omega < \infty.$$

This is called the Paley–Wiener criterion. From this result we can conclude that for a realizable filter the magnitude of $H(\omega)$ may not fall off toward zero faster than a function of simple exponential order (for example, $e^{-k|\omega|}$, but not as $e^{-k\omega^2}$) and the attenuation may not be infinite over any band of frequencies of nonzero width.

often attempt to build realizable filters to approximate the ideal filter characteristic as closely as possible. One measure—called the "shape factor"—of how well this can be done is the ratio of the filter bandwidths at −60 dB and −6 dB. A sharp-cutoff bandpass filter can be made with a shape factor as low as about 2; in contrast, the shape factor of a simple *RC* low-pass filter is almost 600!

The concepts of impulse function and impulse response are valuable analytically, but a more practical system test consists in applying a unit step to the input of a system and observing its response. Using the results obtained above, the step response follows using the convolution property:

$$\begin{aligned} g(t) &= h(t) \circledast u(t) \\ &= \int_{-\infty}^{\infty} h(\tau)u(t-\tau)\,d\tau \\ &= \int_{-\infty}^{t} h(\tau)\,d\tau . \end{aligned} \tag{3.75}$$

Therefore the step response of a linear time-invariant system is the indefinite integral of its impulse response. Combining Eqs. (3.74) and (3.75), we have

$$g(t) = \frac{W}{\pi}\int_{-\infty}^{t} \mathrm{Sa}\,[W(\tau - t_0)]\,d\tau .$$

Changing variable, let $x = W(\tau - t_0)$ so that

$$g(t) = \frac{1}{\pi}\int_{-\infty}^{W(t-t_0)} \mathrm{Sa}\,(x)\,dx .$$

Using the fact that $\int_{-\infty}^{0} \mathrm{Sa}\,(x)\,dx = \pi/2$, we have

$$g(t) = \frac{1}{2} + \frac{1}{\pi}\int_{0}^{W(t-t_0)} \mathrm{Sa}\,(x)\,dx . \tag{3.76}$$

The integral in Eq. (3.76) appears frequently in analytical problems and is tabulated in standard numerical tables or computed easily using numerical integration. It is called the "sine integral," denoted by Si (x), and defined by

$$\mathrm{Si}\,(x) \triangleq \int_{0}^{x} \mathrm{Sa}\,(u)\,du . \tag{3.77}$$

The sine integral has odd symmetry about $x = 0$, is linear for small x, and approaches $\pi/2$ for large x.

Using the sine integral notation, the step response of the ideal low-pass filter is

$$g(t) = \frac{1}{2} + \frac{1}{\pi}\,\mathrm{Si}\,[W(t - t_0)] . \tag{3.78}$$

The impulse response of the LPF and its step response are shown in Fig. 3.20.

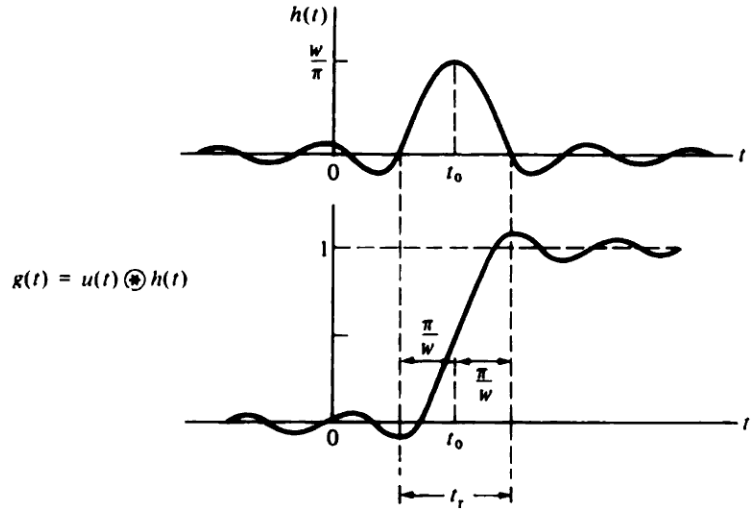


Figure 3.20 The unit step response of an ideal low-pass filter.

Several important observations can be made from Fig. 3.20. Note that as the bandwidth W of the LPF is decreased, the filter output $g(t)$ rises to its peak value more slowly. A measure of the time it takes for this rise is called the "rise time" of the filter. Therefore we conclude that the rise time, t_r, of a filter is inversely proportional to its bandwidth. An exact measure of the rise time, of course, will depend on which particular definition we choose. For pulse circuit applications, a convenient measure of t_r is the time required for the output waveform to rise from 10% to 90% of its final value. For the LPF discussed above, this yields $t_r \approx 0.44/B$, where $B = W/2\pi$. A convenient choice for the LPF is the time required for the output waveform to rise from its minimum to its maximum. From Fig. 3.20, this criterion leads to $t_r = 1/B$. Other criteria could be chosen, but the conclusion is always the same—rise time is inversely proportional to filter bandwidth.

A useful approximation for practical low-pass filters of order greater than second-order is that

$$t_r \approx 1/(2B), \tag{3.79}$$

where t_r is the 10% to 90% rise time in seconds and B is the -3 dB bandwidth in Hz. Because this is the type often used for filtering applications in communication systems, we shall use this approximation in later chapters.

Another important conclusion to be learned from Fig. 3.20 is that the rise time is a measure of an effective width of the system's impulse response. Recall that the impulse response is the derivative of the step response. The inverse relation between rise time and bandwidth is an example of the reciprocal spreading property of the Fourier transform. It is, therefore, a general property and not necessarily constrained to the example above.

Before leaving this discussion, we make one more rather general statement. Any strictly band-limited signal cannot be strictly time-limited. Conversely, a time-limited

signal cannot be strictly band-limited. We can see that this is true for the rectangular-type signals and systems we have been discussing. That this is true in general is beyond the level of this book.

The concept of effective width of an impulse response is tied directly to that of the response of a linear time-invariant system in the above discussion. However, the basic underlying ideas are those of Fourier transform theory and are not constrained to systems and system responses. We wish to extend these concepts also to an effective width for signals. We call this effective width the *time duration*, T_d. If we were to use the same definition for T_d that was used for rise time, it would be that time over which a given signal has the central 80% of its area under a graph of signal waveform versus time. Signal bandwidth B can be defined in the same manner as the -3 dB bandwidth defined earlier for systems. Other definitions could be used, however, and more general analytical definitions are presented in the following section.

An important dimensionless parameter in signal analysis is the *time (duration)-bandwidth product*, $T_d B$. One reason for this importance is that the time-bandwidth product has a lower bound, but it is not constrained by an upper bound. Therefore the time-bandwidth product may be increased intentionally for a given signal, and systems that are intended for this purpose are called "spread-spectrum" systems. Because bandwidth is not dependent on phase, a method to accomplish this spectral spreading is to make alterations to the signal phase to increase the time-bandwidth product. Another reason for this importance is that the time-bandwidth product is a useful parameter in specifying some signal processing systems, particularly those using delay devices and those intended for spread-spectrum applications.

DRILL PROBLEM 3.13.1

Using the 10% to 90% rise time criterion, compare the rise time and the -3 dB bandwidth of the RC low-pass filter (cf. Example 2.12.1).

ANSWER: $t_r = 0.35/B$.

3.14 ★ MINIMUM TIME-BANDWIDTH PRODUCT

Investigation of a minimum time-bandwidth product requires some new definitions of time duration and bandwidth that are easier to handle analytically. We redefine these quantities in terms of a "radius of gyration" definition, so-called because it is analogous to the definition of the moment of inertia about a center of mass used in mechanics. Using this approach, we define a time duration T_d of a given signal $f(t)$ as

$$T_d^2 = \frac{\int_{-\infty}^{\infty} (t - t_0)^2 |f(t)|^2 \, dt}{\int_{-\infty}^{\infty} |f(t)|^2 \, dt}, \tag{3.80}$$

where t_0 is the delay time of $f(t)$. In a similar manner, the "radius-of-gyration" bandwidth W can be found from

$$W^2 = \frac{\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |F(\omega)|^2 \, d\omega}{\int_{-\infty}^{\infty} |F(\omega)|^2 \, d\omega}, \tag{3.81}$$

where ω_0 is the center frequency. Explicit formulas for t_0 and ω_0 can be obtained by minimizing Eqs. (3.80), (3.81). However, because $|F(\omega)|^2$ is independent of time delay and $|f(t)|^2$ is independent of center frequency, we have deleted both t_0 and ω_0 in the following discussion.

Using Eqs. (3.80) and (3.81), and setting t_0 and ω_0 equal to zero, the time-bandwidth product (squared) is

$$(T_d W)^2 = \frac{\int_{-\infty}^{\infty} t^2 |f(t)|^2 \, dt \int_{-\infty}^{\infty} \omega^2 |F(\omega)|^2 \, d\omega}{\int_{-\infty}^{\infty} |f(t)|^2 \, dt \int_{-\infty}^{\infty} |F(\omega)|^2 \, d\omega}. \tag{3.82}$$

Using Parseval's theorem [cf. Eq. (3.21)], and using a dot above the function to indicate a derivative, we have

$$\int_{-\infty}^{\infty} |\dot{f}(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 |F(\omega)|^2 \, d\omega,$$

and Eq. (3.82) becomes

$$(T_d W)^2 = \frac{\int_{-\infty}^{\infty} t^2 |f(t)|^2 \, dt \int_{-\infty}^{\infty} |\dot{f}(t)|^2 \, dt}{\left[\int_{-\infty}^{\infty} |f(t)|^2 \, dt\right]^2}. \tag{3.83}$$

Fixing the signal energy, the term in the denominator of Eq. (3.83) is fixed. Using the Schwarz inequality (cf. Ex. 2.5.2) in the numerator, we have

$$(T_d W)^2 \geq \frac{\left|\int_{-\infty}^{\infty} t f(t) \dot{f}(t)^* \, dt\right|^2}{\left[\int_{-\infty}^{\infty} |f(t)|^2 \, dt\right]^2}. \tag{3.84}$$

Using integration by parts and insisting that $f(t) \to 0$ as $t \to \pm\infty$, the numerator in Eq. (3.84) is

$$\left[\frac{1}{2} \int_{-\infty}^{\infty} |f(t)|^2 \, dt\right]^2,$$

and Eq. (3.84) becomes

$$(T_d W)^2 \geq \tfrac{1}{4}.$$

Thus the minimum time-bandwidth product for the above definitions of T_d, W is given by

$$T_d W \geq \tfrac{1}{2}. \tag{3.85}$$

EXAMPLE 3.14.1

Find the time-bandwidth product of the gaussian-pulse signal: $f(t) = \exp[-t^2/(2\sigma^2)]$.

Solution Using Eqs. (3.80), (3.81), Appendix A, and setting t_0, ω_0 equal to zero we have

$$T_d^2 = \frac{\int_0^\infty t^2 e^{-t^2/\sigma^2}\,dt}{\int_0^\infty e^{-t^2/\sigma^2}\,dt} = \frac{\sigma^2}{2},$$

and

$$W^2 = \frac{\int_0^\infty \omega^2 e^{-\sigma^2\omega^2}\,d\omega}{\int_0^\infty e^{-\sigma^2\omega^2}\,d\omega} = \frac{1}{2\sigma^2}.$$

Thus for this example we have

$$T_d W = \tfrac{1}{2},$$

and we conclude that the gaussian-pulse signal is one that exhibits a minimum time-bandwidth product.

3.15 THE SAMPLING THEOREM

Signals bearing information may be available to us either in analog form or in digital or discrete form. We would like to determine the necessary conditions which will allow us to change an analog signal to a discrete one, or vice versa, without loss of information. As a criterion of how well this can be done, we shall insist that we be able to reconstruct the original signal completely by filtering.

The link between the analog signal and the corresponding discrete signal is provided by what is known as the *sampling theorem*. The sampling theorem can be stated simply as follows:

> A real-valued band-limited signal having no spectral components above a frequency of B Hz is determined uniquely by its values at uniform intervals spaced no greater than $1/(2B)$ seconds apart.†

This statement is a sufficient condition such that an analog signal can be reconstructed completely from a set of uniformly spaced discrete samples in time.

The validity of the sampling theorem can be demonstrated using either the modulation property or the frequency convolution property of the Fourier transform. Consider a band-limited signal $f(t)$ having no spectral components above B Hz. We shall sample this signal using the periodic gate function. Each rectangular sampling pulse is of unit amplitude, τ seconds in width, and occurs at intervals of T seconds.

Denoting the sampled signal by $f_s(t)$ and the periodic gate function by $p_T(t)$, we can write

$$f_s(t) = f(t)p_T(t). \tag{3.86}$$

However, $p_T(t)$ is periodic and can be represented by a Fourier series

$$p_T(t) = \sum_{n=-\infty}^{\infty} P_n e^{jn\omega_0 t}, \tag{3.87}$$

where $\omega_0 = 2\pi/T$. Combining Eqs. (3.86) and (3.87), we have

$$f_s(t) = f(t) \sum_{n=-\infty}^{\infty} P_n e^{jn\omega_0 t}. \tag{3.88}$$

Taking the Fourier transform of both sides of Eq. (3.88) and interchanging the order, we have

$$\begin{aligned} \mathcal{F}\{f_s(t)\} &= \mathcal{F}\left\{f(t) \sum_{n=-\infty}^{\infty} P_n e^{jn\omega_0 t}\right\} \\ &= \sum_{n=-\infty}^{\infty} P_n \mathcal{F}\{f(t)e^{jn\omega_0 t}\}. \end{aligned} \tag{3.89}$$

Using the frequency translation property of the Fourier transform in Eq. (3.89), we now write the spectral density of $f_s(t)$, designated by $F_s(\omega)$, as

$$\begin{aligned} F_s(\omega) &= \sum_{n=-\infty}^{\infty} P_n F(\omega - n\omega_0) \\ &= P_0 F(\omega) + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} P_n F(\omega - n\omega_0). \end{aligned} \tag{3.90}$$

† This statement is known as the "uniform sampling theorem." Requirements on the uniform spacing can be relaxed somewhat, yet the overall conclusions remain the same. See, for example, Chapter 4 of Reference 7.

Therefore we conclude that the spectral density (Fourier transform) of the sampled signal $f_s(t)$ is, within a constant factor, exactly the same as that of $f(t)$ within the original bandwidth. In addition, it repeats itself periodically in frequency every ω_0 radians per second. These replicas of the original spectral density are weighted by the amplitudes of the Fourier series coefficients of the sampling waveform, as shown in Fig. 3.21. Note that the spectral density of the original function $f(t)$ can be retrieved simply by using a LPF on $F_s(\omega)$.

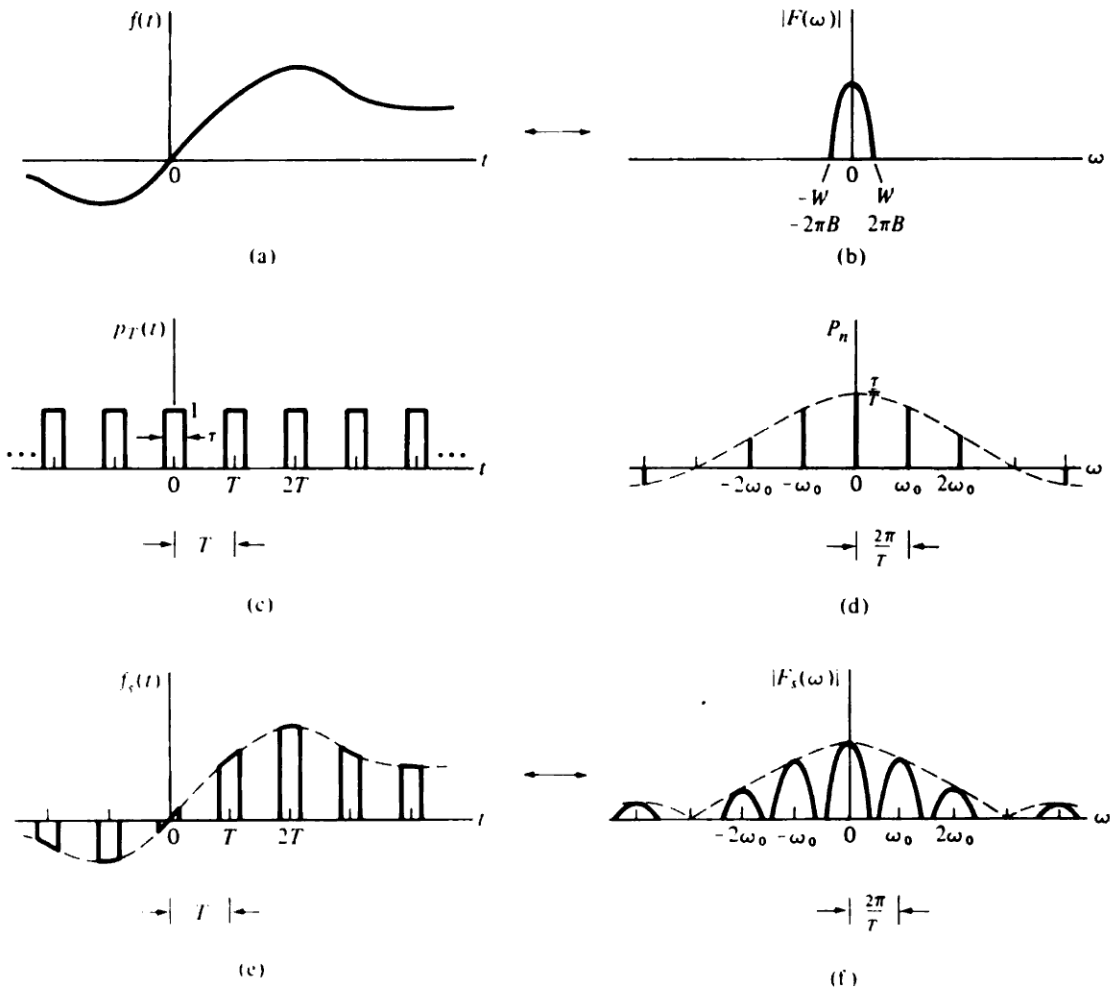


Figure 3.21 Steps in the sampling of a band-limited signal.

Now let us investigate a change in the sampling rate. For an increase in the sampling rate, ω_0 increases, T decreases, and all replicas of $F(\omega)$ move farther apart. On the other hand, as we decrease the sampling rate, ω_0 decreases, T increases, and all replicas move closer. Soon a point is reached beyond which a reduction in sampling

rate will result in overlap between spectral densities. From Fig. 3.21(f) it can be seen that this point is attained when

$$\frac{2\pi}{T} = 2W,$$

or, since $W = 2\pi B$,

$$T = \frac{1}{2B}. \tag{3.91}$$

At the point given by Eq. (3.91), all replicas of the original spectral density are just tangent to each other and an ideal low-pass filter can be used to retrieve the original from the sampled signal. However, if the sampling interval T becomes slightly larger than that given in Eq. (3.91), there is an overlap of spectral densities and the original signal cannot be retrieved by filtering the sampled signal. To avoid this condition, we must insist that

$$T < \frac{1}{2B}. \tag{3.92}$$

Equation (3.92) is a concise mathematical statement of the sampling theorem. The maximum time interval T of sampling, given by Eq. (3.91), is called the *Nyquist interval;* its reciprocal is called the *Nyquist sampling frequency (rate).*

In practice, the full potential of the sampling theorem usually cannot be realized and the statements above serve as upper bounds on actual performance. The first compromise we are forced to face is that we cannot build ideal low-pass filters. If the filter characteristic has a finite slope at the band edges, frequency components from the spectral replicas may be transmitted through the filter even though Eq. (3.92) is satisfied. These spurious frequencies can be suppressed by designing filters with as fast an attenuation rate as possible and then increasing the sampling frequency to allow some frequency space before the next replica appears. For this reason, practical systems oversample.

A second reason that the full potential of the sampling theorem cannot be realized arises from the fact that a time-limited signal is never strictly band-limited. When such a signal is sampled, there will be some unavoidable overlap of spectral components. In reconstruction of the signal, frequency components originally located above one-half the sampling frequency will appear below this point and will be transmitted by the low-pass filter. This is known as *aliasing* and results in distortion of the signal. The effects of aliasing can be combated by doing as good a job of filtering (i.e., band-limiting) as possible *before* sampling and by sampling at rates greater than the Nyquist rate. This is discussed in the next section.

Up to this point in our discussion of sampling we have assumed that all the signals under consideration were low-pass signals. Many signals of importance, however, are bandpass signals. Does the sampling theorem tell us that we must sample a bandpass signal at twice the highest frequency?

The answer to the question is no because the minimum sampling rate is dependent primarily on the bandwidth of a band-limited signal rather than on its highest frequency. For the low-pass case, of course, these two conditions coincide. When sampling a bandpass signal, we can position as many spectral replicas as possible between the original spectral densities, thus making more efficient use of spectral space. For uniform time sampling this results in a minimum sampling rate requirement between two and four times the bandwidth. For a given bandwidth, this minimum rate requirement for a bandpass signal approaches the limit of twice the bandwidth as the center frequency of the signal increases.† This is illustrated in Fig. 3.22. We shall assume that, unless specified otherwise, the minimum sampling rate for a bandpass signal is twice the bandwidth of the signal.

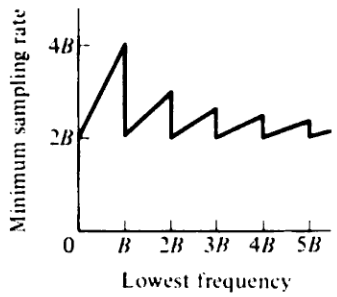


Figure 3.22 Minimum periodic sampling rate for a signal of bandwidth of B Hz.

A reconsideration of Fig. 3.21 shows that for the case of a complex-valued signal with a *one-sided* spectral density it is necessary to sample at a minimum rate equal to the highest frequency, not twice the highest frequency as for real-valued signals [and as given in Eq. (3.92)]. The basic principles illustrated in Fig. 3.21 remain the same, but one cannot jump to a quick conclusion based on Eq. (3.92). In reality, of course, the total number of samples taken remains the same for both cases because each sample of a complex-valued signal is an ordered pair (that is, a real part and an imaginary part). Sampling for the case of a one-sided spectral density is illustrated in the following example.

EXAMPLE 3.15.1

It is possible to reconstruct a function from one of the spectral replicas after sampling. A common sampling device making use of this principle is the *stroboscope*. A stroboscope is a light source that is pulsed periodically. In a study of rotary motion, the repetition rate of the light is adjusted so that the rotary motion appears motionless. The eye, acting as a low-pass filter, reconstructs a low-frequency "alias" from the bandpass sampled signal.

† See, for example, W. W. Harman, *Principles of the Statistical Theory of Communication*, New York, McGraw-Hill, 1963, Chapter 2.

For example, suppose you are watching a horse-drawn wagon in a Western movie. The wagon is being chased by the bad guys *but the wheels do not appear to be rotating!* If each wheel is 4 feet in diameter and has 12 identical spokes, what is the minimum forward velocity of the wagon? The picture frame rate of the movie is 24 frames per second.

Solution The minimum required sampling frequency of the spokes is $\frac{24}{12} = 2$ Hz.

$$v = r\omega = (2)(4\pi) = 8\pi \text{ ft/sec} \approx 17 \text{ mph}.$$

Note that velocities slightly lower than this will cause the wheels to appear to rotate *backward*. The aliases here have a one-to-one correspondence with the phasor rotations discussed earlier in Chapter 2. (Sketch a spectrum of the sampled signal to convince yourself.)

DRILL PROBLEM 3.15.1

Beginning with Eq. (3.86), derive the sampling theorem using the frequency convolution property of the Fourier transform.

3.16 ★ ALIASING EFFECTS IN SAMPLING

We have seen how the sampling of data introduces spectral replicas about multiples of the sampling frequency, and that the frequency band for spectral measurements should be limited to one-half the sampling rate. Now we return to the subject of aliasing which was discussed briefly in Chapter 2.

When a continuous function $f(t)$ is represented by M uniformly spaced samples, replicas of the spectral density of $f(t)$ are present, each spaced at a multiple of M units in frequency. The comparative magnitude of each spectral replica is dependent on the sampling pulse shape and the value of the sampling rate. If the sampling pulse shapes are impulse functions, all spectral replicas have equal area or weight (cf. Problem 2.18.1). We shall assume impulse sampling here both for convenience and because it is appropriate to the application of digital techniques in computation. Following the sampling theorem, we restrict the measurement or computation of spectral content to the interval $0 - M/2$. However, the existence of frequency components in the spectral replicas may fall within the measurement interval, causing errors.

We investigate these effects by considering the spectral density of a certain function $f(t)$, as shown in Fig. 3.23(a). If the waveform $f(t)$ is sampled using M point or impulse samples, the resulting spectral density is periodic in frequency, repeating at intervals of M, as shown in Fig. 3.23(b). Our measurement range is restricted to

$M/2$, but it is seen that the measured frequency components include portions from the spectral replicas. This effect, known as aliasing, is shown by the shaded area in Fig. 3.23(b). (This effect is also known by the descriptive term *fold-over.*) Note that for a given spectral density the aliasing effects decrease as the spectral replicas move apart and as the spectral components decrease more rapidly for frequencies higher than $M/2$. The former can be controlled by the choice of sampling rate. The latter can be controlled through use of a filter to attenuate those frequency components outside the band of interest. Filters used for this purpose are called *pre-alias* filters.

The above concepts can be extended easily to the Fourier series. The exponential Fourier series coefficients are related directly to the Fourier transform [cf. Eq. (3.15)] and the continuous spectral density becomes a line spectrum, as shown in Fig. 3.23(c).

Now let us estimate the accuracy of the measured coefficients in the exponential Fourier series. Using Eqs. (2.87), (2.88), and (2.93) we can write

$$F_n = \frac{1}{T}\int_0^T f(t)e^{-jn\omega_0 t}\,dt,$$

$$F_n \cong \frac{1}{M}\sum_{m=1}^{M} f(mT/M)[\cos(2\pi mn/M) - j\sin(2\pi mn/M)]$$

$$= \frac{a_n}{2} - j\frac{b_n}{2}. \tag{3.93}$$

Designating the true values of the exponential Fourier series coefficients by the F'_n and referring to Fig. 3.23(c), we can write an expression for the measured Fourier coefficients within the frequency range $-M/2 \le n \le M/2$ as

$$F_n = F'_n + F'_{n-M} + F'_{n-2M} + F'_{n-3M} + \cdots + F'_{n+M} + F'_{n+2M} + F'_{n+3M}\cdots. \tag{3.94}$$

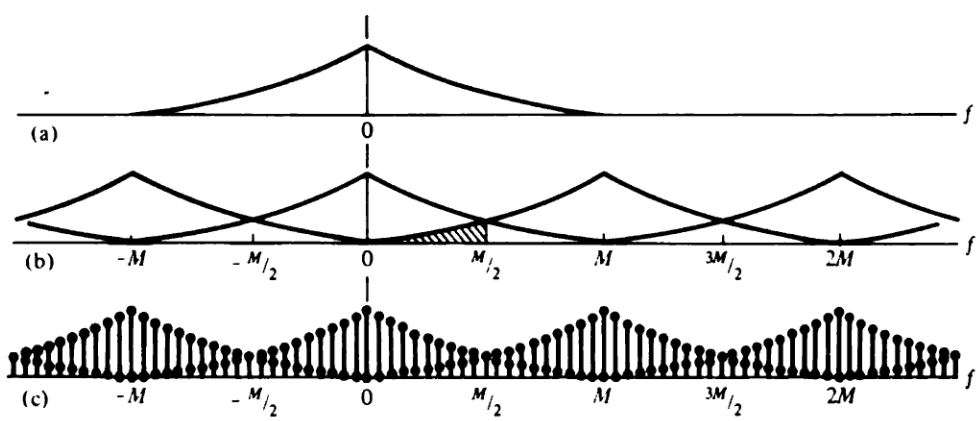


Figure 3.23 (a) A given spectral density; (b) aliasing effects for continuous spectra; and (c) line spectra.

However, for real-valued $f(t)$ we can also write [cf. Eq. (2.46)] $F_{-n} = F_n^*$ so that Eq. (3.94) becomes

$$F_n = F'_n + \sum_{k=1}^{\infty} [F'_{kM+n} + (F'_{kM-n})^*]. \tag{3.95}$$

Note that the alias terms have been grouped into the second term of Eq. (3.95).

Equation (3.95) gives the measured (e.g., computed) value of a given Fourier series coefficient in terms of the true value and the alias terms. By use of Eq. (3.93) in Eq. (3.95), a similar set of relations can be obtained for the aliasing effects for the trigonometric series [cf. Eqs. (2.91) and (2.92)].

The effects of aliasing on the reconstructed waveform representing $f(t)$ depend on the phase relationships as well as the magnitude and are difficult to predict exactly. To generalize the approach, we shall use an energy criterion that disregards the phase. To obtain a simplified measure of aliasing on an energy basis we assume (1) that only the first spectral replica gives rise to aliasing components, and (2) that the aliasing effects are proportional to the spectral components weighted by their frequency span from the origin (this emphasizes those effects near $M/2$). Stating this in equation form, let F_n be the exponential Fourier series coefficients at the input of a pre-alias filter whose magnitude transfer function is $|H(\omega)|$. Therefore $|G_n| = |F_n|\,|H(n\omega_0)|$ and we define the percent aliasing as

$$\% \text{ aliasing} \triangleq \frac{\sum_{n=1}^{M/2} (n - M/2)^2 |G_{n+M/2}|^2}{\sum_{n=1}^{M/2} n^2 |G_n|^2} \times 100\%. \tag{3.96}$$

While the results of this calculation disregard the phase relationships, they turn out to be quite reliable on a comparative basis and are easily computed. An application of this result is demonstrated in the following example.

EXAMPLE 3.16.1

A square wave is applied to a pre-alias filter whose magnitude transfer function is $|H(\omega)| = 1/\sqrt{1 + (\omega/\omega_c)^{2k}}$ and then sampled at a rate of 20 samples per period. (a) Compute the percent aliasing if the pre-alias filter is not used. (b) Compute the percent aliasing for $k = 1, 2, \ldots, 8$ and for ω_c, normalized to the half-sampling frequency, of 0.5, 0.6, 0.7, 0.8, 0.9, 1.0.

Solution a) Using Eq. (3.96) and $H(n) = 1$, we can write

$$\% \text{ aliasing} = \frac{\sum_{n=1}^{10} (n - 10)^2 H^2(n + 10)\, \mathrm{Sa}^2\, [(n + 10)\pi/2]}{\sum_{n=1}^{10} n^2 H^2(n)\, \mathrm{Sa}^2\, (n\pi/2)} = 22.0877\%.$$

b) Here a pre-alias filter (known as a kth order Butterworth filter) with a variable -3 dB frequency (ω_c) and a variable attenuation rate ($\approx -20\ k$ dB/decade) is used. The results are presented in tabular format below.

Percent Aliasing

k	$\frac{\omega_c}{\omega_s/2} =$	0.5	0.6	0.7	0.8	0.9	1.0
1		5.9830	7.2209	8.4080	9.5283	10.5723	11.5357
2		1.3154	2.2456	3.4790	4.9851	6.6896	8.4894
3		0.2676	0.6606	1.4020	2.6271	4.4042	6.6545
4		0.0538	0.1915	0.5584	1.3803	2.9390	5.3708
5		0.0108	0.0556	0.2221	0.7236	1.9719	4.4120
6		0.0022	0.0162	0.0885	0.3792	1.3251	3.6634
7		0.0004	0.0048	0.0354	0.1988	0.8902	3.0604
8		0.0001	0.0014	0.0142	0.1044	0.5976	2.5642

Note that the percent aliasing (or aliasing error) decreases with increasing attenuation rate in the pre-alias filter and as the -3 dB frequency is lowered. However, this also results in some undesirable attenuation within the desired band of frequencies. A solution is to sample at a higher rate, allowing some extra frequency band for the realizable filter characteristic.

√ DRILL PROBLEM 3.16.1

A triangular waveform is sampled at a rate of 20 samples per period. (a) Compute the percentage aliasing. (b) Repeat if a RC low-pass filter is used as a pre-alias filter with the -3 dB frequency set at the half-sampling frequency.

ANSWER: (a) 0.6637%; (b) 0.2903%.

In summary, the accurate measurement of frequency spectral components from sampled data requires a sampling frequency high enough that a pre-alias filter can be used to attenuate the higher frequency components and yet not attenuate those frequency components appreciably that are within the band of interest. All practical measurements using sampled data involve careful compromises between these factors.

3.17 ★ THE DISCRETE FOURIER TRANSFORM

The increasing use of digital methods for computational aids and for signal processing applications has resulted in an increased emphasis on a discrete version of the Fourier transform. We examine this topic briefly here with a particular interest in how the discrete Fourier transform is related to the continuous Fourier transform.

The notation conventionally used for the discrete Fourier transform differs somewhat from that used for the continuous Fourier transform. We follow conventional usage as much as possible and the reader should be aware that the notation used in this section differs slightly from that used in the other sections of this chapter.

Let a sequence of N equally spaced samples over the interval $(0, NT)$ be represented by

$$f(kT) = f(0), f(T), f(2T), \ldots, f[(N-1)T].$$

The *discrete Fourier transform* (DFT) is defined as that sequence of N complex-valued samples in the frequency domain given by[†]

$$F_D(n\Omega) = \sum_{k=0}^{N-1} f(kT)e^{-j\Omega Tnk}, \qquad n = 0, 1, \ldots, N-1, \tag{3.97}$$

where $\Omega = 2\pi/(NT)$. Note that $\Omega T = 2\pi/N$ and Ω and T do not appear explicitly in the DFT. These parameters enter only as scale factors for interpreting the results and are not needed in the computation steps.

In using a numerical approximation to the Fourier transform, it is necessary to restrict the observation interval to a finite length. Therefore let us define the truncated function $\tilde{f}(t)$ in terms of $f(t)$ by

$$\tilde{f}(t) = \begin{cases} f(t) & 0 \le t < NT \\ 0 & \text{elsewhere} \end{cases}. \tag{3.98}$$

The Fourier transform, $\tilde{F}(\omega)$, of this truncated function is

$$\tilde{F}(\omega) = \int_0^{NT} f(t)e^{-j\omega t}\,dt. \tag{3.99}$$

Making the variable changes $\omega \to n\Omega$, $t \to kT$, $dt \to T$, we can approximate Eq. (3.99) by

$$\tilde{F}(n\Omega) \cong \sum_{k=0}^{N-1} f(kT)e^{-jn\Omega kT}T. \tag{3.100}$$

[†] Some authors use definitions that vary slightly from this, and the summation can go from $k = 1$ to $k = N$. Note that T here is used as the interval between samples.

Equations (3.97) and (3.100) show that

$$\tilde{F}(\omega)\big|_{\omega=n\Omega} \cong TF_D(n\Omega)\,. \tag{3.101}$$

A comparison with the continuous Fourier transform shows that the two are analogous if (1) the signal $f(t)$ is truncated to the interval $(0, NT)$; (2) within this interval the signal $f(t)$ is available as a sequence of N equally spaced values; and (3) the interval is extended periodically yielding the discrete harmonic frequencies $n\Omega = 2\pi n/(NT)$. Note that the second condition implies that the computed frequency spectra are periodic with a period of $N\Omega$. To demonstrate this, consider the calculation of $F_D(n\Omega)$ for $n > rN$, r any integer, so that $n = rN + n_1$, $n_1 < N$. Using this in Eq. (3.97), we have

$$F_D[(rN + n_1)\Omega] = \sum_{k=0}^{N-1} f(kT)e^{-j\Omega T(rN+n_1)k}.$$

But we can write $e^{-j\Omega TNr} = e^{-j2\pi r} = 1$ so that

$$F_D[(rN + n_1)\Omega] = \sum_{k=0}^{N-1} f(kT)e^{-j\Omega Tn_1 k} = F_D(n_1\Omega)\,. \tag{3.102}$$

Therefore any value of $F_D(n\Omega)$ for $n > rN$ can be expressed in terms of a smaller argument $n_1\Omega$, where $n_1 = n$ modulo N.† In other words, $F_D(n\Omega)$ is periodic with a period of $N\Omega$.

As a result of the periodicity of $F_D(n\Omega)$, the accuracy of the computation of the continuous Fourier transform using the DFT is affected by aliasing. These effects are analogous to those encountered in our earlier discussion of the Fourier series (cf. Sections 2.17 and 3.16) and can be minimized by choosing a high sampling rate (i.e., small T).

The coefficients of the exponential Fourier series may be computed using the DFT and then multiplying by $1/N$ [cf. Eqs. (2.85), (2.88), and (2.94)]. The highest frequency component which can be determined corresponds to $n = N/2$, or $(N/2)\Omega = (2T)^{-1}$ Hz. This agrees with the sampling theorem.

DRILL PROBLEM 3.17.1

Derive Eq. (3.100) by multiplying $f(t)$ by $\sum_{k=-\infty}^{\infty}\delta(t/T - k)$ in Eq. (3.99) and letting $\omega \to n\Omega$. [Define the limits in Eq. (3.99) as 0_-, NT_-.]

In an analogy to the continuous case, the *inverse discrete Fourier transform* (IDFT) is

$$f(kT) = \frac{1}{N}\sum_{n=0}^{N-1} F_D(n\Omega)e^{j\Omega Tkn}. \tag{3.103}$$

† A number is written as modulo N by expressing the remainder after all integral multiples of N have been subtracted; thus 14 modulo 4 is 2, 14 modulo 5 is 4, etc.

The DFT and the IDFT form an exact transform pair; it is only in the comparisons with the continuous transform that differences may arise.

Properties of the DFT are analogous to those of the continuous Fourier transform within the restrictions given above. As an example, the IDFT of the product of the DFT's of two sequences is the convolution of the sequences. However, the resulting convolution is periodic.† In fact, because the IDFT is basically of the same form as the DFT, all functions having a DFT are automatically extended periodically with a period of NT. For a given truncated function, a convenient way in which the effects of this periodicity can be minimized is to add zeros as extra sample points to the sequence. These added zeros are called *augmenting zeros*. Augmenting zeros decrease the harmonic frequency spacings and the effects of aliasing for a given waveform at the expense of added computation time. Their use is illustrated in the following example.

EXAMPLE 3.17.1

Compute the Fourier transform of $f(t) = u(t) - u(t-1)$ using 4 samples over the interval $(0, 1)$.

Solution The sample sequence is $\{1, 1, 1, 1\}$ and the DFT is

$$F_D(n\Omega) = \sum_{k=0}^{3} f(kT)e^{-j2\pi nk/4} = \begin{cases} 4 & n = 0 \\ 0 & n \neq 0 \end{cases}.$$

Using Eq. (3.101), we have

$$F(\omega) \cong \begin{cases} 1 & n\Omega = 0 \\ 0 & n\Omega \neq 0 \end{cases} \qquad n = 0, 1, 2, 3.$$

However, expressing this function as $f(t) = \text{rect}\,(t - \frac{1}{2})$, the continuous Fourier transform is (cf. Tables 3.1 and 3.2)

$$F(\omega) = e^{-j\omega/2}\,\text{Sa}\,(\omega/2).$$

These two results certainly do not look alike. Let us convert the latter result using Eq. (3.101),

$$F_D(n\Omega) = 4e^{-jn\pi}\,\text{Sa}\,(n\pi) = \begin{Bmatrix} 4 & n = 0 \\ 0 & n \neq 0 \end{Bmatrix} \text{ over } (0, 4),$$

which confirms our earlier result.

The apparent inconsistency here lies in an incorrect interpretation of the truncation of $f(t)$. A periodic extension of $f(t)$ does not yield the desired waveform, but rather gives a constant level of one unit. This constant has only one frequency component located at $\omega = 0$—and that is exactly what we obtained! Figure 3.24(a) illustrates

† Periodic convolution is often called circular convolution; see, for example, Chapter 5 in Reference 1.

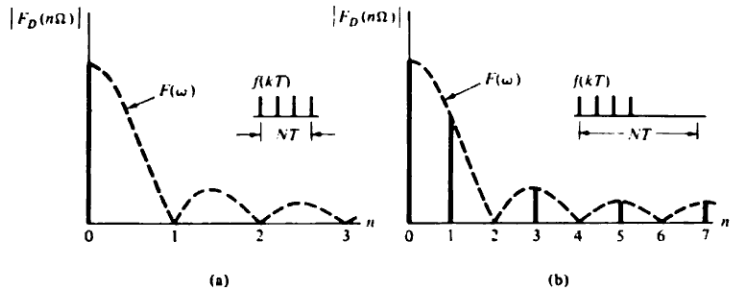


Figure 3.24 Sampled spectrum of a unit pulse: (a) without augmenting zeros; (b) with augmenting zeros.

what is happening by showing both $F(\omega)$ and $F_D(n\Omega)$. Note that the discrete frequency components are located at the nulls in the continuous spectrum for $\omega \neq 0$.

This situation can be improved by placing some augmenting zeros after the sequence of four samples. For example, if 4 zeros are used, the new sequence becomes $\{1, 1, 1, 1, 0, 0, 0, 0\}$. The DFT for this augmented sequence is shown in Fig. 3.24(b). The addition of these augmenting zeros has increased the number of frequency components. Note that a periodic extension of this augmented sequence yields a square wave. (Is this verified by the resulting discrete spectrum?)

DRILL PROBLEM 3.17.2

Find the DFT of the following sequences for $N = 4$: (a) $\{1, 0, 1, 0\}$; (b) $\{1, 1, 0, 0\}$.

ANSWER: (a) $\{2, 0, 2, 0\}$; (b) $\{2, 1 - j1, 0, 1 + j1\}$.

We have assumed that the data is sampled over the integers $k = 0, 1, 2, \ldots, (N - 1)$ [or, equivalently, $k = 1, 2, 3, \ldots, N$†]. When a continuous transform is approximated by the DFT, samples for negative values of time may be assigned to the interval $(N/2, N)$. The resulting symmetry relations are shown in Fig. 3.25. If augmenting zeros are used, they may be added symmetrically about $k = N/2$. Analogous symmetry relations hold in the discrete frequency domain with the added stipulation that the negative and positive frequency components are complex conjugates for real-valued $f(t)$.

Often we are interested in making as good an estimate of the frequency components of $F(\omega)$ as possible based on data in the interval $(0, NT)$. For functions of finite duration, it is possible to choose the interval for computation to correspond to (or be greater than) the duration of the function $f(t)$. This was the case in Example 3.17.1. Another case of interest is when the interval chosen for computation is only a portion

† As a result of periodicity by extension, the sample points $k = 0$ and $k = N$ are identical.

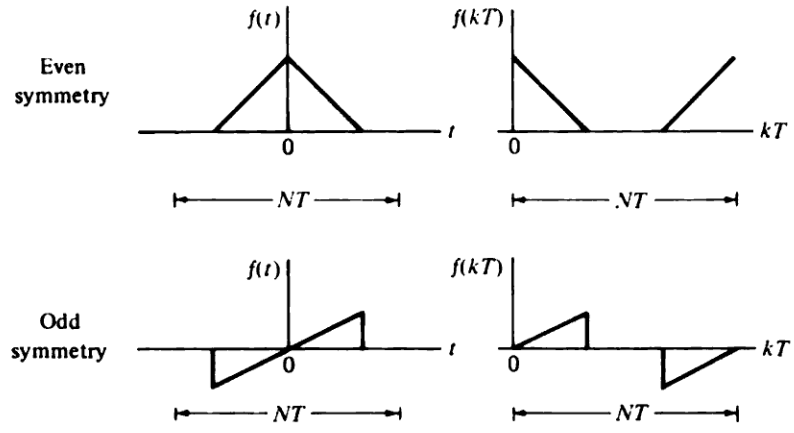


Figure 3.25 Symmetry relations for the discrete Fourier transform.

of the duration of $f(t)$. To investigate some of the effects of this truncation of $f(t)$, we rewrite Eq. (3.99) as

$$\tilde{F}(\omega) = \int_{-\infty}^{\infty} f(t) \text{ rect } [(t - NT/2)/(NT)]e^{-j\omega t}\, dt. \tag{3.104}$$

Written in this way, the truncation can be regarded as a window that permits us to observe only a finite interval of $f(t)$. Hence these functions, such as the second function in the integral in Eq. (3.104), are called *window functions*.

Equation (3.104) can also be expressed as a convolution in frequency:

$$\tilde{F}(\omega) = \frac{1}{2\pi}[F(\omega)] \circledast [NT \text{ Sa } (\omega NT/2)e^{-j\omega NT/2}]. \tag{3.105}$$

Ideally, the second bracketed term in Eq. (3.105) should be an impulse function to give a correct measure of $F(\omega)$. However, this requires that $(NT) \to \infty$ and is not a practical alternative.

For a finite record length NT, a quantity of interest is the minimum measurable frequency separation between frequency components. This minimum separation is called the *frequency resolution* in the estimation of $F(\omega)$ from $\tilde{F}(\omega)$. If two adjacent frequency components of interest have equal amplitudes, the frequency resolution is simply set by the record length NT:

$$\Delta\Omega = 2\pi/(NT). \tag{3.106}$$

If the two components have differing amplitudes, we also require that the Fourier transform of the window function must decrease rapidly for $\omega \neq 0$. The sin x/x pattern corresponding to the rectangular window is relatively poor in this respect because the main lobe (that is, near $\omega = 0$) has unit magnitude and the first sidelobe has a peak magnitude of 0.217 (that is, -13 dB). Therefore adjacent frequency components whose magnitudes differ by more than about 5 may be indistinguishable even though Eq. (3.106) is satisfied.

A remedy is to choose a window function that truncates $f(t)$ and whose Fourier transform has low sidelobes. This has been a topic of much research and no ideal solution exists. One rather simple and yet effective window function is the raised-cosine (Hann) window,

$$\begin{Bmatrix} \frac{1}{2}\left(1 + \cos\frac{2\pi t}{NT}\right) & |t| < NT/2 \\ 0 & |t| > NT/2 \end{Bmatrix}, \tag{3.107}$$

and is shown in Fig. 3.26. The Fourier transform of the raised-cosine window has lower sidelobes than that of the rectangular window at the expense of a wider main lobe and some attenuation. These compromises are typical, to differing extents, of the various window functions commonly employed in practice.

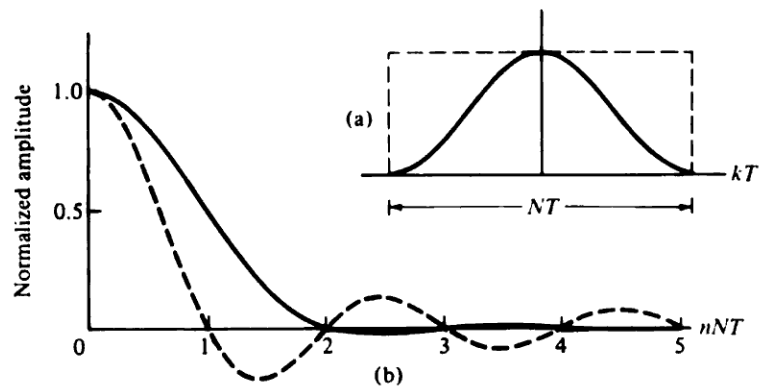


Figure 3.26 (a) The rectangular and raised-cosine window functions and (b) their Fourier transforms.

EXAMPLE 3.17.2

Compute the magnitude of the Fourier transform of $f(t) = \sin\sqrt{2}t + 0.1\sin 7\pi t$ over the interval $(0, 2)$ using the DFT with 32 sample points and (a) a rectangular window; and (b) a raised-cosine window.

Solution $\Omega = 2\pi/2 = \pi$, $T = \frac{2}{32} = \frac{1}{16}$; for the raised-cosine window (the term within the braces is 1 for the rectangular window):

$$F_D(n\Omega) = \frac{1}{2}\sum_{k=0}^{31} f(kT)\{1 + \cos[\pi(k - 16)/16]\}e^{-j\pi nk/16},$$

where $f(kT) = \sin(\sqrt{2}k/16) + 0.1\sin(7\pi k/16)$ and $F(\omega)$ is given by Eq. (3.101). Results of this computation are listed below and, plotted to a logarithmic scale, they are graphed in Fig. 3.27. Effects of lower sidelobes, as well as the wider main lobe, of the raised-cosine window are clearly evident.

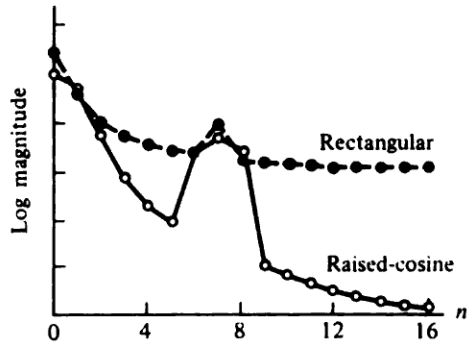


Figure 3.27 Results of the rectangular and raised-cosine window functions in Example 3.17.2.

Computed Magnitude of the Fourier Transform

n	Rectangular	Raised-cosine	n	Rectangular	Raised-cosine
0	1.36930	.86525			
1	.38146	.50423	9	.01622	.00009
2	.09842	.06028	10	.01499	.00006
3	.05336	.00734	11	.01408	.00005
4	.03679	.00211	12	.01340	.00003
5	.02842	.00086	13	.01291	.00002
6	.02343	.02476	14	.01258	.00002
7	.08971	.05015	15	.01239	.00001
8	.01787	.02490	16	.01232	.00001

DRILL PROBLEM 3.17.3

Compute the Fourier transform of $\Lambda(t/2)$ using the DFT and eight samples, each taken at 1 sec intervals.

ANSWER: $\{2, 1 + 1/\sqrt{2}, 1, 1 - 1/\sqrt{2}, 0, 1 - 1/\sqrt{2}, 1, 1 + 1/\sqrt{2}\}$.

3.18 ★ THE FAST FOURIER TRANSFORM

Computation of the DFT requires N^2 multiplications (that is, $0 \le k < N, 0 \le n < N$) and the resulting computation time becomes excessive when N becomes large. The key to more efficient computational methods is to make use of as much symmetry of the complex exponential as possible before the multiplications are performed.

Recent advances in this area have resulted in a class of efficient algorithms, known as the *fast Fourier transform* (FFT), which offer significant reductions in

computation time. The FFT is an algorithm (i.e., a systematic method of performing a series of computations in sequence) which enables the user to compute the DFT with a minimum of computation time. Aside from the algorithm itself, the interpretation of the FFT is the same as that for the DFT.

The particular algorithm introduced here is the commonly used Cooley–Tukey formulation.† This algorithm computes N discrete frequency components from N discrete time samples where $N = 2^r$, r any integer. This restriction to a power of two is not serious in practice as long as 2^r is greater than the number of data points—we can always fill the remainder with augmenting zeros.

As noted in the previous section, the parameters Ω and T are not really involved in the computation of the DFT. Therefore we adopt a shorthand notation by deleting them and Eq. (3.97) can be rewritten as

$$F_D(n) = \sum_{k=0}^{N-1} f(k)W^{nk}, \tag{3.108}$$

where

$$W^{\ell} = e^{-j2\pi\ell/N}, \qquad \ell = 0, 1, 2, \ldots. \tag{3.109}$$

Equation (3.109) describes a phasor of unit magnitude and a phase angle of $\theta_\ell = -2\pi\ell/N$. As an example, let $N = 2^3 = 8$; the corresponding values of Eq. (3.109) are plotted in Fig. 3.28. From this figure, it can be seen that W^{ℓ} and $W^{-\ell}$ are symmetrically located about the real axis and the following symmetry properties hold:

$$\begin{aligned} W^N &= W^0 = 1, \\ W^{N/2} &= -1 = -W^0, \\ W^{N-\ell} &= [W^{\ell}]^*. \end{aligned} \tag{3.110}$$

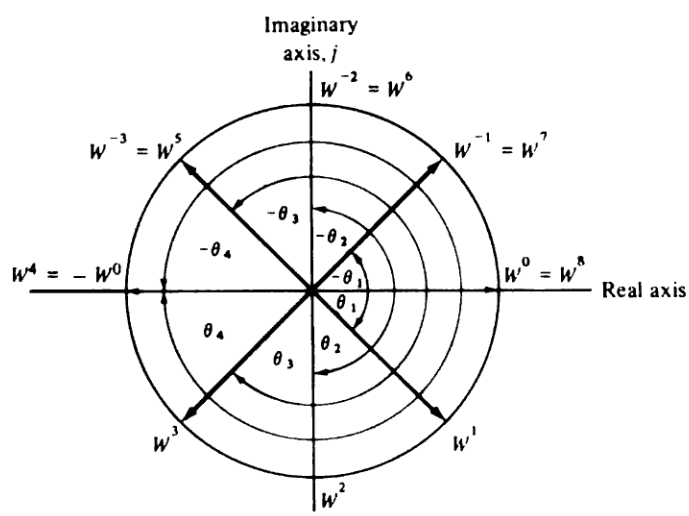


Figure 3.28 Powers of the exponential function W for $N = 8$.

† J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation* 19 (April 1965):297–301.

In order to utilize the full advantage of using $N = 2^r$, we express n, k as binary numbers. Suppose that $r = 2$ (that is, $N = 4$) and

$$k = (k_1, k_0) = \{00, 01, 10, 11\},$$
$$n = (n_1, n_0) = \{00, 01, 10, 11\}, \tag{3.111}$$

where n_0, n_1, k_0, k_1 can take on values of 0 and 1 only. A compact method of writing the numerical value of k and n is

$$k = 2k_1 + k_0,$$
$$n = 2n_1 + n_0. \tag{3.112}$$

Using Eqs. (3.111) and (3.112) in Eq. (3.108), we have

$$F_D(n_1, n_0) = \sum_{k_0=0}^{1} \sum_{k_1=0}^{1} f(k_1, k_0) W^{(2n_1+n_0)(2k_1+k_0)}, \tag{3.113}$$

where the double summation is now needed to enumerate the complete binary representation of k.

Consider the exponential term in Eq. (3.113),

$$W^{(2n_1+n_0)(2k_1+k_0)} = W^{(2n_1+n_0)2k_1} W^{(2n_1+n_0)k_0}$$
$$= W^{2n_0k_1} W^{(2n_1+n_0)k_0}, \tag{3.114}$$

because $W^{4n_1k_1} = 1$ for all integer n_1, k_1. This latter step is crucial to the efficiency of the FFT because now Eq. (3.113) can be rewritten as†

$$F_D(n_1, n_0) = \sum_{k_0=0}^{1} \left[\sum_{k_1=0}^{1} f(k_1, k_0) W^{2n_0k_1} \right] W^{(2n_1+n_0)k_0}. \tag{3.115}$$

The algorithm can now be exhibited by rewriting the summation in the brackets of Eq. (3.115) as

$$f_1(n_0, k_0) = \sum_{k_1=0}^{1} f(k_1, k_0) W^{2n_0k_1}, \tag{3.116}$$

and the outer summation as

$$f_2(n_0, n_1) = \sum_{k_0=0}^{1} f_1(n_0, k_0) W^{(2n_1+n_0)k_0}. \tag{3.117}$$

From Eqs. (3.115) and (3.117) we have

$$F_D(n_1, n_0) = f_2(n_0, n_1). \tag{3.118}$$

This last step is included because the ordering in the output is scrambled in this algorithm. A variant is to scramble the input ordering so that the output is in the correct ordering.

† For a matrix interpretation of this, see, for example Chapters 10 and 11 of Reference 5.

Equations (3.116)–(3.118) are the recursive relations that constitute the Cooley–Tukey algorithm for $N = 4$. A signal flow graph of these relations is shown in Fig. 3.29. Algorithms for higher powers of two follow in the same manner, with one summation in Eq. (3.113) for each power of two considered.

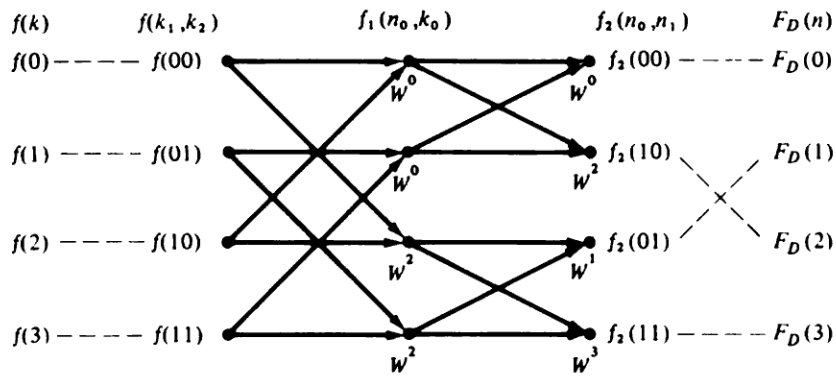


Figure 3.29 A signal flow graph of the Cooley–Tukey algorithm for $N = 4$.

About N^2 complex multiply-and-add operations are required in a direct evaluation of the DFT while the FFT algorithm requires on the order of $N \log_2 N$ operations. The net savings become appreciable for large N. For example, the computation time required for a 1024-point DFT with direct evaluation is almost 100 times greater than that required using the FFT algorithm. However, the FFT algorithm requires considerable computer storage and this may limit its application in situations where only a limited amount of storage is available.

Finally, the following list presents some points that prove helpful in processing continuous time functions with the FFT. Note that many of these arise from our earlier considerations of the DFT and the Fourier transform.

1. The number of samples N is chosen so that $N = 2^r$, r an integer. This number may include augmenting zeros [see (7) below].
2. For N time samples there are N discrete frequencies.
3. As a result of periodic extension, the sample points 0 and N are identical in both domains.
4. Positive frequency components are considered to be those over $(0, N/2)$; negative frequency components are those over $(N/2, N)$. This symmetry may also be used for time samples over negative and positive time.
5. For real-valued functions, the positive frequency components are the complex conjugates of the negative frequency components. The points $n = 0$, $N/2$ are common to both and are real-valued.
6. The highest frequency component (that is, $n = N/2$) is $(2T)^{-1}$ Hz; this may be increased by decreasing the spacing between time samples.

7. The spacing between frequency components is $(NT)^{-1}$ Hz; this may be decreased by adding augmenting zeros to the sample sequence.
8. The exact relationship of the FFT values depends on the particular multiplicative constants assigned in the algorithm; a fairly common procedure is to divide by N so that the computed values are $1/N$ times the DFT.

DRILL PROBLEM 3.18.1

(a) Repeat Example 3.17.2 using a FFT algorithm and check your results with those given. (b) Add 96 augmenting zeros, rerun your program, and check results. (c) Determine the frequency component spacings and the maximum frequency computed for parts (a) and (b).

ANSWER: (b) Every fourth coefficient should check with results given in Example 3.17.2. (c) 0.5000 Hz, 8 Hz; 0.125 Hz, 8 Hz.

3.19 SUMMARY

A nonperiodic signal $f(t)$ can be represented in terms of exponential functions over the entire real line $(-\infty, \infty)$. To be valid for all time, it must be represented by a continuous sum of exponential functions with frequencies lying in the interval $(-\infty, \infty)$. The Fourier transform is a method of expressing a given signal in terms of such a continuous set of exponential components of frequency. The resulting frequency function, $F(\omega)$, gives the relative weighting of each frequency component of the signal and is called the spectral-density function of the signal.

A signal of finite energy can always be described uniquely by a continuous spectral-density function. Periodic components (including dc levels) in the signal can be handled by introducing impulses in the spectral-density functions.

The absolute magnitude squared of the spectral-density function is an indicator of the relative energy per unit of frequency and is called the energy spectral-density function. The area under the energy spectral-density function is equal to the energy in the signal.

It is often convenient to use established properties of the Fourier transform pair without evaluating the transforms themselves. These properties are shown using the integral properties of the Fourier transform. One of the most useful properties is that of convolution. This property states that the Fourier transform reduces convolution operations to multiplications. A direct result of this property is that the system frequency transfer function is the Fourier transform of the impulse response of the system.

A linear time-invariant system acts as a filter on the various frequency components applied to the system. Transversal filters can be synthesized by using an addition of weighted tap outputs of a tapped delay line. Popular ways to construct

transversal filters include charge-coupled devices (CCD) and surface-acoustic-wave (SAW) delay lines.

The bandwidth of a system is defined to be the interval of positive frequencies over which the magnitude of the system transfer function remains constant within a given numerical factor. For a distortionless transmission, the overall response of a system must have a constant (i.e., flat) magnitude response and a linear phase response with frequency.

A convenient experimental test of a linear system is to apply a step function to the input. The output is the indefinite integral of the impulse response of the system. The rise time of the system is that time required for the output to effectively rise to the level of the final value after the step function has been applied. The rise time is a measure of the width of the system impulse response and is inversely proportional to the system bandwidth.

Time duration is a measure of the width of a pulse signal. Time duration multiplied by bandwidth is called the signal's time-bandwidth product. Although there is a lower bound, the time-bandwidth product of a signal may be increased.

Any band-limited signal is determined by its values at uniform intervals spaced no greater than $1/(2B)$ seconds apart. This statement is called the sampling theorem. The minimum sampling rate, $(2B)$ seconds^{-1}, is called the Nyquist sampling rate.

The sampling of data introduces spectral replicas about multiples of the sampling frequency. Frequency components in these spectral replicas may fall within the bandwidth restrictions set by the sampling theorem, causing errors known as aliasing. These errors can be decreased by using higher sampling rates and by using filters before sampling to band-limit the signal.

The discrete Fourier transform (DFT) can be used to represent a sampled signal in terms of a finite set of complex exponentials. The DFT may be used to approximate the Fourier series and the Fourier transform. The accuracy of this approximation is limited by truncation and aliasing effects.

The fast Fourier transform (FFT) is an efficient algorithm for the numerical computation of the DFT. The most commonly used version of the FFT computes N frequency components from N time samples for $N = 2^r$, r any positive integer.

Selected References for Further Reading

1. C. D. McGillem and G. R. Cooper. *Continuous and Discrete Signal and System Analysis*. Second ed. New York: Holt, Rinehart & Winston, 1984.
   A treatment paralleling much of this chapter with added emphasis on numerical convolution methods.
2. R. A. Gabel and R. A. Roberts. *Signals and Linear Systems*. Third ed. New York: John Wiley & Sons, 1987.
   Chapter 3 has a good treatment of convolution and impulse response.
3. R. E. Ziemer, W. H. Tranter, and D. R. Fanin. *Signals and Systems*. New York: Macmillan, 1983.
   This text has a good development and discussion of the FFT, its relation to the DFT, and applications.
4. A. V. Oppenheim, A. S. Willsky, and I. T. Young. *Signals and Systems*. Englewood Cliffs, N.J.: Prentice Hall, 1983.
   A good treatment of the Fourier transform in both continuous and discrete-time systems.

5. E. O. Brigham. *The Fast Fourier Transform*. Englewood Cliffs, N.J.: Prentice Hall, 1974.
A good and fairly complete discussion of the DFT and FFT as well as a concise treatment of the continuous Fourier transform.
6. R. Bracewell. *The Fourier Transform and Its Applications*. Second ed. New York: McGraw-Hill, 1978.
Although more advanced, this book is very readable and stresses physical intuition.
7. H. S. Black. *Modulation Theory*. Princeton, N.J.: D. Van Nostrand, 1953.
Presents an expanded discussion of sampling theory and applications; a classic in its field.
8. A. Papoulis. *The Fourier Integral and Its Applications*. New York: McGraw-Hill, 1962.
An advanced treatment of the Fourier transform and its application from a more mathematical perspective.

PROBLEMS

3.1.1 a) Determine the function $f(t)$ whose Fourier transform is shown in Fig. P–3.1.1(a).
b) Determine the function $g(t)$ whose Fourier transform is shown in Fig. P–3.1.1(b).
c) Sketch $f(t)$ and $g(t)$ near $t = 0$. What effects does the phase have on the symmetry of this waveform?

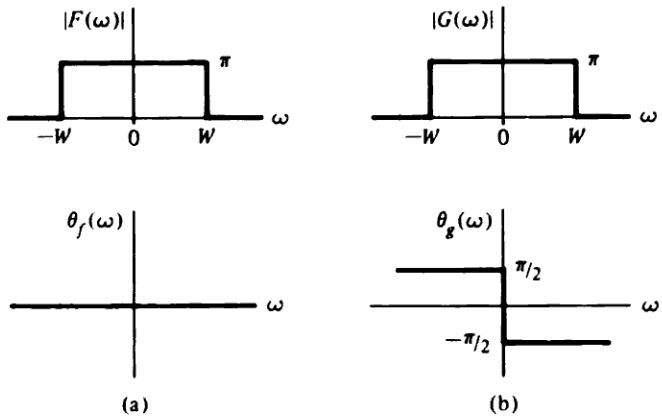


Figure P–3.1.1.

3.1.2 Show that if $F(\omega) = \mathcal{F}\{f(t)\}$, then

a) $F(0) = \int_{-\infty}^{\infty} f(t)\,dt$;

b) $|F(\omega)| \leq \int_{-\infty}^{\infty} |f(t)|\,dt$;

c) $\dfrac{d^2f}{dt^2} \leq \dfrac{1}{2\pi}\int_{-\infty}^{\infty} \omega^2|F(\omega)|\,d\omega$;

d) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t+\tau)e^{-j\omega\tau}\,d\tau\,d\omega\,dt = 2\pi F(0)$.

3.2.1 a) Find the Fourier transform of the raised-cosine pulse signal defined by

$$f(t) = \begin{cases} 1 + \cos \pi t & -1 < t < 1, \\ 0 & \text{elsewhere.} \end{cases}$$

Express your answer in terms of Sa (ω).

b) Use Eq. (3.15) and your answer to part (a) to find the exponential Fourier series coefficients for the following periodic pulse train for the case $T = 2$:

$$\sum_{k=-\infty}^{\infty} f(t - kT).$$

c) Sketch the time waveform of the periodic pulse train in part (b) and then find the exponential Fourier series coefficients directly using Euler's identities.

3.2.2 a) Find the spectral density of the real-valued function:

$$f(t) = \begin{cases} a \exp(-at) & t > 0, \\ b \exp(at) & t < 0. \end{cases}$$

b) Examine your answer to part (a) for the special cases $b = a$ and $b = -a$, particularly with respect to the following chart:

$f(t)$	$F(\omega)$	$F(\omega)$
even	even	real
odd	odd	imag.

3.2.3 A pulse signal described by $f(t) = \exp(-a|t|)$ rect (t/T) is repeated periodically with period T.

a) Find the exponential Fourier series beginning with the Fourier transform of $\exp(-a|t|)$ and then converting to the series.

b) Compare your answer with Problem 2.7.1.

c) Under what conditions can Eq. (3.15) be used to obtain the Fourier series coefficients?

3.2.4 The time function $f(t) = (1/\sigma\sqrt{2\pi})e^{-t^2/2\sigma^2}$ (σ = constant) is known as the gaussian function. This function has finite energy and thus is Fourier transformable. Find its Fourier transform. In carrying out your solution it will be helpful to combine exponents, complete the square in the exponent, and then use the definite integral $\int_{-\infty}^{\infty} e^{-u^2}\,du = \sqrt{\pi}$. Note that $f(t)$ and $F(\omega)$ have the same mathematical form; i.e., the gaussian function is its own Fourier transform.

★ **3.3.1** State whether the Fourier transforms of the following functions, defined over $(-\infty, \infty)$, exist. Defend your answers.

a) $\exp(t^2)$;

b) $\exp(-t^2)$;

c) $1/t$;

d) $\cos(1/t)$;

e) $f(t) = \begin{cases} 1 & t = 0, \\ 0 & \text{elsewhere.} \end{cases}$

3.4.1 Calculate the fraction of the total energy in $f(t)$ within the frequency range $|\omega| < a$ for the first three entries of $f(t)$ in Table 3.1.

3.4.2 Show that a more general statement of Parseval's theorem for energy signals than Eq. (3.21) is

$$\int_{-\infty}^{\infty} f(t)\, g^*(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\, G^*(\omega)\, d\omega .$$

3.4.3 Evaluate the following definite integrals using Parseval's theorem:

a) $\displaystyle\int_{-\infty}^{\infty} [\sin x/x]^2\, dx$;

b) $\displaystyle\int_{-\infty}^{\infty} dx/(a^2 + x^2)$;

c) $\displaystyle\int_{-\infty}^{\infty} dx/(a^2 + x^2)^2$.

3.4.4 Use the result of Problem 3.4.2 to evaluate the following integrals, for $a > 0$, $b > 0$:

a) $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(x^2 + a^2)(x^2 + b^2)}$;

b) $\displaystyle\int_{-\infty}^{\infty} \frac{\mathrm{Sa}\,(bx)}{x^2 + a^2} dx$.

3.5.1 Use Eq. (3.25) and an interchange in the order of integration to show that

a) $\mathcal{F}^{-1}\{\mathcal{F}\{\ \ (t)\}\} = f(t)$;

b) $\mathcal{F}\{\mathcal{F}\{f(t)\}\} = 2\pi f(-t)$.

3.5.2 Determine the average (dc) value of each of the following signals, and verify your answer by first finding $f(t)$, then compute the average value.

a) $F(\omega) = \dfrac{\omega^2 + 2}{\omega^4 + 3\omega^2 + 2} + \pi\delta(\omega)$;

b) $F(\omega) = \dfrac{\omega^2 + 2}{\omega^4 + 3\omega^2 + 2} \pi\delta(\omega)$.

3.5.3 A function $f(t)$ is multiplied by a series of equally spaced impulse functions to form $g(t)$, which can be written as

$$g(t) = f(t) \sum_{k=-\infty}^{\infty} \delta(t - kT) .$$

a) Show that the spectral density of $g(t)$ can be written as a complex exponential series.

b) If $F(\omega)$ is defined over $|\omega| < W$, derive a condition between W and T such that coefficients in the series are linearly independent.

c) Is $G(\omega)$ periodic under the conditions of parts (a) and (b)?

3.5.4 Suppose the exponential term in the Fourier transform is approximated by the first two terms in a MacLaurin series. Designating this integral as $F_1(\omega)$, we have

$$F_1(\omega) = \lim_{T \to \infty} \int_{-T/2}^{T/2} f(t)[1 - j\omega t]\, dt .$$

Find $F_1(\omega)$ for the following functions, compare your results with the corresponding Fourier transforms as given by Eq. (3.13), and then investigate their behavior as $\omega \to 0$.

a) A rect (t/τ);
b) $A \exp(-at)\,u(t)$;
c) $A \exp(-a|t|)$;
d) $A\,\delta(t - t_0)$.

3.6.1 a) Find $F(\omega)$ for the $f(t)$ shown in Fig. P–3.6.1 (a), (b).
b) Sketch $|F(\omega)|$ for $\tau \ll T$ for both cases, and compare.

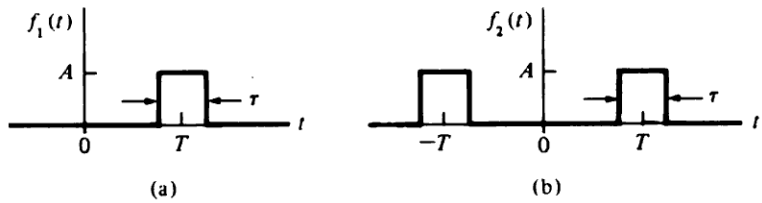


Figure P–3.6.1.

3.6.2 Use the modulation property to find the function $f(t)$ whose Fourier transform is shown in Fig. P–3.6.2 for the conditions:

a) $B = A$; b) $B = -A$.

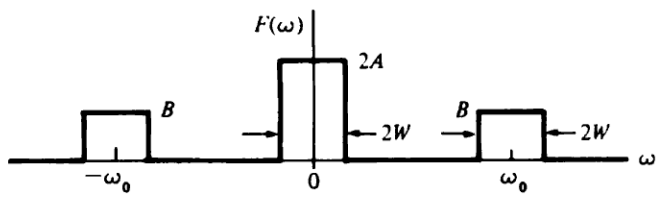


Figure P–3.6.2.

3.6.3 Suppose we define the integral transform $F_1(\omega)$ as described in Problem 3.5.4. Prove the following properties for this transform.

a) $\mathcal{F}_1\{f(at)\} = \dfrac{1}{|a|}F_1\left(\dfrac{\omega}{a}\right)$;

b) $\mathcal{F}_1\{f_{\text{even}}(t)\} = F(0)$ [i.e., a constant, independent of ω];

c) $\mathcal{F}_1\{f_{\text{odd}}(t)\} = \omega\left[\dfrac{dF(\omega)}{d\omega}\right]_{\omega=0}$;

d) $\mathcal{F}_1\{f(t - t_0)\} = F_1(\omega) - j\omega t_0 F(0)$.

3.6.4 If $f(t) \leftrightarrow F(\omega)$, determine the Fourier transform of

a) $f(2 - t)$;
b) $f[(t/2) - 1]$;
c) $f(t) \cos \pi(t - 1)$;
d) $\dfrac{d}{dt}[f(2t)]$.

3.6.5 Find the Fourier transform of the pulse waveform $f(t)$ shown in Fig. P–3.6.5 by differentiating to obtain impulse functions, and then writing the transform using the delay and integration properties. [*Hint:* Consider use of superposition.]

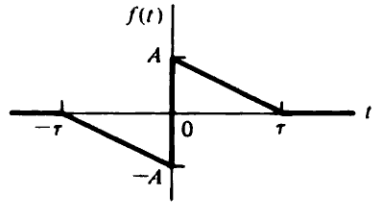


Figure P–3.6.5.

3.6.6 An approximate pulse signal waveform after passage through a *RC* low-pass filter, using piecewise parabolic sections, is shown in Fig. P–3.6.6.

a) Find the Fourier transform of this approximation using the instructions of Problem 3.6.5.

b) From your results, what single characteristic of the waveform predominates in determining the Fourier transform for large values of ω?

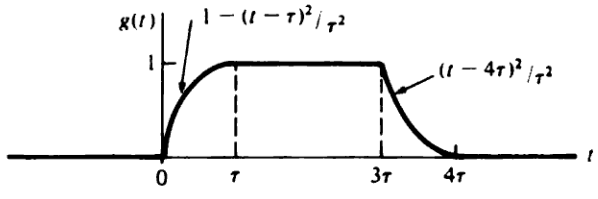


Figure P–3.6.6.

3.6.7 A given pulse waveform is

$$f(t) = \begin{Bmatrix} \cos \pi t & |t| < \frac{1}{2} \\ 0 & \text{elsewhere} \end{Bmatrix}.$$

a) Sketch df/dt and d^2f/dt^2.

b) Using the result for d^2f/dt^2, write a differential equation in terms of $f(t)$ and then solve for $F(\omega)$.

3.6.8 Two functions of time, $f(t)$ and $g(t)$, are known to satisfy the following integral equation:

$$g(t) = \int_{-\infty}^{\infty} g(\tau) f(t - \tau)\, d\tau + \delta(t).$$

a) If $f(t) = \exp(-at)u(t)$, find $g(t)$.

b) If $f(t) = \exp(-a|t|)$, find $g(t)$.

3.6.9 A numeric approximation technique to evaluate the integral of a function $f(t)$ is to divide it into k rectangular strips each of width Δt and height f_k.

a) Derive an approximation for the Fourier transform of $f(t)$ using this approach.

b) Using the properties of the Fourier transform, derive a formula for the exact Fourier transform of $f(t)$ using this approach. Compare this result with that of part (a).

3.6.10 A signal of the form $f(t) = \exp(-at)\,u(t)$ is applied to a linear time-invariant system, and the resulting spectral density of the output $g(t)$ is

$$G(\omega) = c_0/(c_2^2 - \omega^2 + jc_1\omega),$$

where the c_k are constants. If it is known that the impulse response of the system is of the form: $h(t) = b_0 \exp(-b_1 t)\,u(t)$, determine the constants a, b_k in terms of the c_k.

3.7.1 Consider the following (volume) integral:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u)h(v)\,\delta[t - (u + v)]\,du\,dv.$$

a) Show that this integral describes the convolution integral, $f(t) \circledast h(t)$.
b) Using this integral and the result of (a), show that $f(t) \circledast h(t) = h(t) \circledast f(t)$.
c) Use this integral to show that the area under the convolution result of two given functions is equal to the product of the areas under the two functions.
d) Repeat part (c) using the inverse Fourier transform of one function in $f(t) \circledast h(t)$, and then interchanging the order of integration.

3.7.2 Often an overall system is composed of groups of smaller systems, each designed to accomplish a specific objective. Using the convolution integral, show that under these conditions:

a) The order of operation on subsystems is immaterial.
b) The overall system impulse response is the convolution of impulse responses of all subsystems making up the system.
c) The convolution result also holds for bandpass systems; i.e.,

$$f_1(t)\exp(j\omega_0 t) \circledast f_2(t)\exp(j\omega_0 t) = g(t)\exp(j\omega_0 t), \qquad \text{if } g(t) = f_1(t) \circledast f_2(t).$$

d) The overall system delay is given by the sum of the individual delays in the various subsystems. [*Hint:* Use time-limited impulse responses, e.g., rectangularly shaped impulse responses, and define time delay to be the center of the impulse response.]

3.7.3 Use frequency convolution to prove the following trigonometric identities (also see Fig. 3.3):

a) $2\cos^2 \omega_0 t = 1 + \cos 2\omega_0 t$;
b) $2\sin^2 \omega_0 t = 1 - \cos 2\omega_0 t$;
c) $2\cos \omega_1 t \cos \omega_2 t = \cos(\omega_1 + \omega_2)t + \cos(\omega_1 - \omega_2)t$.

3.8.1 Sketch the results of the following convolution operations (where $t_0 > 0$ in all cases); then check your result by writing a Fourier transform of each function, multiplying, and writing the corresponding function of time.

a) $A\,\delta(t) \circledast B\,\delta(t - t_0)$;
b) $A\,\delta(t + t_0) \circledast B\,\delta(t - t_0)$;
c) $A\,\delta(t - t_1) \circledast B\,\delta(t - t_0)$, $t_0 > t_1$;
d) $A[\delta(t + t_0) + \delta(t - t_0)] \circledast B[\delta(t + t_0) + \delta(t - t_0)]$;
e) $A[\delta(t + t_1) + \delta(t - t_1)] \circledast B[\delta(t + t_0) + \delta(t - t_0)]$, $t_0 > t_1$.

3.8.2 Evaluate the following convolution integrals; then check your result by taking a Fourier transform of each function, multiplying, and finding the inverse Fourier transform.

a) $u(t) \circledast e^{-t}u(t)$;
b) $e^{-at}u(t) \circledast e^{-bt}u(t)$;
c) $e^{-a|t|} \circledast \cos \omega_0 t$.

3.8.3 Evaluate $f_1(t) \circledast f_2(t)$, $f_1(t) \circledast f_3(t)$, and $f_2(t) \circledast f_3(t)$ for the functions defined by the waveforms shown in Fig. P–3.8.3.

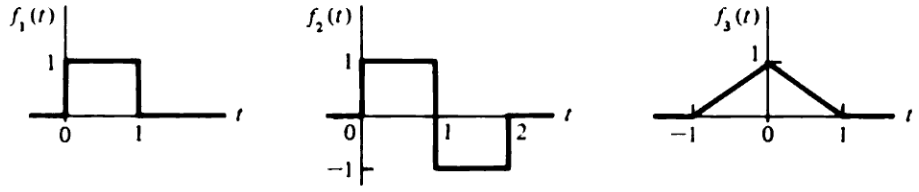


Figure P–3.8.3.

3.8.4 Because the convolution of two impulse functions results in another impulse function, we can handle the convolution of two piecewise linear waveforms by (1) taking derivatives until impulse functions appear, (2) performing the convolution with the impulse functions, and (3) integrating the result as many times as the total number of differentiations. Use this method to perform the convolutions indicated in Problem 3.8.3.

3.8.5 Using the method described in Problem 3.8.4 for the input waveform, find the output of a *RC* low-pass filter for the input pulse waveform shown in

a) Fig. P–3.8.3(a);
b) Fig. 3.9(a).

3.8.6 Sketch $f_1(t) \circledast h(t)$, $f_2(t) \circledast h(t)$, and $f_3(t) \circledast h(t)$ for the functions shown in Fig. P–3.8.6.

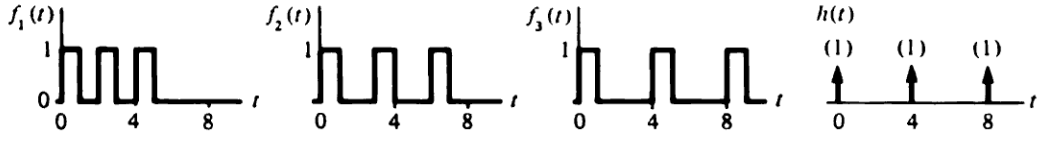


Figure P–3.8.6.

3.8.7 Two functions of time, $f(t)$ and $g(t)$, are defined by

$$f(t) = \begin{cases} e^{-j\omega_1 t} & 0 \le t < T \\ 0 & \text{elsewhere} \end{cases},$$

$$g(t) = \begin{cases} e^{j\omega_1 t} & 0 \le t < T \\ 0 & \text{elsewhere} \end{cases}.$$

Determine and sketch $f(t) \circledast g(t)$ for each of the following conditions:

a) $\omega_1 = 0$;
b) $\omega_1 = \pi/(2T)$.

3.8.8 An averaging, or smoothing, operation used in signal analysis is defined by the integral operator:

$$g(t) = \int_{t-T}^{t} w(t - \tau) f(\tau)\, d\tau,$$

where t indicates present time, T is the averaging duration, $f(t)$ is the function being averaged, and $w(t)$ is a window function.

a) Express $g(t)$ as the convolution of $f(t)$ with a second function, $h(t)$, for the special case in which $w(t) = 1$. Plot and dimension $h(t)$. [*Hint:* Express the limits of integration in terms of step functions in the integrand so that the limits of integration can be extended.]

b) Take the Fourier transform of $g(t)$ to investigate the relative effects of the various factors in the averaging operation in the frequency domain. [*Hint:* Use an interchange in the order of operations.]

3.8.9 A certain function of time, $f(t)$, has a Fourier transform shown in Fig. P–3.8.9.

a) Sketch the Fourier transform of $f(2t)$.

b) Sketch the Fourier transform of $[f(t)]^2$.

c) Sketch the Fourier transform of $[f(t) \cos (15t)]$.

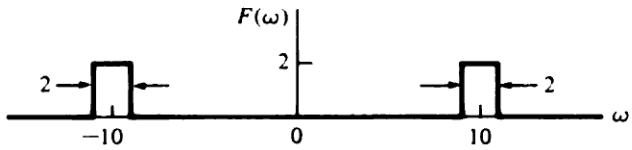


Figure P–3.8.9.

√ **3.8.10** The signal shown in Fig. P–3.8.10 is applied to a filter whose impulse response is $h(t) = ae^{-at}u(t)$. Find the output signal in graphic form. Use numeric methods and space the sample points at 0.1 sec intervals. Assume continuity to the left in the waveform. Compute and plot over (0, 5) for $a = 0.5, 1, 4$.

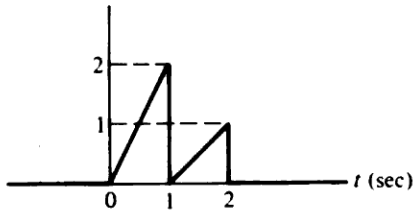


Figure P–3.8.10.

√ **3.8.11** Compute the numeric convolution $x(t) \circledast h(t)$ over the interval $0 \le t < 5$, where

$$x(t) = \begin{cases} \sin \pi t & 0 \le t < t_0 \\ 0 & \text{elsewhere,} \end{cases}$$

$$h(t) = \exp(-t)u(t).$$

Use sample points at increments of 0.1 in each function, and plot the result for a) $t_0 = 1$; b) $t_0 = 2$.

3.9.1 In this problem we investigate the use of phasors, widely used in ac circuit analysis. It is given that $h(t)$ is the impulse response of a linear time-invariant system with frequency transfer function $H(\omega)$.

a) Let $f(t) = Ae^{j\phi}e^{j\omega_0 t}$, and write an integral expression for $g(t) = f(t) \circledast h(t)$. You should recognize a Fourier transform, giving the output in terms of $H(\omega_0)$. Express $H(\omega_0)$ in terms of magnitude and angle, and combine with the magnitude and angle

of $f(t)$. Then take the real part to express the answer in terms of cosines. Note that you do not have to use the Fourier transform (or the convolution integral) in this method, but it requires sinusoidal inputs.

b) Use the method of phasors, as developed in part (a), to find the output $g(t)$ if $f(t) = 10 \cos (20\pi t - \pi/6)$, and $H(\omega) = 30\pi/(j\omega + 30\pi)$.

c) Check your answer to part (b) using Fourier transforms.

d) Check your answer to part (b) using convolution.

3.10.1 a) Find the frequency transfer function of the four-tap transversal filter shown in Fig. P–3.10.1.

b) Determine the value of this frequency transfer function at the following normalized frequencies: $\omega T = 0$; $\omega T = \pi$; $\omega T = \pi/2$.

c) Determine and sketch the magnitude of the frequency transfer function over $-\pi < \omega T < \pi$ for the particular case in which $a_0 = 1$, $a_1 = -2$, $a_2 = 1$, $a_3 = 0$.

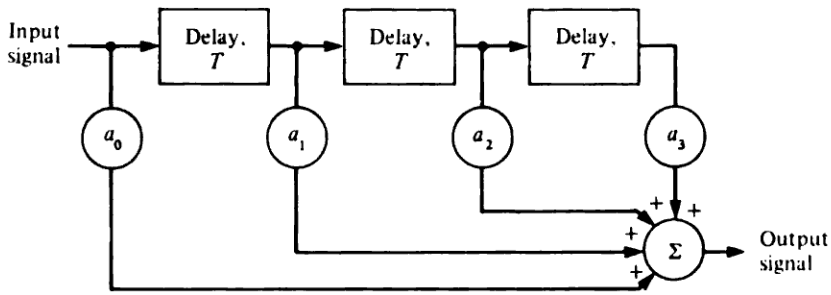


Figure P–3.10.1.

3.10.2 A problem arising in high-frequency communication systems is that if the signal $f(t)$ is sent, the received signal may consist of $f(t)$ plus a delayed replica that arrives later via a longer transmission path. This is called *multipath*. Assuming that the received signal is

$$g(t) = f(t) + \alpha f(t - \tau),$$

design a transversal filter to compensate for this multipath if $\alpha \ll 1$.

√ **3.10.3** A transversal filter such as the one shown in Fig. P–3.10.1 is used in some radar systems for detection of moving targets in the presence of fixed-target signal returns. The delay time T is chosen to be the time between successive radar pulses, so that the radar sampling frequency is $f_s = 1/T$. This filter must reject those frequency components near $\omega = 0$ and emphasize those near $f_s/2$. The filter weights, the a_k, are chosen to accomplish these characteristics.

a) Using a minimum of 50 points, compute and plot the magnitude frequency transfer function of a four-tap transversal filter for the following choice of filter weights: $+1, -1, +1, -1$.

b) Repeat part (a) for the choice of filter weights: $-1, +1, +1, -1$.

3.11.1 Determine the -3 dB bandwidth of the filter whose frequency transfer function is: $H(\omega) = 1/[(j\omega)^2 + 1.4(j\omega) + 1.5]$. Also, calculate the ratio of the -60 dB to -6 dB bandwidths for this filter. (This ratio is often used as a convenient measure of the selectivity of the filter.)

3.11.2 The normalized magnitude frequency transfer functions of a class of low-pass filters known as Butterworth filters are given by $|H_n(\omega)| = 1/\sqrt{1 + \omega^{2n}}$, where n is the order of the filter. The frequency transfer functions of the first three Butterworth filters are

$n = 1 \quad H_1(\omega) = [(j\omega) + 1]^{-1}$
$n = 2 \quad H_2(\omega) = [(j\omega)^2 + \sqrt{2}(j\omega) + 1]^{-1}$
$n = 3 \quad H_3(\omega) = [(j\omega)^3 + 2(j\omega)^2 + 2(j\omega) + 1]^{-1}$

Compute the ratio of the −60 dB to −6 dB bandwidths for these three filters.

3.12.1 An attenuator circuit commonly used in oscilloscope waveform measurements is shown in Fig. P–3.12.1. The oscilloscope probe section of the attenuator has a resistance R_1 and a capacitance C_1. The input impedance of the oscilloscope is R_2 in parallel with C_2.

a) Derive the frequency transfer function of this attenuator circuit.
b) Show that the attenuation is distortionless if $R_1C_1 = R_2C_2$.
c) Derive the impulse response of the attenuator for $R_1C_1 \neq R_2C_2$. What happens as $R_1C_1 \to R_2C_2$? Explain your answer in terms of the impulse response.

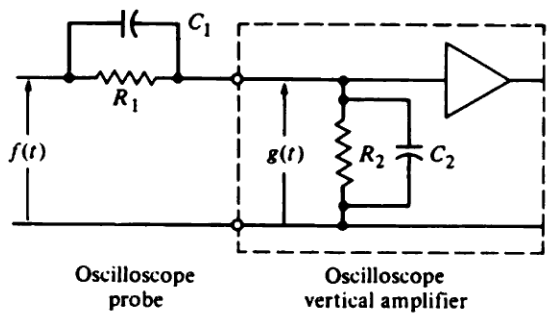


Figure P–3.12.1.

3.12.2 Consider the three Butterworth filters described in Problem 3.11.2. If we insist that the deviation from the requirements for distortionless transmission are to be no greater than 1%, determine the maximum bandwidth, normalized to the −3 dB bandwidth, that may be handled by each of these filters:

a) On the basis of the magnitude of the frequency transfer function.
b) On the basis of the phase of the frequency transfer function.

3.12.3 This problem investigates effects of small nonconstant amplitude and delay deviations from the ideal filter response. The filter output $g(t)$ in terms of the input $f(t)$ and the magnitude response $|H(\omega)|$ and phase response $\theta(\omega)$ is

$$g(t) = \mathcal{R}e\{\mathcal{F}^{-1}[F(\omega)\,|H(\omega)|e^{j\theta(\omega)}]\}.$$

For small deviations, we use the Taylor series approximations:

$$|H(\omega)| \approx A + a_1\omega + a_2\omega^2 + a_3\omega^3 + \cdots$$

$$e^{j\theta(\omega)} \approx 1 + j\theta(\omega) - [\theta(\omega)]^2/2 + \cdots$$

$$\theta(\omega) \approx b_0 - t_0\omega + b_2\omega^2 + b_3\omega^3 + \cdots$$

where A, t_0 are the magnitude and time delay of the ideal filter.

If $H(\omega)$ is a low-pass filter, show that

$$g(t) \approx \mathcal{R}e\{Af(t - t_0) - a_2f^{(2)}(t - t_0) - jAb_3f^{(3)}(t - t_0) + \cdots\},$$

where $f^{(n)}(t)$ denotes the nth derivative of $f(t)$. Note that the term $Af(t - t_0)$ represents the undistorted output, and all other terms are distortion.

3.13.1 The frequency transfer function of a given system is

$$H(\omega) = 10^{18}/(\omega^2 + 10^{16}).$$

a) Determine the −3 dB bandwidth for this system.
b) Determine the ratio of the −60 dB to −6 dB bandwidths.
c) Find the 10% to 90% rise time, in nanoseconds, and compare with the reciprocal of the −3 dB bandwidth for this system. [*Hint:* First find the impulse response.]
d) Is this system physically realizable?

3.13.2 Consider the two linear time-invariant systems whose (idealized) impulse responses are shown in Fig. P–3.13.2.

a) We define the time delay t_d as the time required for the system output to rise to 50% of the final output value in response to a unit step function applied to the input at $t = 0$. We define the rise time t_r as the time required for the system output to rise from 10% to 90% of the final output in response to a unit step applied to the input at $t = 0$. Determine the time delay t_d and the rise time t_r for each of the systems. Note that t_d is a measure of the center of the impulse response and t_r is a measure of the width of the impulse response.
b) The two systems are connected in cascade. Find the impulse response of the overall system.
c) Determine the time delay of the overall system, and compare with the sum of the time delays of the individual systems.
d) Determine the rise time of the overall system. [*Hint:* Use symmetry where possible.]
e) Normalize your results in part (d) by dividing by the square root of the sum of the squares of the individual rise times. Investigate this ratio numerically over a range of (t_1/t_2) of 0.1 to 10. What is your conclusion?

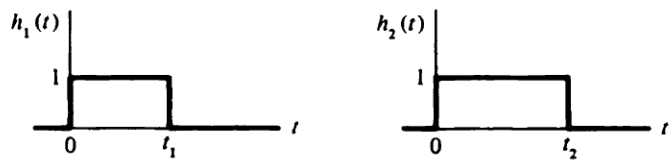


Figure P–3.13.2.

3.13.3 A common rule in the design and operation of scanning spectrum analyzers is that "the frequency sweep rate must be less than the square of the filter bandwidth." In fact, a warning light is used on some spectrum analyzers to warn the user when this inequality is not satisfied. Defend this rule.

3.13.4 If the drive current to a semiconductor light emitting diode (LED) is varied at a frequency ω, the intensity of the optical output is found to vary as

$$I(\omega) = I_0[1 + \omega\tau_{\text{eff}}]^{-1/2}$$

where I_0 is the intensity emitted at $\omega = 0$ and τ_{eff} is the effective carrier lifetime.

a) What is the −3 dB bandwidth of the device?
b) Using the approximation given in Eq. (3.79), estimate the rise time of the device.

★ **3.14.1** Using Eqs. (3.80) and (3.81), determine the time-bandwidth product of the pulse signal $f(t) = t \exp(-at)u(t)$.

★ **3.14.2** Determine the time-bandwidth product of the pulse signal $f(t) = \exp(-a|t|)$.

3.15.1 A device commonly used to check the rotational rate of audio record turntables is a circular disc with equally spaced identical dark lines on a contrasting background and placed radially on the disc. When the disc is viewed under a strobe light, the pattern is designed to remain stationary if the rotational rate is correct. For example, suppose the turntable speed is set for $33\frac{1}{3}$ rpm. A strobe (e.g., a neon or even a fluorescent light) is connected to the 60-Hz ac line and flashes at 120 sec^{-1}. Calculate the minimum number of lines required on the disc. Repeat the calculation if the line frequency is 50 Hz.

3.15.2 A signal $f(t)$ whose spectral density is shown in Fig. P–3.15.2 is sampled using a periodic rectangular pulse waveform. The width of each sampling pulse is 15 msec and the period is equal to the Nyquist interval. Sketch the magnitude spectral density of the sampled signal from zero to 100 Hz, labeling important points.

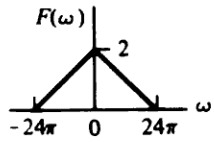


Figure P–3.15.2.

3.15.3 A sampling system composed of an ideal multiplier and an ideal low-pass filter is shown in Fig. P–3.15.3.

a) Let $p(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT)$, and sketch the filter input and filter output for the case in which T is chosen to be equal to the Nyquist sampling interval.

b) Let $p(t) = \sum_{k=-\infty}^{\infty} (-1)^k \delta(t - kT/2)$, and sketch the filter input and filter output for the case in which T is chosen to be equal to the Nyquist sampling interval.

c) Lower the sampling rate in part (b) to one-half the Nyquist sampling rate, and sketch the filter input and filter output. Explain your result.

d) If $z(t)$ in part (c) were applied to the input of a second system that is a replica of the system in (c), will the output of this second system resemble $f(t)$? Explain.

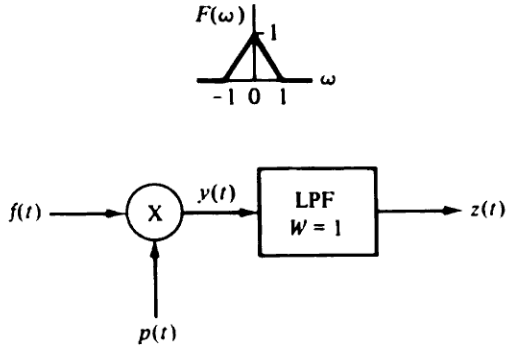


Figure P–3.15.3.

3.15.4 The spectral density of a given band-limited signal $f(t)$ is shown in Fig. P–3.15.4. Using the sampling system in Problem 3.15.3(a), determine the minimum sampling frequency for

a) $\omega_1 = 8$, $\omega_2 = 10$;

b) $\omega_1 = 7$, $\omega_2 = 9$.

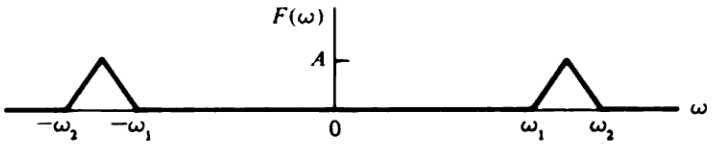


Figure P–3.15.4.

★ **3.16.1** Here we wish to investigate the possible use of signal energy as a measure of the relative amount of aliasing present as a result of sampling a given $f(t)$ with a set of impulse functions (see, e.g., Problem 3.15.3a).

a) Develop an expression for the ratio of the energy introduced into the half-sampling rate bandwidth, divided by the signal energy within that bandwidth with no aliasing present.

b) Use your result to determine the minimum sampling rate, expressed in integer multiples of the -3 dB bandwidth, to make this ratio at least 10 dB for the signal $f(t) = \exp(-at)u(t)$.

★ **3.16.2** While Problem 3.16.1 treats aliasing strictly on an energy (i.e., magnitude) basis, there are also some phase interference effects that can arise as a result of sampling, particularly near the edge of the bandwidth. To investigate this, assume an impulse sampler that samples the signal $f(t) = \cos(\omega_s/2 - \Delta\omega)t$ at the rate $\omega_s = 2\pi/T$. The output of this sampler is filtered by a low-pass filter with frequency transfer function $H(\omega)$. Show that the resulting output can be written as an expression of the form

$$[H(\omega_2) + H(\omega_1)]\cos(\Delta\omega t)\cos[(\omega_s/2)t] + [H(\omega_2) - H(\omega_1)]\sin(\Delta\omega t)\sin[(\omega_s/2)t],$$

where $\omega_1 = \omega_s/2 - \Delta\omega$, and $\omega_2 = \omega_s/2 + \Delta\omega$. Note that in the absence of the filter, the output is of the form:

$$\cos(\Delta\omega t)\cos[(\omega_s/2)t].$$

√★ **3.16.3** The signal $f(t) = 5e^{-t}$, $0 \le t < 2\pi$, is repeated periodically with $T = 2\pi$.

a) Find an expression for the exact value of the F_n using integration.

b) Compute the magnitude of the F_n for the first 20 harmonics using 20 samples per period.

c) Tabulate the percent error in the computed values of the F_n as a function of n.

d) Repeat parts (b), (c), except using 40 samples per period. Compare your results with those obtained in parts (b), (c).

√★ **3.16.4** Compute the percent aliasing error in Problem 3.16.3, based on the exact values for the F_n; repeat using a RC low-pass pre-alias filter whose -3 dB frequency is 90% of the half-sampling frequency.

★ **3.17.1** Compute the DFT of the following sequences for $N = 4$:

a) $\{1, 0, 0, 0\}$;
b) $\{1, -1, 1, -1\}$;
c) $\{1, j, -1, -j\}$;
d) $\exp(-k)$.

√★ **3.17.2** a) Compute the magnitude of the Fourier transform of $f(t) = \exp(-t)u(t)$ over $(0, 2)$ using the DFT and $N = 16$ sample points. Assign a value of 1 to the point at $t = 0$.

b) Compare your result with that of the magnitude of the continuous Fourier transform evaluated at the discrete frequency points, and plot both results for $0 \le n \le N$.

√★ **3.17.3** Compute the magnitude of the Fourier transform of the following functions using the DFT and 16 samples over the interval (0, 1): a) $\sin \pi t$; b) $\sin \sqrt{2}\pi t$; c) $\sin 2\pi t$. Explain any significant differences in their spectral densities.

√★ **3.17.4** Using the DFT, compute and plot the energy spectral density of the following functions over $0 < t \leq 5$:

a) $f(t) = (1/t) - (1/4)$; b) $f(t) = (1/\sqrt{t}) - (1/4)$.

Use 50 equally spaced time samples and plot in log-log coordinates over the integer frequency increments: $1, 2, 3, \ldots, 20$. Then use an approximation of the form: kn^{-r} for each plot, and find the numerical values for k, r.

√★ **3.17.5** Evaluate a) $f_1(t) \circledast f_1(t)$, and b) $f_1(t) \circledast f_3(t)$ in Problem 3.8.3 by taking the DFT of each, multiplying, and then taking the IDFT. Use sample spacings of 0.1, assume continuity to the left, and use augmenting zeros to fill 40 samples in each sequence. Plot your results over (0, 40).

√★ **3.17.6** Use the DFT to compute the real part of the Fourier transform for a waveform $f(t)$ defined by

$$f(t) = \begin{cases} 1 & -0.05r < t \leq 0.05r \\ 0 & \text{elsewhere,} \end{cases}$$

where r is an odd integer (to be assigned). Use 50 sample points, with 0.1 between points and place one sample point at $t = 0$. Plot your answer versus the integer frequency scale n, $0 \leq n \leq N$, with $r = 11$ (or another value assigned by your instructor).

√★ **3.18.1** Using the FFT, compute the Fourier transform of

$$f(t) = (20/\pi)\,\mathrm{Sa}\,(20t)\ \mathrm{rect}\,[t/(r\pi)]; \quad -2\pi \leq t < 2\pi\,.$$

Let $N = 256$, and plot the magnitude of the Fourier transform for $0 \leq n \leq 100$, and a) $r = 1$; b) $r = 2$. Compare results.

√★ **3.18.2** Using a 64-point FFT algorithm, graph a numerical approximation for the output of an ideal low-pass filter that passes only the first seven harmonics (plus zero) when the signal waveform shown in Fig. P–3.8.10 is applied to the input.

√★ **3.18.3** The Hamming window function is defined by

$$w_H(t) = \begin{cases} 0.54 + 0.46 \cos\,[2\pi t/(NT)] & |t| < NT/2 \\ 0 & |t| > NT/2 \end{cases}.$$

a) Using the FFT, compute and plot the magnitude of the spectral density for both the raised-cosine and the Hamming window functions. Choose 64 sample points for each, plus 192 augmenting zeros, and plot for $0 \leq n \leq 64$. Add a small offset (e.g., 1×10^{-5}) to avoid zeros, and plot the magnitude on a logarithmic scale.

b) Compare the two spectral densities on the basis of (1) the width of the main lobes; (2) the relative magnitudes of the first sidelobes; and (3) the rate of decrease of the sidelobe levels for increasing n.

√★ **3.18.4** Repeat Problem 3.18.3 for the case of the Parzen window function defined by

$$w_P(t) = \begin{cases} 1 - 6|x|^2(1 - |x|) & |x| \leq \frac{1}{2}, \\ 2(1 - |x|)^3 & \frac{1}{2} < |x| \leq 1, \\ 0 & |x| > 1, \end{cases}$$

where $x = 2t/(NT)$.

√★ **3.18.5** Compute and plot the output signal in Problem 3.8.10 using the FFT with 100 sample points plus 156 augmenting zeros.

√★ **3.18.6** The discrete signal $f(k)$, defined by

$$f(k) = \begin{Bmatrix} 6 & k = 27, 37 \\ 0 & \text{otherwise} \end{Bmatrix}, \qquad 0 \le k < 64,$$

is passed through a LPF such that only the first five harmonics are not attenuated to zero. Use the FFT to compute the resulting output, and plot. [*Hint:* It may be helpful to refer to Fig. 7.18 as an aid in interpretation of the results.]

√★ **3.18.7** Noting the symmetry of the Fourier transform and its inverse, we can make use of the FFT to investigate time waveforms for given spectral densities. Using this approach, compute the FFT of the following discrete spectral densities for $0 \le k < 128$, and plot the resulting time waveforms for $0 \le n < 64$. [*Hint:* It may be helpful to refer to Fig. 5.3(a), (b) as an aid in interpretation of the results.]

a) $f_1(k) = \begin{cases} 5 & k = 4, 124 \\ 0 & \text{otherwise;} \end{cases}$

b) $f_2(k) = \begin{cases} 5 & k = 60, 68 \\ 0 & \text{otherwise.} \end{cases}$

DESIGN PROBLEMS

Design Problem

3.1 Waveform Synthesizer Design

Using a digital computer, design a simple waveform synthesizer for piecewise linear waveforms by assigning four weights (which form the inputs) to four equally spaced impulse functions, followed by a fifth with the same spacing whose weight is the negative sum (why?) of the preceding four. Provide two displays: (1) a graph of the synthesized time function formed by (double) integration of the weights, and (2) a second graph showing the magnitude of the Fourier transform of the synthesized waveform. Explain your design, and include graphs of several synthesized waveforms in a demonstration of the design.

CHAPTER

4 Spectral Density and Correlation

The Fourier methods of signal analysis that have been discussed in the preceding chapters give some valuable insights into the transmission of signals through systems. A topic of particular importance is that of the distribution of energy and power in both time and frequency. While our motivation for considering this topic is from a deterministic viewpoint, we find that the spectral representation of power can be applied also to a description of random signals and noise. In this chapter we develop some methods for handling the averaged effects of noise on a spectral basis. These methods will prove useful in the succeeding chapters.

4.1 ENERGY SPECTRAL DENSITY

Parseval's theorem gives a relation between a time signal $f(t)$ and its Fourier transform $F(\omega)$ [cf. Eq. (3.21)]:

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2\, d\omega\,. \tag{4.1}$$

The integral on the left-hand side of Eq. (4.1) is the energy in $f(t)$ across a 1-ohm resistance. Thus the quantity $|F(\omega)|^2$ appearing in Eq. (4.1) is the energy per unit of frequency normalized to a resistance of 1 ohm. For resistive loads not equal to 1 ohm the following reasoning is helpful.

The energy in a waveform can be found by integrating the product of the voltage and the current (i.e., the instantaneous power) over the interval under consideration. If we designate the voltage by $v(t)$ and the current by $i(t)$, then the energy is[†]

$$E = \int_{-\infty}^{\infty} v(t) i^*(t)\, dt\,. \tag{4.2}$$

However, we are using the mathematical representation $f(t)$ for a signal whether it is a voltage or a current. If the resistive load is 1 ohm, then it doesn't make any differ-

[†] We are using a generalized complex-valued notation of which the more familiar phasor notation, $v(t) = V \exp(j\omega_0 t)$, $i(t) = I \exp(j\omega_0 t)$, is a special case.

ence and we can proceed. For resistive loads not equal to 1 ohm, the resistance scaling must be introduced using the familiar Ohm's law: $v(t) = Ri(t)$. Suppose, for example, that $f(t)$ represents a voltage. Using Parseval's theorem, we find the energy relations are

$$\left.\begin{aligned} E_f &= \frac{1}{R}\int_{-\infty}^{\infty} |f(t)|^2\,dt \qquad (4.3a)\\ E_f &= \frac{1}{2\pi R}\int_{-\infty}^{\infty} |F(\omega)|^2\,d\omega \qquad (4.3b)\end{aligned}\right\}\ \text{for } f(t) \text{ a voltage.}$$

Analogous relations can be written for the case where $f(t)$ represents a current:

$$\left.\begin{aligned} E_f &= R\int_{-\infty}^{\infty} |f(t)|^2\,dt \qquad (4.4a)\\ E_f &= \frac{R}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\,d\omega \qquad (4.4b)\end{aligned}\right\}\ \text{for } f(t) \text{ a current.}$$

Therefore the resistance is merely a scaling factor which can be taken into account later. The conversion factors are easy to supply for any given condition and we shall return to our assumption of a 1-ohm resistive load. The dimension of energy in the MKS system is the joule.

For a 1-ohm resistance, the quantity $|F(\omega)|^2$ appearing in the above expressions is in terms of energy per unit of frequency. For this reason the quantity $|F(\omega)|^2$ is called the *energy spectral density* of the signal $f(t)$, often shortened to the term *energy density*. Thus the energy spectral density is that function (1) that describes the relative amount of energy of a given signal versus frequency, and (2) whose total area under $|F(\omega)|^2$ is the energy of the signal. Note that the quantity $|F(\omega)|^2$ describes only the relative amount of energy at various frequencies. For continuous $|F(\omega)|^2$, the energy at any given frequency is zero—it is the *area* under $|F(\omega)|^2$ that contributes energy. To find energy, then, we must be given a range of frequencies over which to integrate and the amount of energy is dependent on the magnitude squared of $F(\omega)$ over that given range of frequencies.

The concept of energy spectral density is an important one, for it permits us to account for relative spectral-energy attenuations through linear systems. To see this, let us apply a signal $f(t)$ to the input of a linear time-invariant system whose frequency transfer function is $H(\omega)$. Describing the (amplitude) spectral density of the output by $G(\omega)$, we have

$$G(\omega) = F(\omega)H(\omega)\,.$$

Therefore the (normalized) energy density of $G(\omega)$ is[†]

$$|G(\omega)|^2 = |F(\omega)|^2|H(\omega)|^2, \qquad (4.5)$$

[†] The system frequency transfer function, $H(\omega)$, is not required to be dimensionless; i.e., it can convert a voltage to a current or a current to a voltage. The appropriate unit changes can be made along with the input resistance scale changes after all other calculations have been made. For our purposes, $H(\omega)$ is dimensionless and normalized to 1 ohm.

and the (normalized) energy in the output signal is

$$E_g = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 |H(\omega)|^2 \, d\omega \,. \tag{4.6}$$

In other words, the energy density of the system response is given by the energy density of the system input multiplied by the squared magnitude of the system transfer function. All phase information in both the input signal and the system transfer function is lost in the calculation of energy and energy density. Therefore only the magnitude of the system transfer function need be considered in any calculation of energy density.

An interesting and useful physical interpretation of energy density can be obtained through the use of Eq. (4.6). Consider a signal $f(t)$ applied to the input of a very narrow bandpass filter whose frequency transfer function $H(\omega)$ is shown in Fig. 4.1. Designating the output of the narrow-band filter as $g(t)$, we find the energy in $g(t)$ is

$$\begin{aligned} E_g &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(\omega)|^2 \, d\omega \\ &= \frac{1}{2\pi}\int_{-\omega_0-(\Delta\omega/2)}^{-\omega_0+(\Delta\omega/2)} |F(\omega)|^2 |H(\omega)|^2 \, d\omega + \frac{1}{2\pi}\int_{\omega_0-(\Delta\omega/2)}^{\omega_0+(\Delta\omega/2)} |F(\omega)|^2 |H(\omega)|^2 \, d\omega \\ &\cong \frac{1}{2\pi}|F(-\omega_0)|^2 \Delta\omega + \frac{1}{2\pi}|F(\omega_0)|^2 \Delta\omega \,, \end{aligned} \tag{4.7}$$

where $\Delta\omega$ has been assumed to be small in Eq. (4.7).

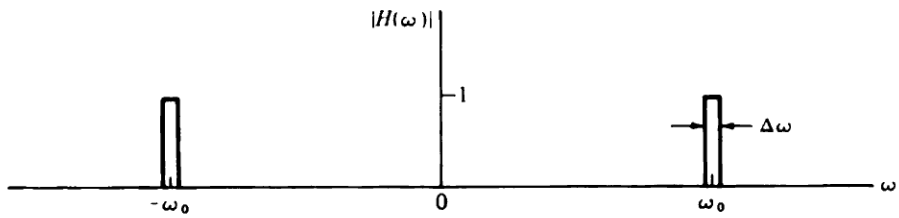


Figure 4.1 An ideal narrow bandpass filter.

If the signal $f(t)$ is real-valued, then $F^*(-\omega) = F(\omega)$ (see Section 3.6.2) and thus $|F(-\omega)| = |F(\omega)|$. An important consequence of this is that *for all real-valued signals the energy spectral density is an even function of* ω. Proceeding, Eq. (4.7) simplifies to

$$E_g \cong \frac{1}{\pi}|F(\omega_0)|^2 \Delta\omega \,, \tag{4.8}$$

if $f(t)$ is a real-valued signal. Note that half of the energy is contributed by the negative frequency components and half by the positive frequency components if the signal $f(t)$ is real-valued.

The practical significance of the above discussion can be realized by reversing the procedure. Given a pulse signal $f(t)$, how can we find its energy spectral density? One way is to extend the above reasoning and to construct a parallel bank of narrow-band filters, all filters positioned in frequency adjacent to each other. If we apply $f(t)$ to this parallel bank of filters, as shown in Fig. 4.2(a), we can obtain an approximation to the energy spectral distribution of $f(t)$. This is illustrated in Fig. 4.2(b). Assigning one-half the energy contributions to negative frequency components yields the final result shown in Fig. 4.2(c). A device available for performing this function is called a "multi-channel spectral analyzer."

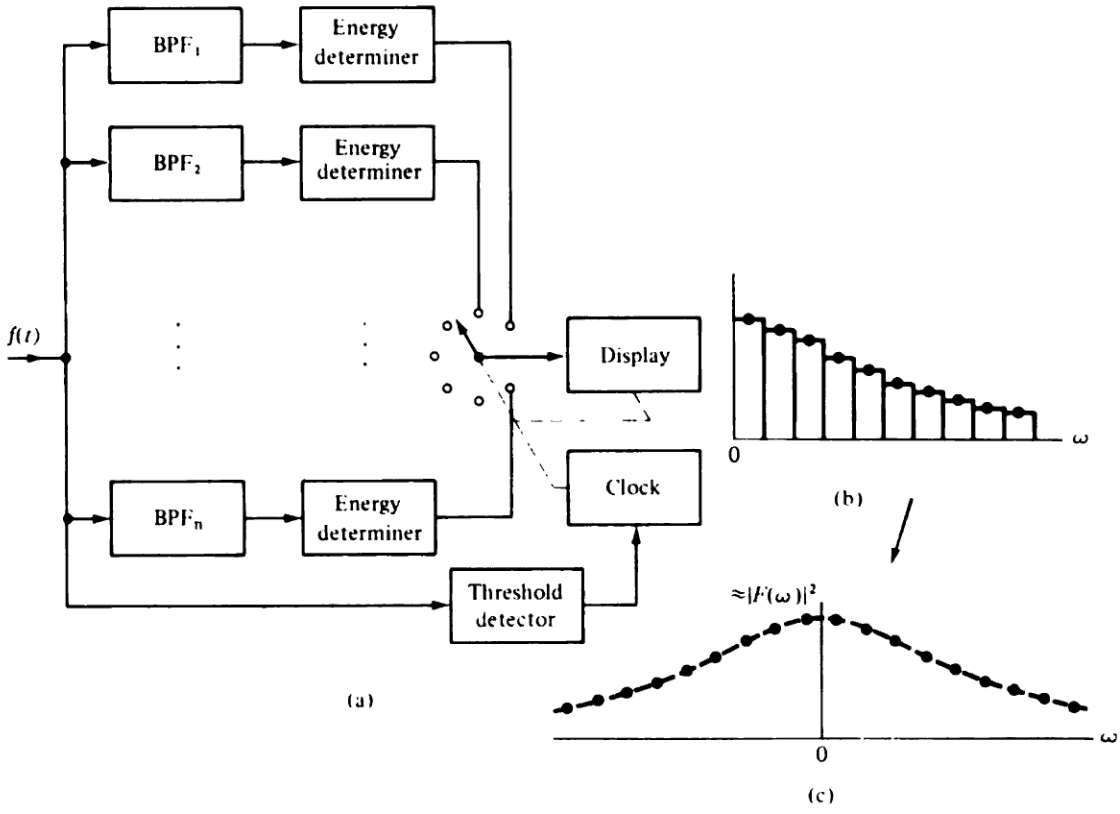


Figure 4.2 Measurement of energy spectral density.

In summary of our discussion, the energy spectral density of a signal represents its energy per unit of frequency and displays the relative energy contributions of the various frequency components. The area under the energy spectral density gives the energy within a given band of frequencies.

EXAMPLE 4.1.1

A voltage signal described by $f(t) = e^{-5t}u(t)$ is applied to the input of an ideal low-pass filter. The low-frequency gain of the filter is unity, the bandwidth is 5 radians

per second, and the resistance levels are 50 ohms. Calculate the energy of the input signal and of the output signal.

Solution The energy in the input signal $f(t)$ is

$$E_f = \frac{1}{R}\int_{-\infty}^{\infty} |f(t)|^2\,dt = \frac{1}{50}\int_0^{\infty} e^{-10t}\,dt$$
$$= \left(\frac{1}{50}\right)\left(\frac{1}{10}\right) = 0.002 \text{ joule.}$$

The energy in the output signal $g(t)$ is

$$E_g = \frac{1}{2\pi R}\int_{-\infty}^{\infty} |G(\omega)|^2\,d\omega = \frac{1}{2\pi R}\int_{-\infty}^{\infty} |F(\omega)|^2|H(\omega)|^2\,d\omega$$
$$= \frac{1}{2\pi R}\int_{-5}^{5} \frac{d\omega}{\omega^2 + 25} = \frac{1}{\pi R}\int_0^5 \frac{d\omega}{\omega^2 + 25}$$
$$= \frac{1}{5\pi R}\tan^{-1}(1) = \frac{1}{5\pi R}\left(\frac{\pi}{4}\right) = 0.001 \text{ joule}.$$

EXAMPLE 4.1.2

A signal $f(t) = e^{-at}u(t)$ is applied to the input of a low-pass filter with a magnitude frequency transfer function $|H(\omega)| = b/\sqrt{\omega^2 + b^2}$. Determine the required relations between the constants a, b such that exactly 50% of the input signal energy, on a 1-ohm basis, is transferred to the output.

Solution The energy in the input signal $f(t)$ is (across 1 ohm)

$$E_f = \int_{-\infty}^{\infty} |f(t)|^2\,dt = \int_0^{\infty} e^{-2at}\,dt = \frac{1}{2a}.$$

The energy in the output signal $g(t)$ is (across 1 ohm)

$$E_g = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(\omega)|^2\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2|H(\omega)|^2\,d\omega$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{b^2}{(\omega^2 + a^2)(\omega^2 + b^2)}\,d\omega$$
$$= \frac{b^2}{\pi}\int_0^{\infty} \frac{d\omega}{(\omega^2 + a^2)(\omega^2 + b^2)}.$$

Using a table of integrals (see, for example, Appendix A),

$$E_g = \frac{b^2}{\pi}\frac{\pi}{2ab(a + b)} = \frac{b}{2a(a + b)},$$

and we require $E_g = \frac{1}{2}E_f$ so that

$$\frac{b}{2a(a+b)} = \frac{1}{2}\frac{1}{2a}.$$

Solving, we find that the required relation is $a = b$.

4.2 POWER SPECTRAL DENSITY

Not all signals of interest have finite energy. Some signals have infinite energy, but they may have a finite time-average of energy. This time-average of energy is called average power and such signals are called power signals.

The time-averaged power of a signal (again, we assume a 1-ohm resistance) is given by

$$P = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} |f(t)|^2\,dt. \tag{4.9}$$

For a periodic signal, each period contains a replica of the function, and the limiting operation in Eq. (4.9) can be omitted as long at T is taken as the period. There are power signals that are not periodic, however, and we shall retain the indicated limiting operation for the general case.

The operation described by Eq. (4.9) is the average (or mean) of the squared signal $f(t)$. Thus this quantity is called the mean-square value of the signal $f(t)$, designated simply as $\overline{f^2(t)}$. Stated another way, the time-averaged power of a signal is its mean-square value, $\overline{f^2(t)}$, if the resistance is 1 ohm. The scaling for other resistance levels follows in the same manner as was discussed for energy and energy spectral density.

In an analogy with the energy signals discussed in the previous section, it would be very convenient and useful for us to define some new function in frequency which would give us some indication of the relative power contributions at various frequencies. Let us define such a function as the *power spectral-density function*, $S_f(\omega)$. This function is in units of power (i.e., watts) per Hz and its integral yields the power in $f(t)$. Writing this mathematically, we have

$$P \triangleq \frac{1}{2\pi}\int_{-\infty}^{\infty} S_f(\omega)\,d\omega. \tag{4.10}$$

The power spectral-density function $S_f(\omega)$ describes the distribution of power versus frequency and hence it is an important measurement in practical systems. Although we are approaching it using deterministic signal theory, it is valid also for random signals and forms a basis for a meaningful measurement of many random signals.

We can use the following intuitive approach to relate the power spectral density $S_f(\omega)$ more closely to the signal $f(t)$. Suppose we are given a power signal $f(t)$ as shown in Fig. 4.3(a). Let us form a truncated version of this power signal by observ-

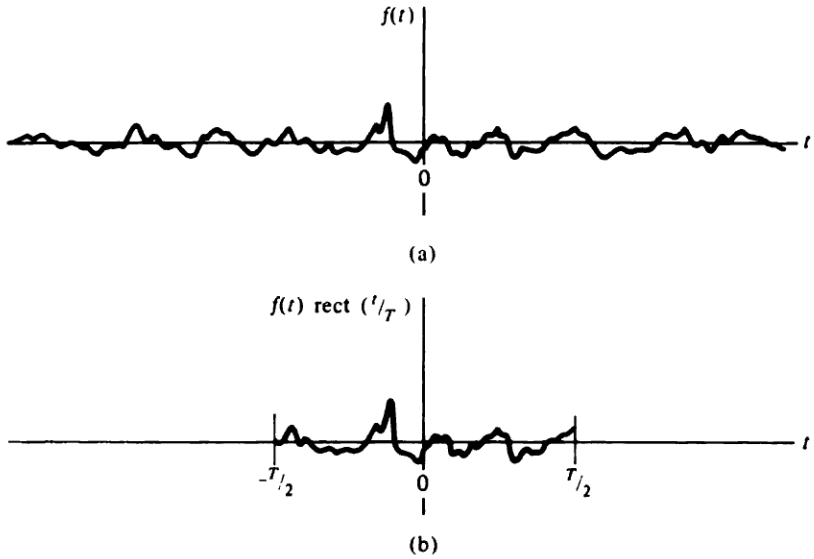


Figure 4.3 (a) A power signal and (b) its finite-interval truncation.

ing it only in the interval $(-T/2, T/2)$, as shown in Fig. 4.3(b). This truncated function can be written as $f(t)$ rect (t/T).

We shall assume that $f(t)$ is finite over the finite interval $(-T/2, T/2)$. Then the truncated function $f(t)$ rect (t/T) has finite energy and its Fourier transform $F_T(\omega)$ is†

$$F_T(\omega) = \mathcal{F}\{f(t) \text{ rect } (t/T)\}. \tag{4.11}$$

Parseval's theorem for this truncated function is

$$\int_{-T/2}^{T/2} |f(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F_T(\omega)|^2 \, d\omega .$$

Hence the average power P across a 1-ohm resistor is given by

$$P = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2 \, dt = \lim_{T\to\infty} \frac{1}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} |F_T(\omega)|^2 \, d\omega . \tag{4.12}$$

Combining Eqs. (4.10) and (4.12), we have

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega) \, d\omega = \lim_{T\to\infty} \frac{1}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} |F_T(\omega)|^2 \, d\omega . \tag{4.13}$$

† We are departing from the notation used formerly where the subscript T was used to indicate a periodic function of period T. The alternative use of the subscript T here should not be confusing because it is used on functions of frequency. Both uses of the symbol T (i.e., to indicate truncation and periodicity) are widely used.

In addition, we insist that this relation should hold over each frequency increment so that Eq. (4.13) becomes

$$G_f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\omega} S_f(u)\,du = \lim_{T\to\infty}\frac{1}{T}\frac{1}{2\pi}\int_{-\infty}^{\omega}|F_T(u)|^2\,du\,, \tag{4.14}$$

where $G_f(\omega)$ represents the cumulative amount of power for all frequency components below a given frequency ω. For this reason, $G_f(\omega)$ is called the cumulative power spectrum or, equivalently, the integrated power spectrum of $f(t)$. The cumulative power spectrum $G_f(\omega)$ always exists for any power signal.

If an interchange in the order of the limiting operation and the integration is valid, Eq. (4.14) becomes

$$2\pi G_f(\omega) = \int_{-\infty}^{\omega} S_f(u)\,du = \int_{-\infty}^{\omega}\lim_{T\to\infty}\frac{|F_T(u)|^2}{T}\,du\,. \tag{4.15}$$

Note that the average or mean power contained in any frequency interval (ω_1, ω_2) is $[G_f(\omega_2) - G_f(\omega_1)]$. In many cases of interest, $G_f(\omega)$ is differentiable and we have

$$2\pi\frac{dG_f(\omega)}{d\omega} = S_f(\omega)\,. \tag{4.16}$$

Under these conditions Eq. (4.15) gives

$$S_f(\omega) = \lim_{T\to\infty}\frac{|F_T(\omega)|^2}{T}\,. \tag{4.17}$$

Equation (4.17) is our desired result for the power spectral density of $f(t)$.

An intuitive explanation of the result given in Eq. (4.17) can be given as follows. The energy of the truncated function increases (or at least does not decrease) with increasing T. Therefore the quantity $|F_T(\omega)|^2$ increases (or at least is not decreasing) with increasing T. As T becomes very large, the fluctuations due to end effects in the integration should become small and thus the quantity $|F_T(\omega)|^2/T$ may approach a limit. Assuming that such a limit exists, we find the power spectral density $S_f(\omega)$ is given by Eq. (4.17). Actually, it is only in a statistical-average sense that the limit in Eq. (4.17) is meaningful for signals with random components (see Chapter 8).

In practical usage the term *power spectral density* is often shortened to *power density spectrum* or just *power spectrum*. Thus we speak of $S_f(\omega)$ as the power spectrum of the signal $f(t)$ even though this terminology is neither as descriptive nor as precise as power spectral density. Note that the power spectral density of a signal retains only the magnitude information and all phase information is omitted. It follows that all signals with the same magnitude spectral densities have identical power spectra, regardless of some possible differences in phase characteristics. Thus, in contrast to the amplitude spectral density $F(\omega)$, $S_f(\omega)$ does not uniquely describe $f(t)$. For a given signal there is a specified power spectral density, but there may be many signals having the same power spectral density.

Equation (4.17) suggests a method used to determine the power spectral density of a signal. A record length of T units is taken from the recording of a physical process—suppose, for example, that we are recording vertical wave motion in an ocean bay. Fixing a float to a recording mechanism, we record T seconds of data. Next, using computational techniques, the corresponding Fourier transform $F_T(\omega)$ is found. (If we use the sampling theorem, analog data can be converted to digital data so that digital computers can be employed.) Taking the magnitude squared, the ratio $|F_T(\omega)|^2/T$ is formed.

Obviously, the limit is not taken (to do so requires infinite record length) and errors are introduced in the measurement. One limitation is that the best resolution in frequency which can be obtained will be on the order of $1/T$ [cf., for example, Eq. (3.106)]. To get better frequency resolution, then, requires longer record lengths. Second, most physical processes have random components and therefore fluctuations arise in the power spectral density. If one is reasonably certain that the true power spectrum does not vary with time, then a possible remedy is to make several consecutive recordings, each of length T, and then average the power spectral densities after computation. These and other methods for obtaining power spectral densities and enhancing their frequency resolution form the subject of an area known as *spectral analysis*. A discussion of spectral analysis requires a knowledge of random signal processes and will have to be postponed.

The above discussion holds for any general power signal. We can proceed a little farther if we have a periodic power signal. Assume that $f(t)$ is periodic and that it is represented by the exponential Fourier series

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t}.$$

Using Parseval's theorem [cf. Eq. (2.69)], we have

$$\overline{f^2(t)} = \sum_{n=-\infty}^{\infty} |F_n|^2. \tag{4.18}$$

Equation (4.18) gives the power across a 1-ohm resistor at each harmonic frequency for a given $f(t)$ and, when all terms are added, the total average power.

For a periodic signal, we can use Eq. (4.18) to plot a line power spectrum, as shown in Fig. 4.4(a). The corresponding cumulative power spectrum $G_f(\omega)$ is found simply by summing the terms in Eq. (4.18) over all harmonic numbers (n) up to the given frequency ω. Because the inclusion of each harmonic frequency adds a discrete amount of power, this $G_f(\omega)$ will be a series of step functions forming a staircase-type graph, as shown in Fig. 4.4(b). (By definition, the cumulative power spectrum is always a nondecreasing function of ω because power cannot be a negative quantity.)

Writing $G_f(\omega)$ in equation form, we have

$$G_f(\omega) = \sum_{n=-\infty}^{\omega/\omega_0} |F_n|^2 u(\omega - n\omega_0). \tag{4.19}$$

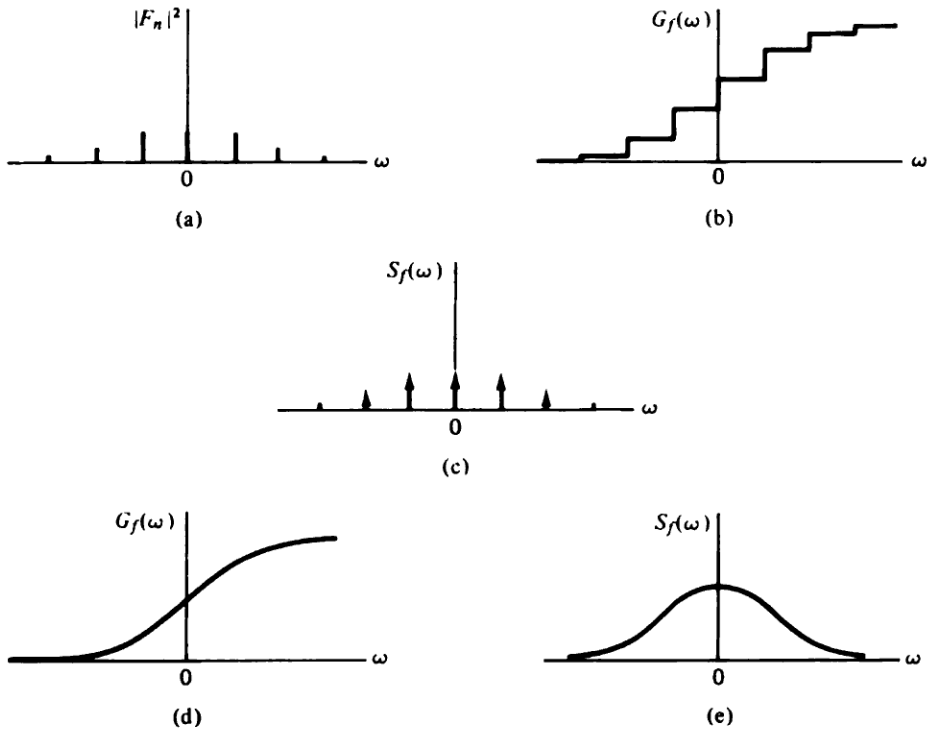


Figure 4.4 Power spectra of periodic functions: (a) line power spectrum of a periodic function; (b) integrated power spectrum of a periodic function; (c) power spectral density of a periodic function; (d) integrated power spectrum of an aperiodic function; and (e) power spectral density of an aperiodic function.

According to our definition that the derivative of a step function is an impulse function, Eqs. (4.16) and (4.19) now yield the desired result

$$S_f(\omega) = 2\pi \sum_{n=-\infty}^{\infty} |F_n|^2 \delta(\omega - n\omega_0) . \tag{4.20}$$

Therefore the power spectral density of a periodic function is a series of impulse functions with weights (areas) corresponding to the magnitude squared of the respective Fourier series coefficients. This is shown in Fig. 4.4(c).

Note that for aperiodic power signals, $G_f(\omega)$ will be a smoothly varying function of frequency and the power spectral density $S_f(\omega)$ follows directly from an application of Eq. (4.16). This case is illustrated in Fig. 4.4(d), (e). We conclude that if a given signal $f(t)$ contains a periodic component of frequency ω_c, then $G_f(\omega)$ will have a step discontinuity at ω_c and $S_f(\omega)$ will contain an impulse at ω_c. This assumes an infinite observation time, as noted earlier.

Now that we have obtained our result, we can formally convert any line power spectrum to a power spectral density simply by changing the lines to impulses. The weights (areas) of these impulses are equal to the squared magnitudes of the line

heights and multiplied by 2π if in radian frequency. Integrating this power spectral density over all frequencies yields

$$P = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_f(\omega)\,d\omega\,,$$

which, for a 1-ohm resistor, gives

$$\overline{f^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} 2\pi \sum_{n=-\infty}^{\infty} |F_n|^2\,\delta(\omega - n\omega_0)\,d\omega = \sum_{n=-\infty}^{\infty} |F_n|^2.$$

This result, of course, checks with Parseval's theorem.

EXAMPLE 4.2.1

Find the power spectral density of the periodic signal

$$f(t) = A\cos(\omega_0 t + \theta)\,.$$

Solution

$$f(t) = \frac{A}{2}e^{j\theta}e^{j\omega_0 t} + \frac{A}{2}e^{-j\theta}e^{-j\omega_0 t}$$

Writing an exponential Fourier series for $f(t)$ (or using Euler's identities), we find

$$F_{-1} = \left(\frac{A}{2}\right)\exp(-j\theta); \qquad F_1 = \left(\frac{A}{2}\right)\exp(j\theta)\,.$$

Using Eq. (4.20), we have

$$S_f(\omega) = \tfrac{1}{2}\pi A^2\,\delta(\omega + \omega_0) + \tfrac{1}{2}\pi A^2\,\delta(\omega - \omega_0)\,.$$

Note that the mean power across a 1-ohm load can be found from

$$\begin{aligned}\overline{f^2(t)} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_f(\omega)\,d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{2}\pi A^2\,\delta(\omega + \omega_0)\,d\omega + \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{2}\pi A^2\,\delta(\omega - \omega_0)\,d\omega \\ &= \tfrac{1}{4}A^2 + \tfrac{1}{4}A^2 = \tfrac{1}{2}A^2.\end{aligned}$$

This result can be checked easily in the time domain. Also note that the phase information is lost in the calculation.

EXAMPLE 4.2.2

Develop an expression for the power spectral density of $f(t) = A\exp(j\omega_0 t)\,\mathrm{rect}\,(t/T)$, both for finite T and as $T \to \infty$.

Solution From Table 3.1, we have

$$\mathcal{F}\{A \text{ rect } (t/T)\} = AT \text{ Sa } (\omega T/2).$$

Using the frequency translation property of the Fourier transform, we have

$$\mathcal{F}\{A \exp (j\omega_0 t) \text{ rect } (t/T)\} = AT \text{ Sa } [(\omega - \omega_0)T/2].$$

Using Eq. (4.17) for a finite observation interval T, we have

$$S_f(\omega) \approx \frac{|F_T(\omega)|^2}{T} = |A|^2 T \text{ Sa}^2 [(\omega - \omega_0)T/2].$$

As the observation interval is made increasingly longer, the $[\sin x/x]^2$ pattern becomes more and more concentrated about $\omega = \omega_0$ and in the limit [cf. Eq. (2.110)],

$$\begin{aligned}\lim_{T\to\infty} \frac{|F_T(\omega)|^2}{T} &= |A|^2 \lim_{T\to\infty} \{T \text{ Sa}^2 [(\omega - \omega_0)T/2]\} \\ &= 2\pi|A|^2 \lim_{T\to\infty} \left\{\frac{T}{2\pi} \text{ Sa}^2 [(\omega - \omega_0)T/2]\right\} \\ &= 2\pi|A|^2 \delta(\omega - \omega_0).\end{aligned}$$

This result agrees with that obtained by the Fourier series expression in Eq. (4.20). (If we were to derive Eq. (4.20) in this manner, however, we would have to use a sum of exponentials and then show that the cross-products vanish as $T \to \infty$.)

EXAMPLE 4.2.3

The power spectral density of a certain signal is found using Eq. (4.17). Several recordings each of length T seconds are taken and the results averaged. The resulting graph is shown in Fig. 4.5. Determine what periodic components, if any, are contained in the signal. Also estimate the length of record taken.

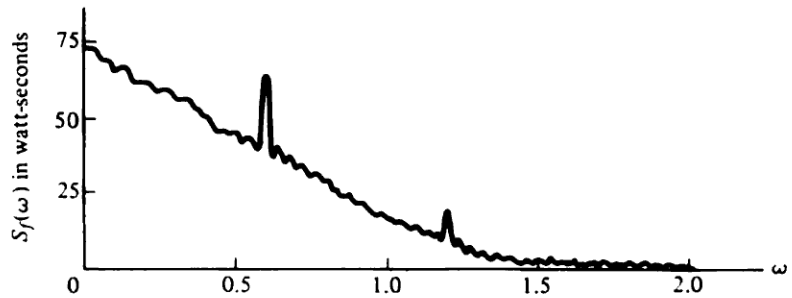


Figure 4.5 Signal power spectral density of Example 4.2.3.

Solution For an infinite observation time, the given signal will have a periodic component at any frequency for which its power spectral density has an impulse function. How-

ever, because the observation time here is finite, the periodic component will show up in the power spectral density in the form

$$T|F_n|^2 \, \mathrm{Sa}^2 \, [(\omega - n\omega_0)T/2]$$

(see previous example). Therefore we can recognize periodic components at the frequencies

$$\omega_1 = 0.6 \text{ rad/sec } (f_1 = 0.0955 \text{ Hz}),$$

$$\omega_2 = 1.2 \text{ rad/sec } (f_2 = 0.191 \text{ Hz}).$$

Other small periodic components may be present but are masked in the fluctuations or noise in the measurement. From Fig. 4.5, the widths of these peaks are about 0.05 rad/sec so that the observation time is

$$T = \frac{4\pi}{0.05} \cong 250 \text{ sec}.$$

Using Fig. 4.5 and the preceding equation for the (known) form of the power spectral density, the amplitudes of the periodic components are (across a 1-ohm load)

$$2T|F_1|^2 = 25 \text{ W-sec}, \quad \text{or} \quad |F_1| = 0.22 \text{ V},$$

$$2T|F_2|^2 = 10 \text{ W-sec}, \quad \text{or} \quad |F_2| = 0.14 \text{ V}.$$

[The factor of 2 was introduced here because $S_f(\omega)$ was displayed as a one-sided spectrum.]

The transmission of power spectra through linear systems follows the same type of reasoning that was used for energy density. Suppose we apply our truncated function to a linear time-invariant filter whose frequency transfer function is $H(\omega)$. The truncated response function, $G_T(\omega)$, is

$$G_T(\omega) \approx F_T(\omega)H(\omega).$$

The power spectral density of the output signal is then

$$S_g(\omega) = \lim_{T\to\infty} \frac{|F_T(\omega)H(\omega)|^2}{T}$$

$$= \lim_{T\to\infty} \frac{|F_T(\omega)|^2}{T} |H(\omega)|^2,$$

$$S_g(\omega) = S_f(\omega)\,|H(\omega)|^2. \tag{4.21}$$

Therefore the power spectral density of the output signal is the power spectral density of the input signal modified by the squared magnitude of the system transfer

function. Again, note that all phase information in both the input signal and the system transfer function is lost. The average power in the output signal is given by

$$P_g = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega) \, |H(\omega)|^2 \, d\omega \, . \tag{4.22}$$

Equations (4.21) and (4.22) give us some insight into why the magnitude transfer function is such a popular way to rate power amplifiers. High-fidelity audio amplifiers, for example, are rated on the basis of power-response curves—graphs of log $|H(\omega)|^2$ versus log ω.

Power spectral density has a physical interpretation very similar to that of energy spectral density. This resemblance can be based on a comparison of Eqs. (4.22) and (4.6). The MKS units of power spectral density are watts per Hz.

DRILL PROBLEM 4.2.1

Derive a relation between the power spectral density of the input, $S_f(\omega)$, and the output, $S_g(\omega)$, of an ideal differentiator.

ANSWER: $S_g(\omega) = \omega^2 S_f(\omega)$.

4.3 TIME-AVERAGED NOISE REPRESENTATIONS

The concept of power spectra allows us to handle some of the averaged effects of the random fluctuations that are present in physical processes. These fluctuations in voltage or current tend to obscure and mask the desired signals and are commonly called noise. In a general sense, noise consists of any unwanted signals, random or deterministic, which interfere with the faithful reproduction of a desired signal in a system. These unwanted signals arise from a variety of sources and can be classified as man-made or naturally occurring.

Man-made types of interference (noise) include such things as electromagnetic pick-up of other radiating signals, inadequate power supply filtering, alias terms arising from poor sampling choices, mechanical vibrations resulting in electrical disturbances, etc. Man-made sources of noise all have the common property that their effects can be eliminated or at least minimized by careful engineering design and practice. We shall not concern ourselves here with these.

Interference caused by naturally occurring disturbances are not controllable in such a direct way and their characteristics can be best described statistically. For some, the fluctuations are very erratic (e.g., those arising from nearby electrical storms) and a valid analytical description is difficult. In others, the time-averaged power may remain quite constant and a treatment on an average power basis is possible and quite meaningful. The concept of a power spectral density is useful in the latter and enables us to treat the effects of noise on an average power basis.

In forming averages of any signal (random or nonrandom), we find parameters that tell us something about the signal. Much of the detailed information about the signal, of course, is lost in the process. In the case of random noise, however, "something is better than nothing" and we will be content with averaged quantities.

Suppose $n(t)$ is a noise voltage or current (again, assume a 1-ohm resistive load). A typical waveform is illustrated in Fig. 4.6(a). We now define the following averages of $n(t)$.

1. *Mean value,* $\overline{n(t)}$:

$$\overline{n(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} n(t)\,dt. \tag{4.23}$$

The parameter $\overline{n(t)}$ is often referred to as the dc, or average, value of $n(t)$. Typically the time interval T in Eq. (4.23) can be finite for a good estimate of $\overline{n(t)}$ if it is large enough to smooth the fluctuations of $n(t)$ adequately.

2. *Mean-square value,* $\overline{n^2(t)}$:

$$\overline{n^2(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |n(t)|^2\,dt. \tag{4.24}$$

The square root of $\overline{n^2(t)}$ is called the root-mean-square (rms) value of $n(t)$. The advantage of the rms notation is that the units of $\sqrt{\overline{n^2(t)}}$ are the same as those of $n(t)$. Aside from a resistance scaling factor, Eq. (4.24) gives the time-averaged power of $n(t)$. From the preceding section, we found it is also related to the integral of the power spectral density, $S_n(\omega)$.

3. *AC component,* $\sigma(t)$:

$$\sigma(t) \triangleq n(t) - \overline{n(t)}. \tag{4.25}$$

The ac, or fluctuation, component of $n(t)$ is that component which remains after the mean value, $\overline{n(t)}$, has been taken out; a typical waveform is shown in Fig. 4.6(b).

Using Eq. (4.25) in Eq. (4.24), we get

$$\overline{n^2(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |\overline{n(t)} + \sigma(t)|^2\,dt,$$

$$\overline{n^2(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |\overline{n(t)}|^2\,dt + \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |\sigma(t)|^2\,dt. \tag{4.26}$$

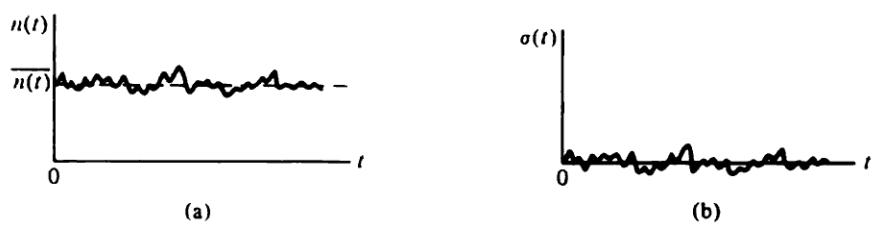


Figure 4.6 (a) A random noise waveform and (b) its ac component.

In Eq. (4.26) we have used the fact that $\overline{n(t)}$ is a constant and the mean of $\sigma(t)$ is zero by definition. The term on the left-hand side of Eq. (4.26) is the time-averaged power in $n(t)$ across a 1-ohm resistor. Likewise the first term on the right-hand side of Eq. (4.26) is the dc power and the second term, the ac power in $n(t)$. An inspection of Eq. (4.26) reveals that the rms value of $n(t)$ is equal to the rms value of $\sigma(t)$ only if the mean value $\overline{n(t)}$ is zero.

The foregoing definitions can be applied to any signal whether random or nonrandom. For periodic signals, T is taken as the period and the limiting operation is not necessary.

EXAMPLE 4.3.1

Calculate the (a) average value, (b) ac power, and (c) rms value of the periodic waveform $v(t) = 1 + \cos \omega_0 t$.

Solution Because $v(t)$ is periodic, we can integrate over one period rather than taking the limit.

$$\text{a)}\quad \overline{v(t)} = \frac{1}{T}\int_{-T/2}^{T/2} (1 + \cos \omega_0 t)\, dt = 1\,,$$

$$\text{b)}\quad \overline{\sigma^2(t)} = \frac{1}{T}\int_{-T/2}^{T/2} (\cos \omega_0 t)^2\, dt = \tfrac{1}{2}\,,$$

$$\text{c)}\quad \overline{v^2(t)} = \frac{1}{T}\int_{-T/2}^{T/2} (1 + \cos \omega_0 t)^2\, dt$$

$$= \frac{1}{T}\int_{-T/2}^{T/2} (1 + 2\cos \omega_0 t + \cos^2 \omega_0 t)\, dt = \tfrac{3}{2}\,,$$

$$v_{\text{rms}} = \sqrt{\overline{v^2(t)}} = \sqrt{3/2}\,.$$

An important measure of the performance of systems, particularly those involving the amplification of low-level signals, is how little noise is introduced in the system. Because the noise varies in an unpredictable manner from one point in time to the next, however, we prefer to analyze this on the basis of average noise power. It is convenient, in fact, to speak of a dimensionless ratio of signal power to noise power. This ratio, called the *signal-to-noise ratio*, can be formed by taking the ratio of the mean-square signal to the mean-square noise because the resistance factor drops. Designating the signal-to-noise ratio by S/N, we have

$$S/N = \overline{s^2(t)}/\overline{n^2(t)}\,. \tag{4.27}$$

It is quite common to express the signal-to-noise ratio in decibels[†]

$$[S/N]_{\text{dB}} = 10 \log_{10}[\overline{s^2(t)}/\overline{n^2(t)}]\,. \tag{4.28}$$

[†] Appendix B has a brief discussion of the decibel.

Both $\overline{s^2(t)}$ and $\overline{n^2(t)}$ are assumed to be measured at the same point. Because a dc level can be added arbitrarily, both of these mean-square values assume zero mean value unless stated otherwise. Thus the power spectral density of the signal and/or the noise may contain an impulse function at $\omega = 0$ which is usually neglected in the calculation of S/N.

DRILL PROBLEM 4.3.1

A given low-level amplifier has a bandwidth 0.01–10 Hz and a noise power spectral density of $S_n(\omega) = 10^{-14}\,|\omega|^{-1}$ W/Hz referred to its input. (Assume 1 ohm.) Calculate the signal-to-noise ratio of the amplifier output for a 1-μV rms sinusoidal input at 1 Hz. Assume that the amplifier gain is a constant over the bandwidth given.

ANSWER: 16.6 dB.

4.4 CORRELATION FUNCTIONS

In a previous section we have seen how signals can be handled using the power spectral-density function, $S_f(\omega)$. The question now naturally arises: Is there some operation in the time domain which is equivalent to finding the power spectral density in frequency?

To answer this important question, let us begin by assuming that our definition of power spectral density is satisfied, that is [cf. Eq. (4.17)],

$$S_f(\omega) = \lim_{T\to\infty} \frac{1}{T}\,|F_T(\omega)|^2. \tag{4.29}$$

The corresponding operation in time will be the inverse Fourier transform of Eq. (4.29). Writing this out gives us

$$\mathcal{F}^{-1}\{S_f(\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{1}{T}\,|F_T(\omega)|^2 e^{j\omega\tau}\,d\omega\,. \tag{4.30}$$

We have purposely chosen a new time variable, τ, in Eq. (4.30) because the time variable t is already in use in the definition of $F_T(\omega)$. Interchanging the order of operations yields

$$\begin{aligned}
\mathcal{F}^{-1}\{S_f(\omega)\} &= \lim_{T\to\infty} \frac{1}{2\pi T}\int_{-\infty}^{\infty} F_T^*(\omega)F_T(\omega)e^{j\omega\tau}\,d\omega \\
&= \lim_{T\to\infty} \frac{1}{2\pi T}\int_{-\infty}^{\infty}\int_{-T/2}^{T/2} f^*(t)e^{j\omega t}\,dt\int_{-T/2}^{T/2} f(t_1)e^{-j\omega t_1}\,dt_1\,e^{j\omega\tau}\,d\omega \\
&= \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} f^*(t)\int_{-T/2}^{T/2} f(t_1)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega(t-t_1+\tau)}\,d\omega\right]dt_1\,dt\,.
\end{aligned} \tag{4.31}$$

The integration over ω within the brackets in Eq. (4.31) is now recognized as [cf. Eq. (3.25)] $\delta(t - t_1 + \tau)$, so that

$$\mathcal{F}^{-1}\{S_f(\omega)\} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t) \int_{-T/2}^{T/2} f(t_1)\,\delta(t - t_1 + \tau)\,dt_1\,dt$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t) f(t + \tau)\,dt\,. \tag{4.32}$$

Equation (4.32) describes the operations in the time domain that correspond to the determination of $S_f(\omega)$ in frequency. The inverse Fourier transform of $S_f(\omega)$ is called the *autocorrelation function* of $f(t)$, designated by $R_f(\tau)$. Then our desired result can be written as[†]

$$R_f(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t) f(t + \tau)\,dt\,. \tag{4.33}$$

Also, taking the Fourier transform of both sides of Eq. (4.32) and using Eq. (4.33), we can write

$$S_f(\omega) = \mathcal{F}\{R_f(\tau)\}\,. \tag{4.34}$$

We now have another method to find the power spectral-density function, i.e., first determine the autocorrelation function and then take a Fourier transform. This method is applicable to both random and deterministic signals. We have assumed that the power and the power spectral density do not vary with time. If this assumption is not valid, one can still find an autocorrelation function but it will depend on both τ and t.

For real-valued signals, a computation of the autocorrelation function $R_f(\tau)$ turns out to be similar to that of taking the convolution of $f(-t)$ with $f(t)$. Because these operations have been discussed in Chapter 3, we only illustrate with examples.

EXAMPLE 4.4.1

Determine and sketch the autocorrelation function of a periodic square wave with peak-to-peak amplitude A, period T, and mean value $A/2$.

Solution Because $f(t)$ is periodic, the limiting operation in the determination of $R_f(\tau)$ can be replaced by a computation over one period. Using an alternate form of Eq. (4.33) [see footnote to that equation], we have

For $-T/2 < \tau < 0$:

$$R_f(\tau) = \frac{1}{T} \int_{-T/4}^{(T/4)+\tau} A^2\,dt = A^2\left(\frac{1}{2} + \frac{\tau}{T}\right),$$

[†] With a change of variable, this can also be written as

$$R_f(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t - \tau) f(t)\,dt\,.$$

For $0 < \tau < T/2$:

$$R_f(\tau) = \frac{1}{T}\int_{\tau-T/4}^{T/4} A^2\,dt = A^2\left(\frac{1}{2} - \frac{\tau}{T}\right).$$

A graph of $f(t)$ and $R_f(\tau)$ is shown in Fig. 4.7. Since $f(t + T) = f(t)$, all calculations repeat over every period. It follows that the autocorrelation function of a periodic waveform is periodic. Similarly, the autocorrelation function of an aperiodic waveform is aperiodic.

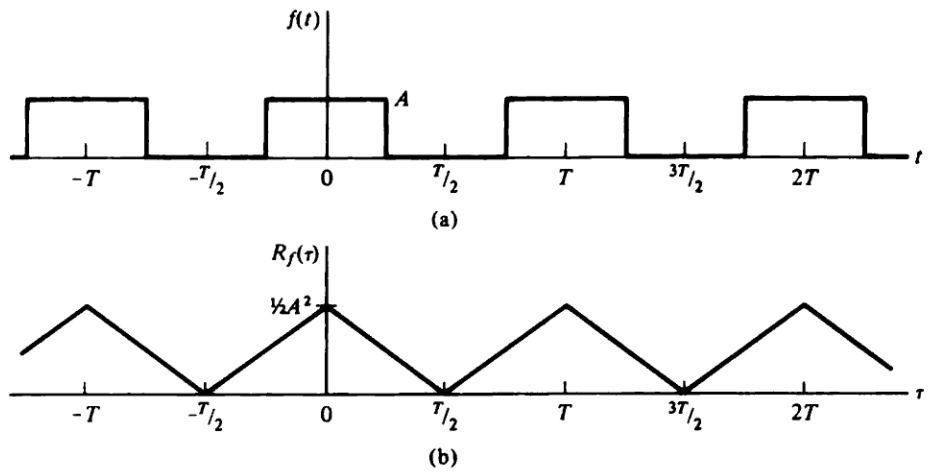


Figure 4.7 (a) A periodic waveform and (b) its autocorrelation function.

EXAMPLE 4.4.2

Find the autocorrelation function of $\sqrt{2}\cos(\omega_0 t + \theta)$.

Solution

$$R_f(\tau) = \frac{1}{T}\int_{-T/2}^{T/2} 2\cos(\omega_0 t + \theta)\cos(\omega_0 t + \omega_0\tau + \theta)\,dt,$$

$$R_f(\tau) = \frac{1}{T}\int_{-T/2}^{T/2} \cos\omega_0\tau\,dt + \frac{1}{T}\int_{-T/2}^{T/2} \cos(2\omega_0 t + \omega_0\tau + 2\theta)\,dt,$$

$$R_f(\tau) = \cos\omega_0\tau.$$

Note that the autocorrelation function is independent of the phase θ.

The autocorrelation function is widely used in signal analysis. It is especially useful for the detection or recognition of signals that are masked by additive noise. For example, consider a periodic square wave such as that shown in Fig. 4.8(a). Its autocorrelation function, as developed in Example 4.4.1, is shown in Fig. 4.8(b). A band-limited white random-noise waveform is shown in Fig. 4.8(c) and its autocorrelation function is shown in Fig. 4.8(d). The random-noise waveform is added to the periodic

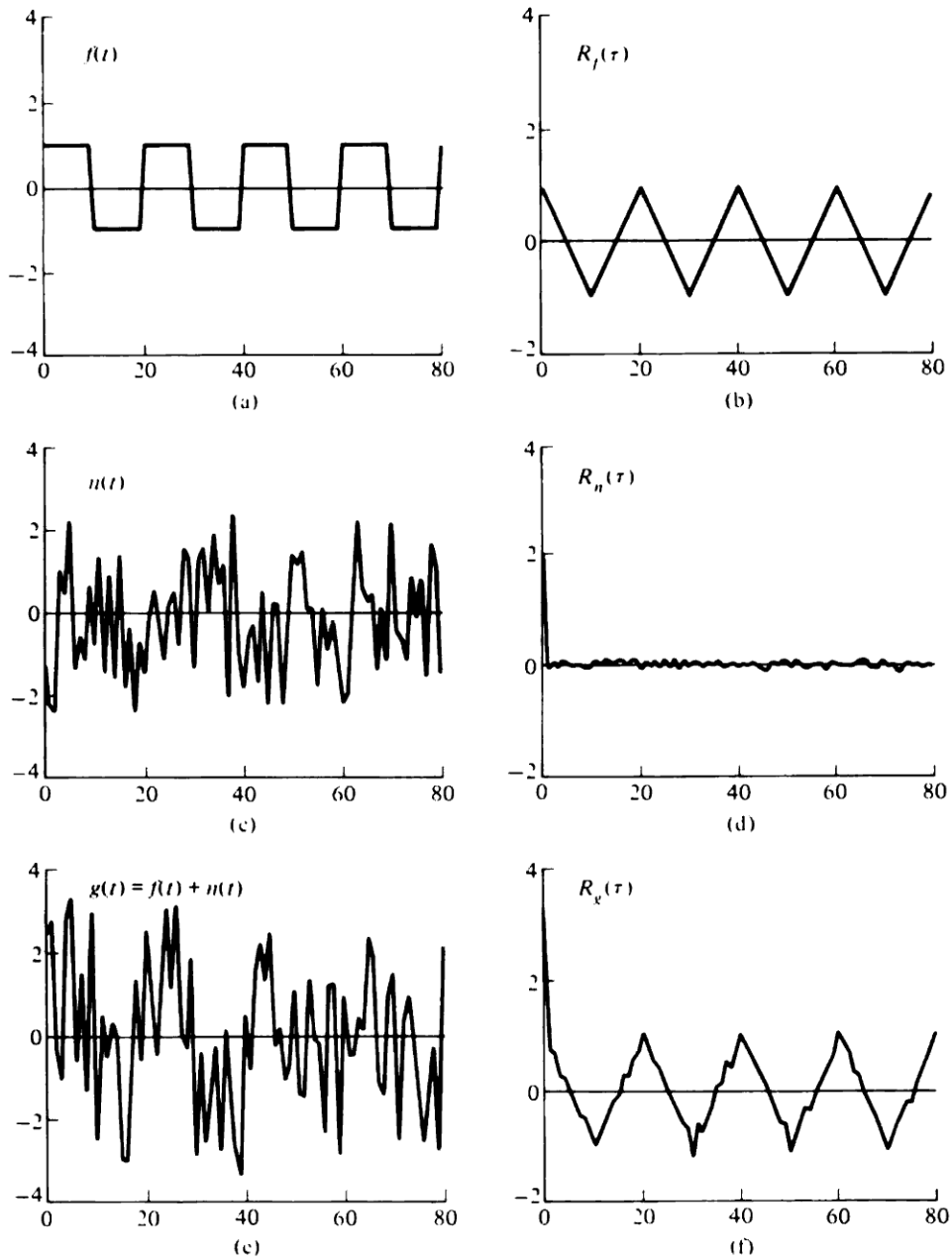


Figure 4.8 Autocorrelation of a periodic signal plus noise.

square wave and the resulting waveform is shown in Fig. 4.8(e) and its autocorrelation function is shown in Fig. 4.8(f). Note that, even though the square wave is immersed in a considerable amount of noise, the autocorrelation function of the square wave is clearly recognizable in the final result.

One of the drawbacks in using autocorrelation for this type of application is that the autocorrelation of the noise is present in the output as well as the autocorrelation of the signal (e.g., compare Figs. 4.8d and f). This may make detection of aperiodic

signals difficult. Also, the relative time shift (phase) between signals is lost (cf. Example 4.4.2). These drawbacks can be avoided by use of a closely related operation known as *cross-correlation*.

Suppose that we have two waveforms $f(t)$ and $g(t)$. The cross-correlation function $R_{fg}(\tau)$ is defined as

$$R_{fg}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t)g(t+\tau)\,dt. \tag{4.35}$$

As an example of an application of cross-correlation, we choose a random waveform $f(t)$ as shown in Fig. 4.9(a). (Its autocorrelation function, $R_f(\tau)$, will be similar to that shown in Fig. 4.8d.) For the second function, $g(t)$, we choose a delayed replica of $f(t)$ plus a second random waveform $n(t)$ so that $g(t) = f(t - t_0) + n(t)$. The composite waveform $g(t)$ is shown in Fig. 4.9(b). It is presumed that the receiver has a replica of the waveform $f(t)$ available (e.g., in memory). On the basis of this

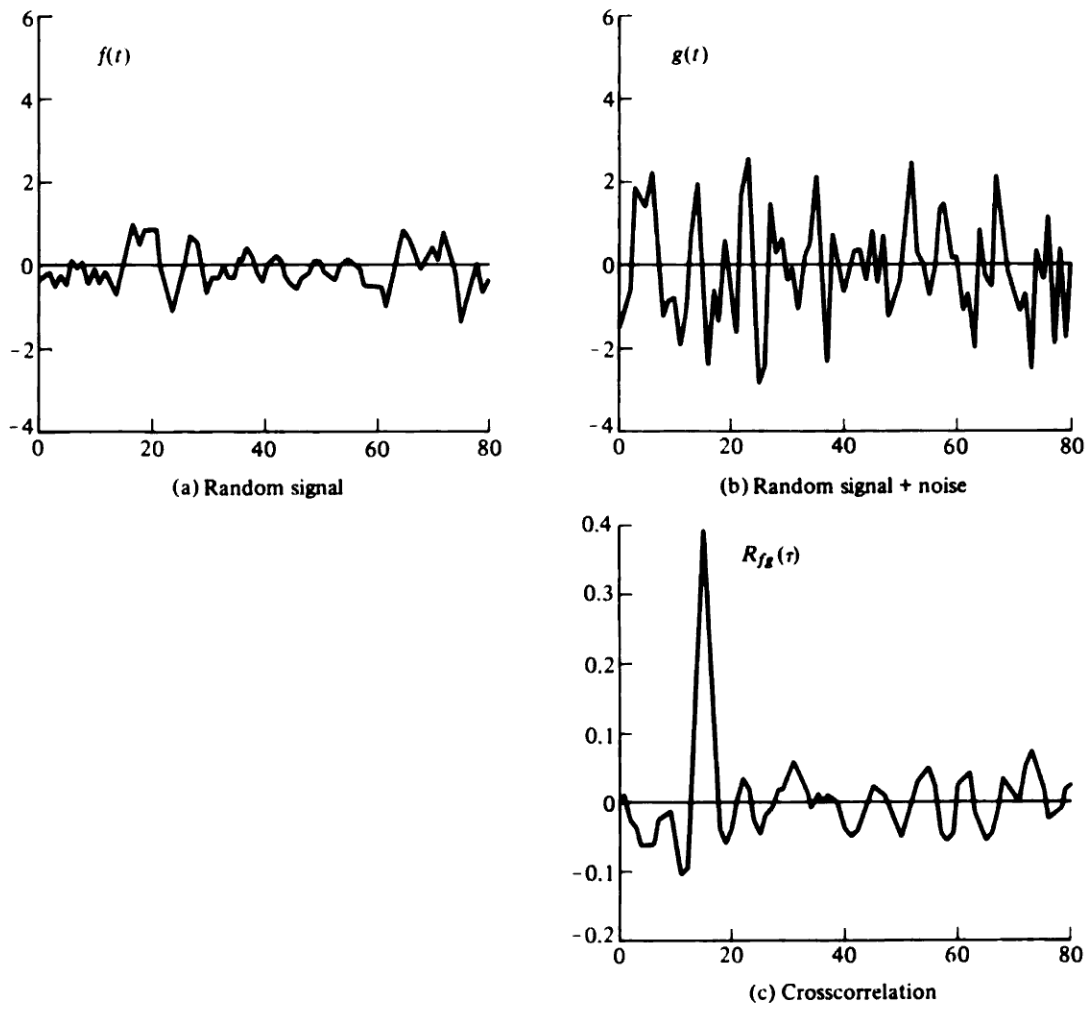


Figure 4.9 Cross-correlation of a random signal plus noise.

knowledge of $f(t)$, we wish to have the receiver make a measurement of the time delay t_0. To do this, we take the cross-correlation function $R_{fg}(\tau)$, as defined in Eq. (4.35). The result is shown in Fig. 4.9(c). The value of the time delay t_0 is clearly evident by measuring the time delay from the origin to the large peak in the result.

Correlation functions furnish measures of the similarity of a signal $f(t)$ either with itself (in the case of autocorrelation) or with another signal (in the case of cross-correlation) versus a relative shift by an amount τ. For dissimilar signals, the peak of the correlation function is an indicator of how good this match is between signals.

Both autocorrelation and cross-correlation are powerful tools in signal analysis in analytical and practical work. We shall have occasion to use them in future chapters.

4.5 SOME PROPERTIES OF CORRELATION FUNCTIONS

We have already found, in the preceding section, that the Fourier transform of the autocorrelation function gives the power spectral density. We shall examine briefly several additional properties of correlation functions.

4.5.1 Symmetry

Examining the autocorrelation function for negative arguments, we have

$$R_f(-\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t) f(t-\tau)\, dt\,,$$

$$R_f(-\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(\zeta+\tau) f(\zeta)\, d\zeta\,,$$

$$R_f(-\tau) = R_f^*(\tau)\,. \tag{4.36}$$

Therefore the real part of $R_f(\tau)$ is an even function; and if $f(t)$ is real-valued, then $S_f(-\omega) = S_f^*(\omega)$ [cf. Eqs. (3.38) and (4.34)].

4.5.2 Mean-Square Value

The autocorrelation function $R_f(\tau)$ evaluated at $\tau = 0$ is equal to the mean-square value of the signal $f(t)$ [cf. Eq. (4.24)],

$$R_f(0) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^*(t) f(t)\, dt\,,$$

$$R_f(0) = \overline{f^2(t)}\,. \tag{4.37}$$

It is left for the reader to verify that Eq. (4.37) agrees with Eqs. (4.10) and (4.34) when referenced to 1 ohm.

4.5.3 Periodicity

If $f(t + T) = f(t)$ for all t, then

$$R_f(\tau + T) = R_f(\tau) \quad \text{for all } \tau. \tag{4.38}$$

Proof follows easily from writing out the integrals and using the definition of periodicity.

4.5.4 ★ Average Value

Let $f(t)$ be represented by a function $x(t)$ with zero average value, plus an average value designated by m_1. Similarly, we represent $g(t)$ as a function $y(t)$ with zero average value, plus an average value designated m_2. In equation form, this can be written as

$$f(t) = x(t) + m_1,$$
$$g(t) = y(t) + m_2.$$

The cross-correlation of $f(t)$, $g(t)$ is

$$R_{fg}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [x^*(t) + m_1][y(t + \tau) + m_2]\,dt.$$

Noting that the average values of $x(t)$ and $y(t)$ are zero by definition, we have

$$R_{fg}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x^*(t)y(t + \tau)\,dt + m_1m_2.$$

The average value of the cross-correlation function is

$$\overline{R_{fg}(\tau)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x^*(t)y(t + \tau)\,dt\,d\tau + m_1m_2.$$

Interchanging the order of integration, we have

$$\overline{R_{fg}(\tau)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x^*(t) \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} y(t + \tau)\,d\tau\,dt + m_1m_2.$$

Because $\overline{y(t + \tau)}$ is zero, we obtain the result

$$\overline{R_{fg}(\tau)} = m_1m_2. \tag{4.39}$$

Therefore the average value of the cross-correlation of two functions $f(t)$ and $g(t)$ is equal to the product of their average values. If the average value of either function is zero, then the average value of their cross-correlation is zero. The case for autocorrelation follows readily from this result.

DRILL PROBLEM 4.5.1

Making use of Fig. 4.7, sketch the autocorrelation function of a symmetric square wave (with zero average value).

ANSWER: See Fig. 6.22(b).

4.5.5 ★ Maximum Value

We can show that $R_x(0) \geq R_x(\tau)$ for any τ [see, e.g., Fig. 4.7] by taking the magnitude squared of the autocorrelation function and using the Schwarz inequality (cf. Example 2.5.2). Thus we have

$$\left|\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} f^*(t)f(t+\tau)\,dt\right|^2 \leq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} |f(t)|^2\,dt \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} |f(t+\tau)|^2\,dt$$
$$\leq R_f(0)R_f(0)\,.$$

Taking the square root of both sides of this result, we have

$$|R_f(\tau)| \leq R_f(0)\,. \tag{4.40}$$

Therefore the autocorrelation function $R_f(\tau)$ is bounded by the mean-square value of the signal $f(t)$. For a periodic signal, the equality in Eq. (4.40) is valid at multiples of the period from the origin (see Fig. 4.7). For nonperiodic $f(t)$, $R_f(\tau)$ is strictly less than $R_f(0)$ for all $\tau \neq 0$.

4.5.6 Additivity

If two signals are added, the autocorrelation function of the sum may or may not be the sum of their respective autocorrelation functions. To investigate this, we write $z(t) = x(t) + y(t)$. The autocorrelation function for the sum of the two signals $x(t)$ and $y(t)$ is

$$R_z(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} [x^*(t) + y^*(t)][x(t+\tau) + y(t+\tau)]\,dt\,,$$

$$R_z(\tau) = R_x(\tau) + R_y(\tau) + R_{xy}(\tau) + R_{yx}(\tau)\,. \tag{4.41}$$

We conclude that only if the cross-correlation functions are zero [i.e., if $R_{xy}(\tau) = R_{yx}(\tau) = 0$] can we write

$$R_z(\tau) = R_x(\tau) + R_y(\tau)\,. \tag{4.42}$$

For the condition $R_{xy}(\tau) = 0$ for all τ, we say that $x(t)$ and $y(t)$ are *uncorrelated*. Moreover, we can show that $R_{yx}(\tau) = R_{xy}^*(-\tau)$, so that if $R_{xy}(\tau) = 0$, then $R_{yx}(\tau) = 0$.

Note that if $x(t)$ and $y(t)$ are orthogonal, then they are uncorrelated. In addition, as will be discussed in Chapter 8, if $x(t)$ and $y(t)$ are statistically independent they are also uncorrelated.

Because power spectral density is the Fourier transform of the autocorrelation function, our conclusion here is that if two signals $x(t)$ and $y(t)$ are uncorrelated then their power spectral densities are additive. In other words, the average power of the sum of two signals is the sum of the average powers of the two signals only if the signals are uncorrelated. In cases in which the cross-correlation functions are not zero, the signals must be added first and then the average power may be determined, or, equivalently, the cross-correlations in Eq. (4.41) must be included.

DRILL PROBLEM 4.5.2

A sinusoidal waveform, $3\sqrt{2}\cos\omega_1 t$, is added to a second, $4\sqrt{2}\cos\omega_2 t$. Determine the rms value of the sum if (a) $\omega_1 = \omega_2$; (b) $\omega_1 \neq \omega_2$.

ANSWER: (a) 7; (b) 5.

4.6 CORRELATION FUNCTIONS FOR FINITE-ENERGY SIGNALS

The concept of correlation can be extended to signals of finite energy. Specifically, we define the autocorrelation function $r_f(\tau)$ for a signal $f(t)$ of finite energy as

$$r_f(\tau) = \int_{-\infty}^{\infty} f^*(t)f(t+\tau)\,dt. \tag{4.43}$$

Similarly, for signals $f(t)$ and $g(t)$, both of finite energy, we define the cross-correlation function $r_{fg}(\tau)$ as

$$r_{fg}(\tau) = \int_{-\infty}^{\infty} f^*(t)g(t+\tau)\,dt. \tag{4.44}$$

Note that for real-valued functions these operations are the same as for convolution, except that the second function is not reversed.

The Fourier transform of Eq. (4.43) gives

$$\mathcal{F}\{r_f(\tau)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^*(t)f(t+\tau)\,dt\,e^{-j\omega\tau}\,d\tau. \tag{4.45}$$

Interchanging the order of integration in Eq. (4.45), we have

$$\mathcal{F}\{r_f(\tau)\} = \int_{-\infty}^{\infty} f^*(t)\int_{-\infty}^{\infty} f(t+\tau)e^{-j\omega\tau}\,d\tau\,dt. \tag{4.46}$$

Use of the time-delay property of the Fourier transform in the right-hand side of Eq. (4.46) gives

$$\int_{-\infty}^{\infty} f^*(t)F(\omega)e^{j\omega t}\,dt = |F(\omega)|^2. \tag{4.47}$$

Combining Eqs. (4.46) and (4.47), we have

$$\mathcal{F}\{r_f(\tau)\} = |F(\omega)|^2. \tag{4.48}$$

Identifying the right-hand side of Eq. (4.48) as the energy spectral density of $f(t)$, we conclude that the energy spectral density is the Fourier transform of the autocorrelation function for finite-energy signals.

DRILL PROBLEM 4.6.1

Determine the autocorrelation function of the rectangular pulse waveform shown in Fig. 4.10(a).

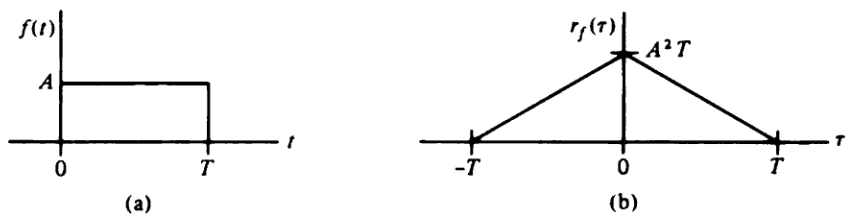


Figure 4.10 (a) A pulse signal and (b) its autocorrelation function.

ANSWER: See Fig. 4.10(b).

4.7 BAND-LIMITED WHITE NOISE

The power spectral-density function plays a central role in our time-averaged description of random noise. The particular type of power spectral density in which we are interested is that one which tends to be a constant for all frequencies. Such a flat power spectrum contains all frequency components with equal power weighting and is designated as white, in an analogy to white light.

If we have a constant power spectral density of η watts per Hz (measured over positive frequencies), and if $n(t)$ has zero mean value, then the power spectral density of white noise is

$$S_n(\omega) = \eta/2 \qquad \text{for all } \omega. \tag{4.49}$$

The factor of one-half in Eq. (4.49) is necessary to have a two-sided power spectral density. Note that we are defining $S_n(\omega)$ on a power basis. In other words, for a resis-

tor of R ohms, one must multiply Eq. (4.49) by R to convert to mean-square voltage and divide by R to convert to mean-square current.

Strictly speaking, Eq. (4.49) cannot be used to describe any physical process because it implies an infinite amount of power; that is,

$$\overline{n^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (\eta/2)\, d\omega \to \infty .$$

However, it turns out to be a good model for many cases in which the bandwidth of the measuring device is narrower than the bandwidth limitations of the physical process being observed. Because our measurements are restricted to finite bandwidths, what we are really interested in, then, is *band-limited white noise*. In other words, if a noise waveform has a flat power spectral density extending beyond the bandwidth of a given system, the noise appears to the system as if it were indeed white.

For band-limited white noise, the noise power is independent of the choice of operating frequency. For example, suppose that $n(t)$ is a zero-mean white noise whose power spectral density is $\eta/2$ watts per Hz. Across a bandwidth B, the noise power P_n is

$$P_n = \frac{1}{2\pi}\int_{-2\pi B}^{2\pi B} (\eta/2)\, d\omega = \eta B \qquad \text{W}. \tag{4.50}$$

Assuming that this is developed across a resistor R, we find the mean-square noise voltage is[†]

$$\overline{n^2(t)} = RP_n = \eta RB \qquad \text{V}^2. \tag{4.51}$$

If $n(t)$ is a current, then

$$\overline{n^2(t)} = P_n/R = \eta GB \qquad \text{amp}^2. \tag{4.52}$$

The transmission of white noise through linear time-invariant systems follows in the manner described for power spectral density. Suppose that we wish to find the rms voltage at the output of a filter whose transfer function $H(\omega)$ is known. Designating the input by $n_i(t)$ and the output by $n_o(t)$, we can write

$$S_{n_o}(\omega) = S_{n_i}(\omega)\,|H(\omega)|^2, \tag{4.53}$$

$$\overline{n_o^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{n_o}(\omega)\, d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{n_i}(\omega)\,|H(\omega)|^2\, d\omega . \tag{4.54}$$

[†] Note that here the average power is given (from the spectral density) and we wish to find the mean-square voltage.

If the spectral density of the noise is white (1-ohm resistor assumed), Eq. (4.54) becomes

$$\overline{n_o^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\eta}{2}|H(\omega)|^2\,d\omega$$

$$= \frac{\eta}{2\pi}\int_{0}^{\infty} |H(\omega)|^2\,d\omega\,. \tag{4.55}$$

An important physical noise source giving rise to band-limited white noise over a very broad frequency range is discussed in the following section.

4.7.1 Thermal Noise

Thermal noise is produced as a result of the thermally excited random motion of free electrons in a conducting medium, such as a resistor. The path of each electron in motion is randomly oriented due to collisions. The net effect of the motion of all electrons is an electric current in the resistor that is random with a mean value of zero. From thermodynamic and quantum mechanical considerations, the power spectral density of thermal noise is given by

$$S_n(\omega) = \frac{h|\omega|}{\pi[\exp\,(h|\omega|/2\pi kT) - 1]}, \tag{4.56}$$

$$S_n(\omega) \cong 2kT \text{ watts per Hz} \quad \text{for} \quad |\omega| \ll 2\pi kT/h\,, \tag{4.57}$$

where

T = temperature of the conducting medium in Kelvin (K),

k = Boltzmann's constant = 1.38×10^{-23} joule/K,

h = Planck's constant = 6.625×10^{-34} joule-sec.

For frequencies above kT/h, thermal noise is no longer white. However, these frequencies are so high for electrical signals that we can safely assume that thermal noise is white for our purposes here (for example, $kT/h \approx 6000$ GHz for $T = 290$ K).

In practice, resistors may actually produce slightly more thermal noise than that indicated by the spectral density above. This increase is a function of materials and geometry and we shall neglect these factors here. Note that an ideal capacitor has no thermal noise source because there are no free electrons present in an ideal dielectric. On the other hand, an ideal inductor has no thermal noise source because an ideal conductor cannot have a lattice structure to impede the flow of electrons.

Using Eq. (4.57) and Eq. (4.51), we find the mean-square (open-circuit) voltage generated by a resistor R in a bandwidth B is

$$\overline{v^2(t)} = RP_n = \frac{1}{2\pi}\int_{-2\pi B}^{2\pi B} 2kTR\,d\omega\,,$$

$$\overline{v^2(t)} = 4kTRB \qquad \text{V}^2. \tag{4.58}$$

The mean-square (short-circuit) current generated, using Eq. (4.57) and Eq. (4.52), is

$$\overline{i^2(t)} = GP_n = \frac{1}{2\pi}\int_{-2\pi B}^{2\pi B} 2kTG\, d\omega\,,$$

$$\overline{i^2(t)} = 4kTGB \qquad \text{amp}^2. \tag{4.59}$$

The voltage and current equivalent circuit models for band-limited thermal noise are shown in Fig. 4.11. The resistance R and the conductance G are assumed noise-free and the bandwidth B is that of the measurement device.

Noise circuit problems involving only resistive components can be solved using these circuit models. The general procedure is to replace each noisy resistor in a series configuration with the equivalent voltage model and each noisy resistor in a parallel configuration with its equivalent current model. Note that voltage and current sources are added on a mean-square basis. The procedure is illustrated in the following example.

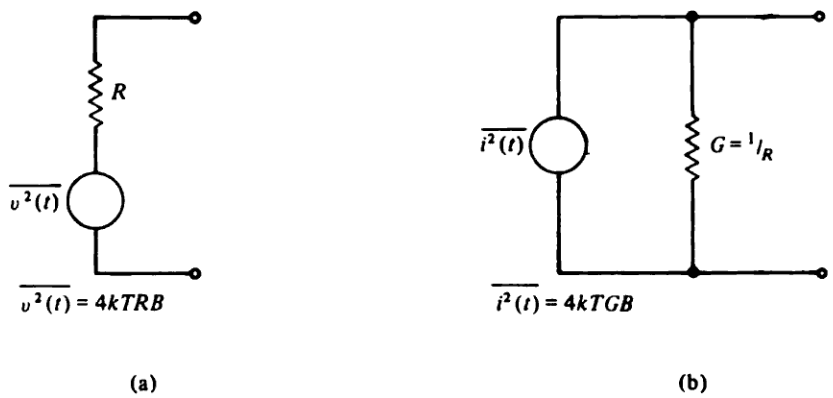


Figure 4.11 Thermal-noise equivalent circuit models: (a) voltage model; (b) current model.

EXAMPLE 4.7.1

Calculate the rms noise voltage arising from thermal noise in two resistors, 100 Ω and 150 Ω respectively, at $T = 300$ K within a bandwidth of 1 MHz if (a) the resistors are connected in series; and (b) the resistors are connected in parallel.

Solution a) Placing the two equivalent circuit models in series and using Eq. (4.58), we get

$$\begin{aligned}\overline{v^2(t)} &= 4kTB(R_1 + R_2)\\ &= 4(1.38 \times 10^{-23})(300)(10^6)(250)\\ &= 4.14 \times 10^{-12}\ \text{V}^2,\\ \sqrt{\overline{v^2(t)}} &= 2.03\ \mu\text{V}.\end{aligned}$$

The equivalent source resistance is 250 Ω.

b) Placing the two in parallel and using Eq. (4.59) gives us

$$\begin{aligned}\overline{i^2(t)} &= 4kTB(G_1 + G_2)\\ &= 4(1.38 \times 10^{-23})(300)(10^6)(0.0167)\\ &= 2.76 \times 10^{-16}\ \text{amp}^2,\\ \sqrt{\overline{i^2(t)}} &= 0.0166\ \mu\text{A}.\end{aligned}$$

The equivalent parallel resistance is

$$R' = \frac{1}{G_1 + G_2} = 60\ \Omega,$$

so that

$$\begin{aligned}\sqrt{\overline{v^2(t)}} &= \sqrt{\overline{i^2(t)}}R'\\ &= (0.0166 \times 10^{-6})(60)\\ &= 0.997\ \mu\text{V}.\end{aligned}$$

DRILL PROBLEM 4.7.1

Repeat Example 4.7.1 with the sole exception that the 100-Ω resistor is at a temperature of 600 K.

ANSWER: (a) 2.41 μV; (b) 1.26 μV.

4.7.2 Transmission of Thermal Noise through Linear Systems

We have been assuming that there is no relation between the random motion of free electrons in different resistors. This means that noise contributions arising from this random motion add on a power basis for any given frequency band. Under this condition we can add the power spectral densities of different noise sources or, if the integration remains the same, we can add the mean-square values of the current or voltage directly. This was done above.

If any energy storage components (i.e., capacitors or inductors) are present, the bandwidth of the system may have more effect on the noise measurement than the measuring device itself. In this case we must be careful to "go back to the fundamentals" and treat the noise source(s) in terms of power spectral density, then integrate. The integration may change drastically as a function of the system.

Suppose, for example, that we connect a resistor to the input terminals of a linear system containing only noiseless components, as shown in Fig. 4.12(a). The proce-

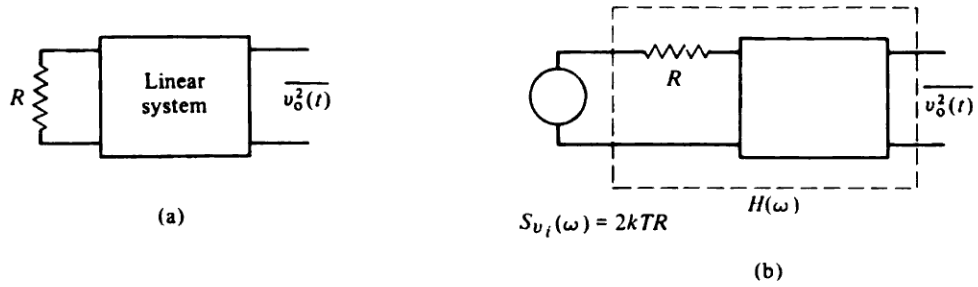


Figure 4.12 Transmission of thermal noise through a linear system: (a) system model; (b) equivalent noise model.

dure is the following. First we replace the input resistance with a noise voltage source and a noise-free resistor R, as shown in Fig. 4.12(b). The noise-free resistor is now included as part of the transfer function of the system. On a voltage basis,

$$S_{v_i}(\omega) = 2kTR \tag{4.60}$$

and

$$S_{v_o}(\omega) = S_{v_i}(\omega)\,|H(\omega)|^2. \tag{4.61}$$

The mean-square output voltage is

$$\overline{v_o^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{v_i}(\omega)\,|H(\omega)|^2\,d\omega\,. \tag{4.62}$$

The analogous relations for current follow readily.

But what if the system itself contains noisy resistive components? If the system, in addition to being linear, is passive and bilateral, then the effective noise resistance referred to the input is†

$$R_{eq}(\omega) = \mathscr{R}e\{Z(\omega)\}\,, \tag{4.63}$$

where $Z(\omega)$ is the complex-valued input impedance of the system. All other calculations follow in a similar way to that discussed above. For example, the noise voltage spectral density would be [cf. Eq. (4.60)]

$$S_v(\omega) = 2kTR_{eq}(\omega)\,. \tag{4.64}$$

Note that in general $R_{eq}(\omega)$ is a function of frequency.

EXAMPLE 4.7.2

Calculate the rms noise voltage developed across a capacitor C when it is connected in parallel with a noisy resistor R.

† For more complicated networks involving active and unilateral devices it is common practice to state the equivalent noise voltage referenced to the input as a part of the measured data.

Solution The circuit and its equivalent noise model are shown in Fig. 4.13. The resistor R is replaced by a noise-free resistor in series with a noise voltage source with power spectral density,

$$S_{v_i}(\omega) = 2kTR \qquad \mathrm{V}^2/\mathrm{Hz}\,.$$

The system transfer function is that of a RC low-pass filter:

$$H(\omega) = \frac{1/RC}{j\omega + 1/RC}\,.$$

The mean-square value of the output is given by Eq. (4.62):

$$\begin{aligned}
\overline{v_o^2(t)} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} 2kTR\frac{(1/RC)^2}{\omega^2 + (1/RC)^2}\,d\omega \\
&= \frac{1}{\pi}\int_{0}^{\infty} 2kTR\frac{(1/RC)^2}{\omega^2 + (1/RC)^2}\,d\omega \\
&= \frac{2kT}{\pi C}\tan^{-1}(\omega RC)\Big|_0^{\infty} \\
&= kT/C\,.
\end{aligned}$$

Then the rms output noise voltage is

$$\sqrt{\overline{v_o^2(t)}} = \sqrt{\frac{kT}{C}}\,.$$

At first glance, it seems strange that this result should be independent of R because the resistance is the source of the noise! A closer look, however, reveals that even though the mean-square noise voltage is proportional to R, the bandwidth of the low-pass filter formed is inversely proportional to R. These two effects cancel, yielding the above result.

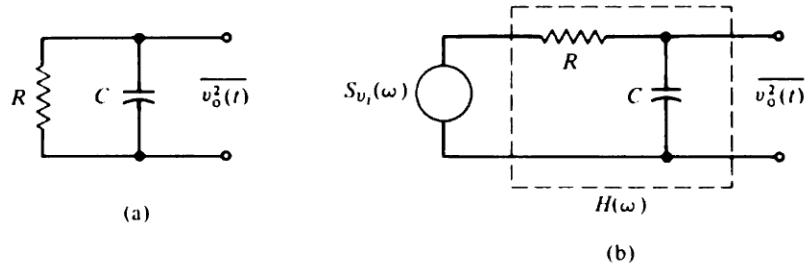


Figure 4.13 (a) The RC network and (b) its equivalent noise model as discussed in Example 4.7.2.

DRILL PROBLEM 4.7.2

Calculate the rms noise current developed through an ideal inductor L when it is connected to a noisy resistor R.

ANSWER: $\sqrt{kT/L}$.

4.7.3 Equivalent Noise Bandwidth

Expressions such as Eqs. (4.58) and (4.59) assume an ideal filter of bandwidth B for the noise measurement. In practice, it is convenient to combine the bandwidth-limiting characteristics of a system by defining an equivalent noise bandwidth. The equivalent noise bandwidth, B_N, is that ideal filter bandwidth which gives the same noise power as the actual system. The equivalent noise bandwidth for white noise can be determined easily as follows.

Within the white noise assumption, the input power spectral density is a constant, $\eta/2$. The mean-square voltage output of a linear system is given by Eq. (4.62). Then the mean-square output voltage, $\overline{v_o^2(t)}$, across a 1-ohm resistor is [cf. Eq. (4.55)]

$$\overline{v_o^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\eta}{2}|H(\omega)|^2\,d\omega$$

$$= \frac{\eta}{2\pi}\int_{0}^{\infty} |H(\omega)|^2\,d\omega\,. \tag{4.65}$$

The definite integral in Eq. (4.65) is a constant for a given system frequency transfer function $H(\omega)$. (We are using the voltage transfer function here because it is the one most commonly used.)

We would like to use a simplified approach (such as was used for resistors) in which the power spectral density is white within some effective bandwidth of the system. We can do this by defining an *equivalent noise bandwidth* B_N such that (1) the power spectral density at the filter output is white within the bandwidth B_N and zero elsewhere, forming an equivalent rectangular spectral density; and (2) the area under this rectangular spectral density is equal to the area under the spectral density at the filter output. This is illustrated in Fig. 4.14.

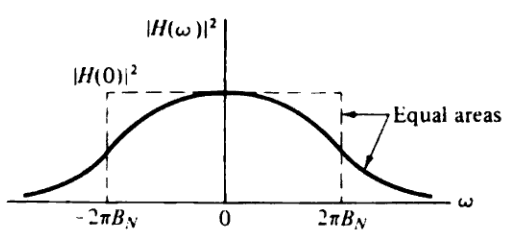


Figure 4.14 A graphic definition of noise equivalent bandwidth.

Designating the midband frequency of the system by ω_0 ($\omega_0 = 0$ for a low-pass system), the midband system voltage gain is $|H(\omega_0)|^2$ and we can write

$$\overline{v_o^2(t)} = \frac{1}{2\pi}\int_{-2\pi B_N}^{2\pi B_N} (\eta/2)\,|H(\omega_0)|^2\,d\omega\,,$$

$$\overline{v_o^2(t)} = \eta|H(\omega_0)|^2 B_N\,. \tag{4.66}$$

Equating the right-hand sides of Eqs. (4.65) and (4.66), we have

$$B_N = \frac{1}{2\pi}\frac{\int_0^\infty |H(\omega)|^2\,d\omega}{|H(\omega_0)|^2}\,. \tag{4.67}$$

This definition of the equivalent noise bandwidth B_N allows us to discuss practical linear systems by using their idealized equivalents. For example, the equivalent noise bandwidth of the RC low-pass filter in Example 4.7.2 is $B_N = 1/(4RC)$. Therefore the mean-square voltage at the filter output is $\overline{v_o^2(t)} = 4kTRB_N = kT/C$.

EXAMPLE 4.7.3

Compute the -3-dB bandwidth and the equivalent noise bandwidth of a filter with the following magnitude transfer characteristic:

$$|H(\omega)| = \frac{1}{\sqrt{1+\omega^4}}\,.$$

Solution Because $H(0) = 1$, the -3 dB bandwidth is found by solving

$$|H(\omega_1)| = \frac{1}{\sqrt{2}}, \quad \text{or} \quad \omega_1 = 1; \qquad f_1 = \frac{1}{2\pi} = 0.159 \text{ Hz}\,.$$

Using Eq. (4.67), we find the equivalent bandwidth is (cf. Appendix A)

$$B_N = \frac{1}{2\pi}\int_0^\infty \frac{1}{1+\omega^4}\,d\omega = \sqrt{2}/8$$

$$= 0.177 \text{ Hz}\,.$$

Thus for this particular filter (known as a second-order Butterworth filter) the equivalent noise bandwidth is about 11% greater than the -3-dB bandwidth.

DRILL PROBLEM 4.7.3

Compute the equivalent noise bandwidth of the (RC-type) filter whose magnitude transfer characteristic is $|H(\omega)| = 1/\sqrt{1+\omega^2}$. Compare your result to the -3-dB bandwidth.

ANSWER: 0.250 Hz; 57% greater.
[*Note:* As the order of the filter increases, the −3-dB bandwidth and the noise equivalent bandwidth agree more closely.]

4.7.4 Available Power and Noise Temperature

From Eqs. (4.49), (4.50), and (4.57), the thermal noise power generated in a resistor R is

$$P_n = 4kTB\,. \tag{4.68}$$

How much of this noise power can be extracted? Using a matched resistive load R (noise-free) for maximum power transfer, we find that the voltage transferred is exactly one-half of the open circuit voltage. The maximum available power, P_a, is then one-fourth that given in Eq. (4.68), or

$$P_a = kTB\,. \tag{4.69}$$

This available noise power, P_a, is the maximum thermal noise power that can be extracted from a noisy resistor.

Examining Eq. (4.69), we see that k is a constant and B is the equivalent noise bandwidth—constant for a given system. The temperature T is then directly related to the available noise power. A convenient way to describe the input noise power is to specify it as a noise temperature. Thus the *noise temperature* specifies the thermal noise power into a matched resistance.

The noise-free resistor used for matching purposes above is, of course, fictitious. In practice, we usually wish to connect an amplifier (receiver) that has an input resistance R for maximum power transfer. A simplified model of this amplifier is an input resistance R at an equivalent noise temperature T_e followed by a power gain G_p. In other words, the noise temperature T_e is the effective temperature of a white thermal noise source at the system input that would be required to produce the same noise power at the output of an equivalent noiseless system. Some very low noise amplifiers, for example, have equivalent noise temperatures as low as 10 K to 30 K while standard broadcast receivers may have noise temperatures on the order of 1000 K.

Note that the equivalent noise temperature is not necessarily the ambient temperature of the amplifier. Equivalent noise temperatures below ambient temperature are possible by use of amplifiers with a low resistive component in the gain characteristic (e.g., parametric amplifiers). Sometimes cryogenic cooling is also employed to lower the noise temperature.

An equivalent circuit for a receiver operating in the presence of thermal noise is shown in Fig. 4.15. Here all of the noise (white) generated by the amplifier is included in the equivalent noise temperature T_e. Thermal-noise effects of the output resistance R_o can also be included in T_e. However, if the power gain G_p is

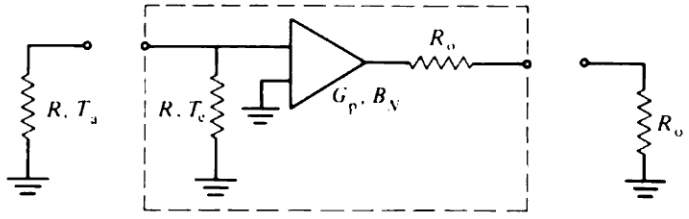


Figure 4.15 Receiver equivalent circuit for thermal noise.

appreciable, the effect of R_o on the overall noise output is negligible, as shown in the following example.

The noise temperature of the matched resistance connected to the receiver input is usually set at 290 K for test purposes. In operating systems, T_a represents the effective noise temperature of the sky and the surrounding noise environment as seen through the antenna, as well as of any transmission-line losses from the antenna to the receiver input terminals.

EXAMPLE 4.7.4

We wish to design a high-gain cascaded-stage amplifier. The first stage power gain is fixed at 20 dB. Succeeding stages have provisions for gain control but the maximum gain per stage is 20 dB. The maximum net power gain of the amplifier to to be such that the thermal noise power level from internally generated noise is 20 milliwatts at the output. Determine the minimum number of stages needed if $T_e = 600$ K, $B_N = 10$ MHz.†

Solution The internally generated noise power in the first stage, referred to the input, is

$$P_a = kT_eB$$
$$= (1.38 \times 10^{-23})(600)(10^7) = 8.28 \times 10^{-14} \text{ W}.$$

After 20 dB of gain, this becomes 8.28×10^{-12} watt of noise power at the input to the second stage. The available input noise power to the second stage is then 8.28×10^{-12} watt coming from the first stage, plus 8.28×10^{-14} watt originating from thermal noise in the second stage (and referred to its input). The contribution from the second stage is almost negligible (= 1%) compared to the noise power coming from the first stage. Our conclusion, then, is that *the noise performance of cascaded amplifiers is primarily dependent only on the noise performance of the first stage if that stage has appreciable gain.*

† The equivalent noise bandwidth does vary somewhat with the number of amplifier stages but we shall ignore this change here.

The required maximum power gain, G_p [$G_p = |H(\omega_0)|^2$] is

$$\begin{aligned} G_p &= P_o/P_a \\ &= 20 \times 10^{-3}/(8.28 \times 10^{-14}) = 2.42 \times 10^{11} \\ &\cong 114 \text{ dB}. \end{aligned}$$

Thus the amplifier will require a minimum of six (6) stages to meet the given specifications.

4.7.5 ★ Noise Figure

It is convenient to develop a concise way to state the equivalent noise temperature of an amplifier relative to a fixed standard. This leads to the definition of a noise figure which can then be used as a figure-of-merit in comparisons between amplifiers for low-level signals.

Let the input and output signal voltages (or currents) in a given system be $s_i(t)$, $s_o(t)$, respectively, and let the input and output noise voltages (or currents) be $n_i(t)$, $n_o(t)$. The input signal-to-noise ratio, $(S/N)_i$, is[†]

$$(S/N)_i = \overline{s_i^2(t)}/\overline{n_i^2(t)}, \tag{4.70}$$

and the output signal-to-noise ratio is

$$(S/N)_o = \overline{s_o^2(t)}/\overline{n_o^2(t)}. \tag{4.71}$$

The system always adds some noise so that the input signal-to-noise ratio is higher than the output signal-to-noise ratio. To measure the amount of degradation, we define a noise figure, F, to be the ratio of the input signal-to-noise ratio divided by the output signal-to-noise ratio:

$$F \triangleq \frac{(S/N)_i}{(S/N)_o}. \tag{4.72}$$

By definition, the input noise power in Eq. (4.72) *is equivalent to the thermal noise power provided by a resistor matched to the input and at a temperature of* $T_0 = 290$ K. The noise figure of a perfect system is unity, and the introduction of additional noise causes the noise figure to be larger than 1.

Some convenient simplifications can be made when only thermal noise is present in the system. Let us apply a signal $s_i(t)$ and thermal noise $n_i(t)$ at a temperature T_0 to the input of an amplifier with a power gain G_p and noise bandwidth B. The available input noise power is

$$N_i = kT_0B. \tag{4.73}$$

[†] The resistance factor drops out so that this is also a power ratio.

The output signal power is

$$S_o = S_i G_p . \tag{4.74}$$

The amplifier adds some thermal noise. Representing this noise by an equivalent noise temperature T_e referred to the amplifier input, we have

$$N_o = kT_0 BG_p + kT_e BG_p . \tag{4.75}$$

When we substitute (4.73)–(4.75) into Eq. (4.72), the noise figure of the amplifier is

$$F = 1 + \frac{T_e}{T_0} . \tag{4.76}$$

This result is so simple and easy to apply that even if not all noise sources are thermal, the effects are often included in an equivalent noise temperature using test data.[†] It is quite common to express the noise figure in decibels:

$$F_{dB} = 10 \log_{10}(F) .$$

Note that the noise figure of an ideal system which generates no noise itself is $F = 1$, and that portion of the noise figure of any system arising from internally generated noise is $(F - 1)$.

EXAMPLE 4.7.5

A given amplifier has a 4-dB noise figure, a noise bandwidth of 500 kHz, and an input resistance of 50 Ω. Calculate the rms signal input that yields an output signal-to-noise ratio of unity when the amplifier is connected to a 50-Ω input at 290 K.

Solution The available input noise power is

$$P_n = kT_0 B = 2.00 \times 10^{-15} \text{ W},$$

so that

$$\overline{n_i^2(t)} = P_n R = 1.00 \times 10^{-13} \text{ V}^2,$$

$$F = 4 \text{ dB} = 2.51 .$$

Using Eq. (4.72), we have

$$(S/N)_i = F(S/N)_o = F .$$

The required input signal is

$$\overline{s_i^2(t)} = (2.51)(1.00 \times 10^{-13}) \text{ V}^2$$

$$\sqrt{\overline{s_i^2(t)}} = 0.501 \ \mu\text{V} .$$

[†] The standard value for T_0 is 290 K.

EXAMPLE 4.7.6

A resistive attenuator (e.g., a coaxial cable or waveguide) at a temperature T_0 has matched input and output resistances and an attenuation (in power) of α, where $\alpha > 1$. Determine the equivalent noise temperature and noise figure of the attenuator when the input source and the attenuator are at temperature T_0.

Solution Using Eq. (4.75) and noting that $G_p = 1/\alpha$, we have

$$N_o = kT_0B(1/\alpha) + kT_eB(1/\alpha).$$

Looking back into the output terminals, we see that the attenuator appears entirely resistive and at a temperature T_0 so that

$$N_o = kT_0B.$$

Equating these two expressions, we have

$$T_e = (\alpha - 1)T_0.$$

Use of Eq. (4.76) then yields

$$F = \alpha.$$

A similar procedure can be used if the attenuator is at some other temperature.

EXAMPLE 4.7.7

A multistage amplifier has stage power gains $G_1, G_2, G_3, \ldots$, and stage noise figures $F_1, F_2, F_3, \ldots$, respectively. Show that the overall noise figure is

$$F = F_1 + [(F_2 - 1)/G_1] + [(F_3 - 1)/G_1G_2] + \cdots$$

and hence that the first stage is the most significant in the determination of the overall noise figure if $G_1 \gg 1$.

Solution The result follows easily when one considers that the noise introduced by each stage is amplified only by the gains of that stage and succeeding stages. Assume a three-stage amplifier with equivalent noise temperatures $T_{e_1}, T_{e_2}, T_{e_3}$:

$$S_o = S_iG_1G_2G_3,$$

$$N_i = kT_0B,$$

$$\begin{aligned} N_o &= kT_0BG_1G_2G_3 + kT_{e_1}BG_1G_2G_3 + kT_{e_2}BG_2G_3 + kT_{e_3}BG_3 \\ &= N_iG_1G_2G_3 + (F_1 - 1)N_iG_1G_2G_3 + (F_2 - 1)N_iG_2G_3 + (F_3 - 1)N_iG_3 \\ &= F_1N_iG_1G_2G_3 + (F_2 - 1)N_iG_2G_3 + (F_3 - 1)N_iG_3, \end{aligned}$$

$$F = \frac{S_i/N_i}{S_o/N_o} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2}.$$

Extensions of this result are straightforward.

This result demonstrates the advantages of using a first stage of amplification with not only a low noise figure (F_1) but also high gain (G_1). From Example 4.7.6 we see that the performance of a lossy transmission line fails on both counts. Therefore some amplification with low-noise stages is often used near or at the antenna in low-noise receiving systems before the signal is transferred (via a transmission line) to the main receiver.

DRILL PROBLEM 4.7.4

A receiver for geostationary satellite transmissions at 2 GHz consists of an antenna preamplifier with a noise temperature of 127 K and a gain of 20 dB. This is followed by an amplifier with a noise figure of 12 dB and a gain of 80 dB. Compute the overall noise figure and equivalent noise temperature of the receiver.

ANSWER: 2.0 dB; 170 K.

4.7.6 ★ Sky-Noise Temperature

When an antenna is connected to the input of a receiver, it is convenient to represent it by a resistor that is matched to the receiver input and whose temperature represents the effective noise of the sky and the surrounding noise environment as seen through the antenna. This antenna temperature will in most cases differ substantially from 290 K.

If the input resistor, now representing an antenna, is at some arbitrary temperature T_a, we can modify Eq. (4.75) to give

$$N_o = kT_aBG_p + kT_eBG_p,$$

$$N_o = kB(T_a + T_e)G_p. \tag{4.77}$$

The equivalent noise temperature of an antenna is therefore easy to interpret and can be compared directly with the receiver noise temperature. Note that when T_a and T_e contribute equally to the output noise, the largest possible improvement by going to a perfect receiver amounts to only a factor of 2 in signal-to-noise improvement. Therefore efforts to obtain a low equivalent noise temperature in a receiver really pay large dividends only if the antenna noise temperature is low.

Average antenna noise temperatures are mainly a function of frequency and the pointing of the antenna. A graph illustrating some of the main sources of antenna noise temperature is shown in Fig. 4.16. Although the concept of noise figure is useful for the testing and comparisons of receivers, the concept of equivalent noise temperature is more useful in computations of actual system performance.

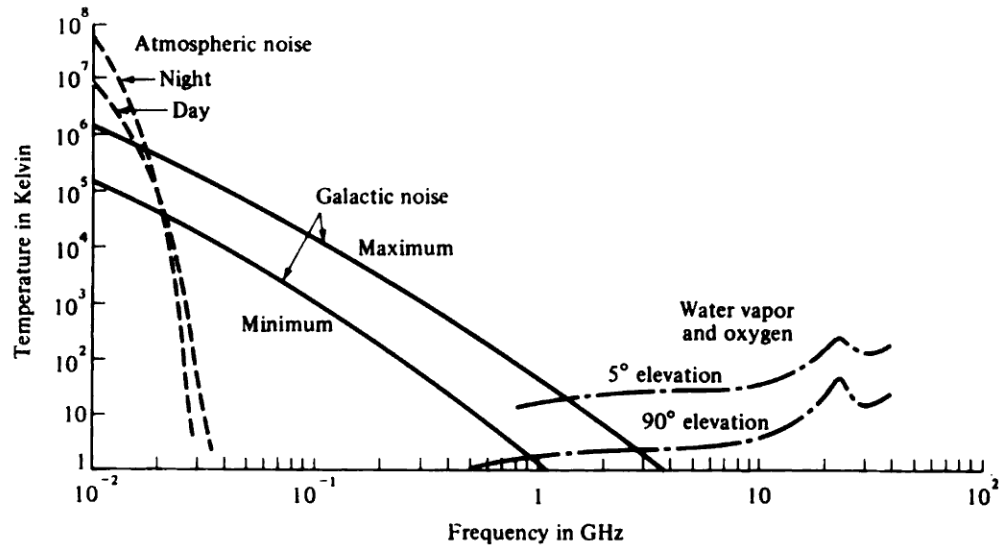


Figure 4.16 Some average sky noise temperatures.

The principal source of antenna noise below 30 MHz is atmospheric noise that results mainly from lightning discharges. Propagation effects that make long-distance communication possible at these lower frequencies also provide good propagation for electrical storm activities occurring throughout the world. Galactic or cosmic noise is a major contributor above 30 MHz and up to the GHz range. This noise arises from radiation from outer space; for narrow beamwidths, the intensity of galactic noise is a function of antenna pointing direction. Our sun is also an active source of radiation, but its effects are more localized in angular direction.

Water vapor and oxygen act as attenuators of rf energy, particularly around 23 GHz. These attenuation effects not only decrease signal strength but also act as thermal noise sources (see Example 4.7.6). The frequency range between about 2 to 8 GHz, bounded by the effects of galactic noise and noise due to oxygen and water vapor, is referred to as the "low-noise window." This is a preferred range for low-noise receivers for space telemetry and radio telescopes.

EXAMPLE 4.7.8

The receiver in Drill Problem 4.7.4 has a bandwidth of 1 MHz. The receiving antenna gain is 40 dB and the antenna noise temperature is 59 K. If the satellite antenna gain is 6 dB and expected path losses are 190 dB, what is the minimum required satellite transmitter power to achieve a 14-dB S/N ratio at the output of the receiver?

Solution See Table 4.1.

TABLE 4.1 ***S/N*** **Budget for the Ground-to-Geostationary Satellite Communications Problem of Example 4.7.8.**

Equivalent receiver noise temperature	170 K	
Antenna noise temperature (given)	59 K	
Total noise temperature	229 K	
Total noise temperature—dB above 1 K	23.6	
Boltzmann's constant—dB/K/Hz	−228.6	
Bandwidth (given)—dB above 1 Hz	60.0	
Total receiver noise power	−145.0 dB	
Receiver S/N allowance (given)	14.0 dB	
Minimum receiver signal power	−131.0 dB	
Receiving antenna gain (given)	40 dB	
Satellite transmitting antenna gain (given)	6 dB	
Minimum detectable signal power		−177.0 dB
Path losses (given)		190 dB
Minimum required satellite transmitter power		13.0 dB (20 W)

4.7.7 ★ Antennas and Communication Links

Because thermal noise is a fundamental limitation in communication systems, we extend our discussion a little farther into how the propagation of power from transmitter to receiver can be handled in terms of signal-to-noise calculations.

Figure 4.17 illustrates a model of a communication system for the transmission of power through free space, as, for example, in a satellite communication system. Such a model is referred to as a *communication link*. The transmitter supplies power to the transmitter antenna, which is directive and radiates power within a beamwidth θ toward the receiver.

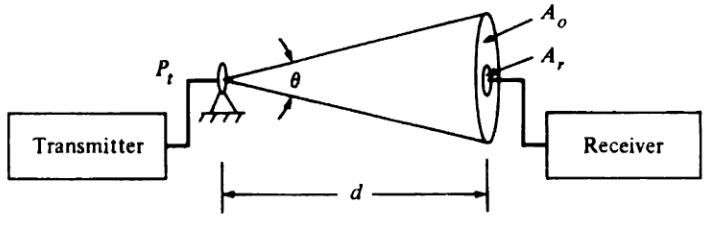


Figure 4.17 Model of a communication link.

If the transmitting antenna were isotropic, it would radiate power uniformly in all directions, and a surface of constant power density at distance d would be a spherical surface of area $A = 4\pi d^2$. A parameter used to describe the directivity or "gain" of

the antenna is defined as the ratio of the area illuminated by an isotropic antenna to that illuminated by the directive antenna:

$$G_t = \frac{4\pi d^2}{A_o}. \tag{4.78}$$

The beamwidth θ is inversely proportional to the gain of the antenna. The antenna gain G_t is often expressed in dB, $10 \log_{10} G_t$.

The receiver is characterized by an effective area A_r of its antenna and by the net receiver noise temperature $T = T_a + T_e$. The received power is the fraction A_r/A_o of the transmitted power (see Fig. 4.17):

$$P_r = P_t \frac{A_r}{A_o},$$

or, using Eq. (4.78),

$$P_r = P_t G_t \frac{A_r}{4\pi d^2}. \tag{4.79}$$

The product $P_t G_t$ is referred to as the *effective isotropic radiated power* (EIRP) of the transmitter. It is the effective power that an isotropic antenna would radiate toward the receiver.

The role of a properly designed receiving antenna is to focus the received power, by means of phasing, to extract maximum signal power. There are limitations, based on optical criteria, on how well this can be accomplished. By an analogy with optics, the directivity or gain G_r of the antenna and its effective area A_r are related by

$$G_r = \frac{4\pi A_r}{\lambda^2}, \tag{4.80}$$

where λ is the wavelength of the radiation. The effective area A_r is related principally to the physical shape and dimension of the antenna. For example, a parabolic reflector antenna has an effective area, A_r, in the range 54–65% of the actual physical area $\pi D^2/4$, where D is the diameter of the "dish." Thus, the power gain for a paraboloidal reflector is $G_r \approx 0.55(\pi D/\lambda)^2$.

There is a reciprocity relationship between an antenna used for transmitting and the same antenna used for reception. Thus the phasing effects used in the receiving antenna are also those that can be used to make the beam more directive in the transmitting antenna. A given antenna has the same directivity patterns in both cases. Combining Eqs. (4.79), (4.80), we obtain

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2, \tag{4.81}$$

where the factor $(4\pi d/\lambda)^2$ is called the free-space path loss.† The fact that this loss is frequency dependent arises because the receiving antenna is described in terms of antenna gain, rather than its effective area. For a fixed antenna size, this apparent frequency dependence is canceled by the inverse frequency dependence of the receiving antenna gain.

The important receiver parameter here is not the received signal power, but its ratio to the received noise power. If the dominant source of noise is thermal noise, then the noise power is $P_n = kT_rB$, where T_r is an effective noise temperature resulting from both receiver and antenna noise. Dividing Eq. (4.81) by the noise power, the received signal-to-noise ratio is

$$\frac{P_r}{P_n} = \frac{1}{kB}(P_tG_t)\left(\frac{\lambda}{4\pi d}\right)^2\left(\frac{G_r}{T_r}\right). \tag{4.82}$$

Aside from the scaling constant $1/(kB)$, this result is dependent only on (1) (P_tG_t), the EIRP of the transmitter; (2) the free-space path loss, $(4\pi d/\lambda)^2$, and (3) (G_r/T_r), the ratio of the receiving antenna gain to the net receiver noise temperature. The ratio G_r/T_r is referred to as the *receiver figure of merit*. For a given EIRP and path losses, it is the ratio G_r/T_r that really counts. This is very important in satellite-to-ground receiver designs, where large antennas and low noise amplifiers are essential for adequate performance in consideration of relatively low EIRP and high free-space path losses.

DRILL PROBLEM 4.7.5

For a transmitter EIRP of 80 W in Example 4.7.8, and undetermined receiving antenna gain G_r and total receiver noise temperature T_r, show that

$$(S/N)_{dB} = (G_r/T_r)_{dB} - 2.4\,,$$

where (G_r/T_r) is the receiver figure of merit. [*Note:* T_r is referenced to 1 K to yield correct units in this expression.]

4.7.8 ★ Other Sources of Band-limited White Noise

There are other types of noise which can be described in terms of white noise sources, although the approximation may not hold over as wide a range of frequencies as for thermal noise. One example is that of shot noise.

† For computations, it is convenient to express the free-space path loss as an attenuation factor (in dB):

$$\alpha = 36.6 + 20\log_{10} f + 20\log_{10} d\,,$$

where f is the center frequency in MHz and d is the path distance in miles. Path loss formulas can be found in handbooks; see, for example, *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, Seventh ed., E. C. Jordan, editor in chief, Indianapolis: Howard W. Sams and Company, 1985, Chapter 33.

Shot noise, like thermal noise, is due to the discrete nature of physical matter. It arises in physical devices when a charged particle moves through a potential gradient without collisions and with a random starting time. Averaging over many such particles we have an average flow, but there will always be fluctuations about this average.

In vacuum tubes, shot noise arises as a result of the random emission of electrons from the cathode. In semiconductor devices, shot noise arises as a result of the random diffusion of minority carriers and the random generation and recombination of hole-electron pairs. For these cases, the power spectral density is approximately flat up to frequencies on the order of $1/\tau$, where τ is the transit time or lifetime of the charge carriers. In terms of mean-square current, the power spectral density is

$$S_i(\omega) = q\overline{i(t)} + 2\pi\overline{i(t)}^2\,\delta(\omega) \tag{4.83}$$

where q is the charge of an electron $= 1.6 \times 10^{-19}$ coulomb. The first term in Eq. (4.83) corresponds to the ac or fluctuation part of the noise current and the second term corresponds to the nonzero mean value, $\overline{i(t)}$.

EXAMPLE 4.7.9

If there are $\bar{n}$ electrons per second (time-averaged) crossing an evacuated region in a potential gradient, determine the spectral density of the fluctuation current due to shot noise. Assume a white spectral density.

Solution Each electron leaves at a random starting time and traverses the given region. Let the current induced in the circuit as a result of one electron be $i_e(t)$. Because the charge on the electron is q coulombs,

$$\int_{-\infty}^{\infty} i_e(t)\,dt = q.$$

Let the Fourier transform of $i_e(t)$ be $I_e(\omega)$:

$$I_e(\omega) = \int_{-\infty}^{\infty} i_e(t)e^{-j\omega t}\,dt,$$

$$I_e(0) = \int_{-\infty}^{\infty} i_e(t)\,dt = q.$$

The energy spectral density of one electron is $|I_e(\omega)|^2$. Because the starting times are random, the induced current pulses are independent of each other and we can add their energy spectral densities† so that

$$S_i(\omega) = \frac{N}{T}\,|I_e(\omega)|^2 = \bar{n}|I_e(\omega)|^2.$$

† This is an exception to the rule that spectral densities can be added without regard to phase only when their relative phases are 0° or 180°.

Within the white noise assumption,

$$S_i(\omega) = \bar{n}|I_e(0)|^2 = \bar{n}q^2 = q(\bar{n}q) = q\overline{i(t)}.$$

Thus $S_i(\omega)$ varies in direct proportion to the average current $\overline{i(t)}$.

Another noise component that arises in the division of current in multielectrode devices is called partition noise. This partitioning, which is a random process, introduces noise because the charge carriers are discrete.

Both transistors and multielectrode vacuum tubes have all three sources of noise present—shot noise, partition noise, and thermal noise. All three can generally be treated as if their power spectral densities were flat across the bandwidths of interest. Sometimes the effects of shot noise and partitioning noise are accounted for by placing an equivalent thermal noise source in the equivalent circuit model.

4.8 SUMMARY

The quantity $|F(\omega)|^2$ describes the relative amount of energy of a given signal $f(t)$ versus frequency and is called the energy spectral density of $f(t)$. Likewise, the function describing the relative power contributions of a signal $f(t)$ versus frequency is called the power spectral density, $S_f(\omega)$. The power spectral density of an aperiodic signal is a continuous function of frequency. The power spectral density of a periodic signal approaches a series of impulse functions located at the harmonic frequencies of the signal as the observation time becomes long.

The mean-square value of a waveform can be found from the area under its power spectral-density function. The ratio of the mean-square signal to the mean-square noise is called the signal-to-noise ratio.

The inverse Fourier transform of the power spectral density is the autocorrelation function. For real-valued signals of finite duration, the autocorrelation function is, aside from a normalizing constant, given by the convolution of $f(-t)$ with $f(t)$. Similarly, the convolution of $f(-t)$ with a different signal is called the cross-correlation function.

Two signals are uncorrelated if their cross-correlation is zero. The power in the sum of two signals is equal to the sum of the powers in the two signals if they are uncorrelated. This rule can be extended to power spectral densities.

The inverse Fourier transform of the energy spectral density is the autocorrelation function for finite-energy signals.

Noise having a flat power spectral density is called *white*. Thermal noise, arising from the random motion of free electrons in a conducting medium, is white up to infrared frequencies with $S_n(\omega) = 2kT$ watts per Hz.

The available power from thermal noise is kTB watts, where B is the equivalent noise bandwidth of the system. A thermal-noise source may be specified as an equivalent temperature. Noise figure is a convenient figure-of-merit for systems when

a reference temperature is set at $T_0 = 290$ K. Antenna input noise can be specified as an equivalent noise temperature, called the sky-noise temperature.

Shot noise and partition noise are two other types of noise which are white over fairly wide frequency ranges. All three types are present to varying degrees in thermionic and semiconductor devices. An equivalent thermal-noise source is sometimes used to account for these and other sources of band-limited white noise.

Selected References for Further Reading

1. S. J. Mason and H. J. Zimmerman. *Electronic Circuits, Signals, and Systems*. New York: John Wiley & Sons, 1960.
   This book has a good discussion of autocorrelation and cross-correlation for deterministic signals.
2. D. Roddy. *Satellite Communications*. Englewood Cliffs, N.J.: Prentice Hall, 1989.
   Chapter 6 is a good introductory discussion of a system link budget for satellite communications. The text also contains general information about satellites; easy reading.
3. W. R. Bennett. *Electrical Noise*. New York: McGraw-Hill, 1960.
   A general reference on the sources of electrical noise and the methods of handling the effects of noise in linear systems.
4. W. W. Mumford and E. H. Scheibe. *Noise Performance Factors in Communication Systems*. Norwood, Mass.: Artech House, 1968.
   A thorough treatment of noise and the use of noise figure in communication systems at the level of this chapter.
5. B. Sklar. *Digital Communications*. Englewood Cliffs, N.J.: Prentice Hall, 1988.
   Chapter 4 is a good discussion of systems link budget calculations and considerations, with examples from satellite systems.
6. P. F. Panter. *Communication Systems Design: Line-of-Sight and Troposcatter Systems*. New York: McGraw-Hill, 1972.
   Chapter 6 has a practical discussion of S/N calculations for communication systems design at a slightly more advanced level.
7. E. J. Baghdady (ed.). *Lectures on Communication Systems Theory*. New York: McGraw-Hill, 1960.
   Chapters 15–17 discuss antenna noise and Chapters 21 and 22 discuss some system S/N considerations at a more advanced level.
8. A. Van der Ziel. *Noise*. Englewood Cliffs, N.J.: Prentice Hall, 1954.
   A classic reference on noise and noise calculations.

PROBLEMS

4.1.1 The pulse signal $f(t)$ shown in Fig. P–4.1.1 is applied to the input of an ideal low-pass filter with a variable cut-off frequency, f_c. Find the approximate energy, on a 1-ohm basis, in the output signal for

a) $f_c = 0.1$ Hz;
b) $f_c = 10$ Hz.

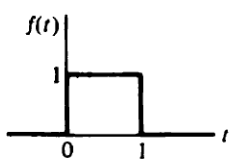


Figure P–4.1.1.

4.1.2 A given pulse generator produces very narrow voltage waveforms, which we approximate here as impulses. Assume that the generator has a negligibly low output impedance. The generator is first connected to an ideal low-pass filter with bandwidth B Hz and unity low-frequency gain, and the resulting output energy is measured for an input pulse. Next, the generator is connected to a RC low-pass filter (cf. Fig. 2.16) and the output energy is measured for an input pulse. Derive a relation between B and RC if the output energies are equal.

4.1.3 The voltage $f(t) = 10te^{-t}u(t)$ is developed across a 50-ohm resistor.

a) Calculate the total energy dissipated in the resistor.
b) What fraction of this energy is contained within a (low-pass) bandwidth of 1 rad/sec?
c) What fraction of this energy is contained within a bandwidth of 2 rad/sec, with a center frequency of 4 rad/sec?

4.1.4 The following pulse signals are each applied to four filters (assume ideal, with unity gain) in parallel: 0–1, 1–2, 2–3, 3–4 radians per second. Assume that the input impedance of this simple spectral analyzer is normalized to 1 ohm. Energy detectors at the outputs of these filters provide output indicators of the four spectral contributions. Compute the four outputs, in decibels relative to 1 joule, for each of the following input signals and compare results.

a) $f_1(t) = e^{-t}u(t)$
b) $f_2(t) = te^{-t}u(t)$
c) $f_3(t) = e^{-|t|}$

4.2.1 A certain signal $f(t)$ has the following power spectral density (assume a 1-ohm resistive load):

$$S_f(\omega) = \left[\frac{1}{1+\omega^2} + \delta(\omega - 2) + \delta(\omega + 2)\right].$$

a) What is the total mean power in $f(t)$?
b) What is the mean power in $f(t)$ within the bandwidth 0.9 to 1.1 rad/sec?
c) What is the mean power in $f(t)$ within the bandwidth 1.9 to 2.1 rad/sec?
d) Find a signal $f(t)$, in terms of cosines and exponentials, that will satisfy this power spectral density. Are other solutions possible?

4.2.2 A given voltage is $f(t) = 4 \cos 20\pi t + 2 \cos 30\pi t$ across 1 Ω. Note that this is a Fourier series.

a) Determine and sketch the power spectral density of $f(t)$, and identify the component at the fundamental frequency.
b) Calculate the average power, both in the time domain and in the frequency domain, dissipated by $f(t)$ across the 1-ohm resistor.
c) Determine and sketch the power spectral density of $f^2(t)$. (First perform the squaring operation, then write the result in the form of a Fourier series and from this determine the power spectral density.)

4.2.3 A symmetric square wave (i.e., zero average value) with a peak amplitude of 1 V and period T is applied to the input of an amplifier whose magnitude transfer function is

$$|H(\omega)| = \begin{cases} K(1 + \cos \omega)/2 & |\omega| < 4\pi/T \\ 0 & \text{elsewhere} \end{cases}.$$

where K is the voltage gain. Assume that the input and output impedances are resistive and equal to 1 ohm. The voltage gain K is adjusted so that the amplifier output is equal to 1 W.

a) Determine the value of K.
b) The square wave is replaced with a symmetric triangular wave (cf. Table 2.2) with the same peak amplitude, and a period T_1. What is the output power under the above conditions if $T_1 = T$?
c) Repeat part (b) if $T_1 = 2T$.

4.2.4 A sinusoidal generator produces the waveform $v_i(t) = A \cos \omega_0 t$ as the input to the low-pass filter shown in Fig. P–4.2.4.

a) Using the methods of ac circuit analysis, find an expression for the output waveform, $v_o(t)$.
b) Find the average power in $v_o(t)$ across R_2 from the result of part (a).
c) Determine the power spectral density of $v_i(t)$.
d) Determine the power spectral density of $v_o(t)$, and from this result find the average power in $v_o(t)$ across R_2.

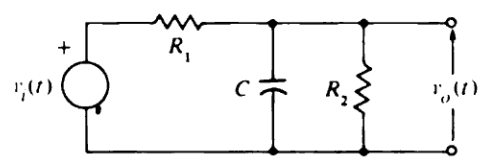


Figure P–4.2.4.

4.3.1 A full-wave rectified sinusoid (cf. Table 2.2) has a peak amplitude A and a period T.

a) Determine its mean-square value in the time domain.
b) Find an expression for the power spectral density (assume 1 ohm).
c) Making use of the above results, show that

$$\sum_{\substack{n=-\infty \\ n \text{ even}}}^{\infty} \frac{1}{(1 - n^2)^2} = \frac{\pi^2}{8}$$

d) Do you reach the same conclusion as in part (c) when using a half-wave rectified sinusoid?

4.3.2 White noise with a two-sided power spectral density of $\eta/2$ watts per Hz (assume 1-ohm source), added to a signal described by $\sqrt{2} \cos 30\pi t$, is applied to the input of a RC low-pass filter (cf. Fig. 2.16). The rms value of the signal at the filter output is 10 mV, and the ratio of the mean-square signal to the mean-square noise at the output is 10 dB.

a) Determine the value of η in terms of the RC time constant of the filter.
b) Determine the RC time constant of the filter.

4.3.3 Repeat Problem 4.3.2 for the case in which the power spectral density of the input signal is $S_f(\omega) = 1/(1 + \omega^2)$.

4.4.1 a) Determine the power spectral density of $F_1 \exp(j\omega_0 t)$ by first finding the autocorrelation function, then taking the Fourier transform of the autocorrelation function.
b) Repeat part (a) for $[F_1 \exp(j\omega_0 t) + F_2 \exp(j2\omega_0 t)]$.
c) Extend your result of part (b) to $\sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t}$.

4.4.2 The complex-valued pulse signal

$$f(t) = \begin{cases} e^{j\omega_c t} & 0 \le t < t_1 \\ 0 & t_1 \le t < T \end{cases}$$

is repeated periodically with period T, $T > 2t_1$.

a) Find the autocorrelation function $R_f(\tau)$, and sketch the real part.
b) Find and sketch the power spectral density $S_f(\omega)$.

4.4.3 Sketch to scale the autocorrelation function of (a) $x(t)$, and (b) $y(t)$ shown in Fig. P–4.4.3 for $T = 16$.

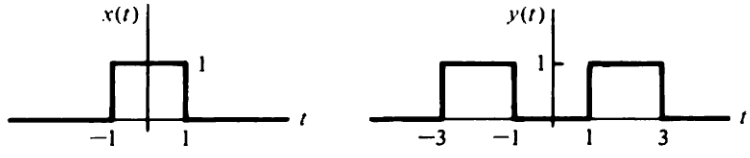


Figure P–4.4.3.

4.5.1 a) Find the autocorrelation function for the sum of the two sinusoids in Drill Problem 4.5.2, assuming a finite averaging time T.
b) Determine the rms value for a finite averaging time for the two cases in Drill Problem 4.5.2, using the result of part (a).
c) Determine the smallest value of T, in terms of ω_1, ω_2, for the second answer in Drill Problem 4.5.2 to be valid.

4.5.2 A white noise source, $\eta/2$ V^2/Hz, is connected to the input of the RC filter shown in Fig. P–4.2.4.

a) Determine the spectral density of the output.
b) Determine the autocorrelation function of the output.
c) Determine the rms value of the output using each of the two preceding results, and compare.
d) What happens as $C \to 0$?

4.5.3 A sinusoidal signal is transmitted; on reception, the signal is present together with additive noise. Assume that the signal and noise are uncorrelated. The autocorrelation function of the observable quantity at the receiver input is $R(\tau) = a \cos \omega_0\tau + b \exp(-c|\tau|)$. The receiver input contains a bandpass filter (BPF—assume ideal) with bandwidth B Hz, centered at $\pm\omega_0$. Find an expression for the S/N ratio at the output of the BPF; express your answer in terms of the S/N ratio at the input to the BPF.

4.6.1 Determine the autocorrelation function of each of the following pulse waveforms by first taking a Fourier transform, next taking the magnitude squared, and then taking an inverse Fourier transform.

a) $e^{-at}u(t)$
b) rect (t/t_1)
c) Sa (Wt)

4.6.2 Two pulse waveforms $x(t)$ and $y(t)$ are shown in Fig. P–4.6.2. Determine and sketch $r_x(\tau)$, $r_y(\tau)$, and $r_{xy}(\tau)$.

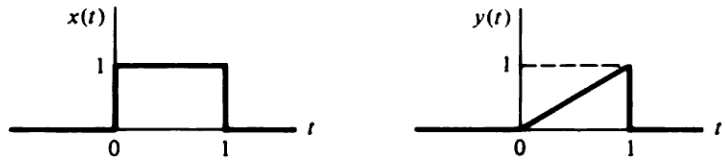


Figure P–4.6.2.

4.6.3 Two complex-valued pulse waveforms $f(t)$ and $g(t)$ are defined by

$$f(t) = \begin{cases} e^{-j\omega_1 t} & 0 \le t < T \\ 0 & \text{elsewhere} \end{cases},$$

$$g(t) = \begin{cases} e^{j\omega_1 t} & 0 \le t < T \\ 0 & \text{elsewhere} \end{cases}.$$

a) Determine the autocorrelation function of $f(t)$.
b) Sketch the real part of $r_f(\tau)$ for $\omega_1 \gg \pi/T$.
c) Determine and sketch $r_{fg}(\tau)$ for each of the following conditions:
   (i) $\omega_1 = 0$;
   (ii) $\omega_1 = \pi/T$.

4.6.4 A pulse waveform $f(t)$ is applied to the input of the two-tap transversal filter shown in Fig. P–4.6.4. Find an expression for the autocorrelation function of the output, $r_g(\tau)$, in terms of $r_f(\tau)$, assuming that a_0, a_1, and T are given constants.

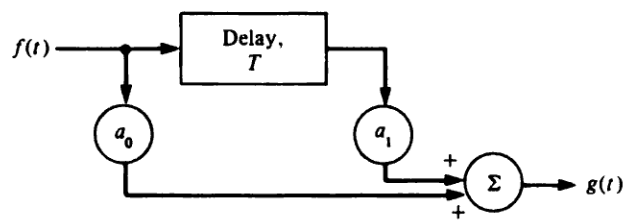


Figure P–4.6.4.

√ **4.6.5** Use numerical methods to compute and plot the autocorrelation function of the signal waveform in Problem 3.8.10. Use either a direct approach (similar to convolution) or the FFT; for the latter, use 64 sample points, assigning zeros to points beyond the signal duration.

4.7.1 In a certain condition it is given that the rms thermal voltage developed across the series combination of two resistors R_1, R_2 is γ times that developed across the parallel combination of these two resistors.

a) Determine a lower bound on γ if both resistors are at the same temperature.
b) Find the relation required between R_1, R_2 to achieve this lower bound.

4.7.2 The frequency transfer function for the RC bandpass filter shown in Fig. P–4.7.2 can be approximated by that for two independent cascaded RC filter sections if $R_1 \ll R_2$. Using this approximation, determine an expression for the spectral density of the thermal noise voltage at the output terminals of the filter.

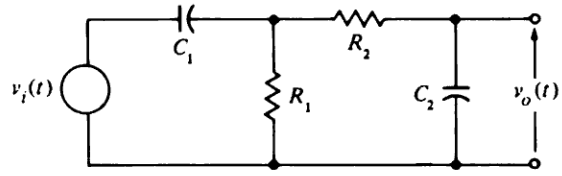


Figure P–4.7.2.

4.7.3 Find the rms thermal noise arising from the resistance bridge circuit in Fig. P–4.7.3 under the following limiting conditions: a) $R_0 \to 0$; b) $R_0 \to \infty$.

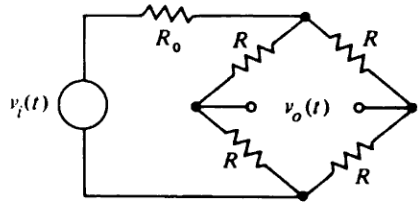


Figure P–4.7.3.

4.7.4 The input of a voltage amplifier is connected to a 1-kΩ resistor and the output to a 100-kΩ resistor, as shown in Fig. P–4.7.4. The voltage gain of the amplifier is 10, the input impedance is 1 MΩ (assumed noise-free), and the output impedance is 100 Ω (noise-free). The bandwidth of the amplifier is 1 MHz and the amplifier noise is 10 μV rms, referred to the input. The temperature of the entire system is 300 K. Compute the rms thermal noise at the output (point 2).

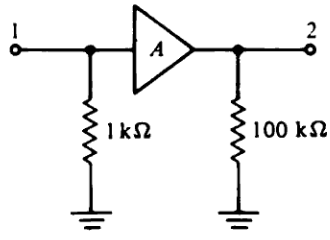


Figure P–4.7.4.

4.7.5 A sinusoidal generator develops the waveform $v_i(t) = A \cos \omega_0 t$ as the input to the RC filter shown in Fig. P–4.2.4.

a) Derive an equation for the rms value of the thermal noise at the output of the filter.
b) Determine that value of C which will yield the highest S/N ratio (i.e., ratio of average signal power to average noise power) at the output.

4.7.6 Compute the noise-equivalent bandwidth of the nth order Butterworth low-pass filter whose magnitude frequency transfer function is: $|H(\omega)| = 1/\sqrt{1 + \omega^{2n}}$, and compare your answer to the -3 dB bandwidth for $n = 3, 4, 5$. Show that the higher the order of the filter, the closer the agreement between these two measures of bandwidth.

4.7.7 Compute the noise-equivalent bandwidth of the RC bandpass filter shown in Fig. P–4.7.2 for $R_1C_1 = R_2C_2$, and using the approximation described in Problem 4.7.2 for $R_1 \ll R_2$. Choose the midband frequency to be that frequency at which $|H(\omega)|$ is maximum.

4.7.8 A bandpass amplifier is to be designed to meet the following specifications:

$$B_N = 200 \text{ kHz}$$

$$(S/N)_{\text{output}} \geq 20 \text{ dB}$$

Max. output = 1 V rms across 300 Ω

Estimated noise temp. of input stage = 1000 K

What is the maximum gain, in dB, for which you would design the amplifier to meet these specifications?

4.7.9 A certain amplifier has input and output resistances of 50 Ω and a noise-equivalent bandwidth of 140 kHz. When connected to a matched source and a matched load, the net gain is 50 dB. When a 50-Ω resistor at 290 K is connected to the input, the output rms noise voltage across 50 Ω is 100 μV. Determine the equivalent noise temperature, T_e, of the amplifier.

★ **4.7.10** Suppose that receiving system specifications are set such that the antenna noise temperature T_a and the receiver equivalent noise temperature T_e contribute equally to the output noise power. Using Fig. 4.16, estimate the range of receiver noise figures (in dB) required if the center frequency is

a) 10 MHz;
b) 100 MHz;
c) 1 GHz.

★ **4.7.11** A resident in a rural area finds that line-of-sight TV reception from a given direction is blocked by a hill, and decides to place an antenna on top of the hill and use a transmission line down to the TV receiver. It is estimated that the noise figure of the TV receiver is 15 dB and that the transmission line losses are 4 dB. Determine the following, in dB:

a) The overall noise figure of the TV receiving system with the antenna on top of the hill.
b) The overall noise figure if a preamplifier with a 5 dB noise figure and 20 dB gain is connected between the transmission line and the input of the TV set.
c) The overall noise figure if the preamplifier above is moved to the antenna site and connected between the antenna and the transmission line.

★ **4.7.12** In Problem 4.7.11, assume that when the TV receiver is tuned to the desired station the S/N ratio, measured at the output of the radio-frequency amplifier in the receiver, is 20 dB for the conditions of part (a) and an assumed sky noise temperature of 2000 K. Calculate the expected S/N ratio for the two conditions in which a preamplifier is used. (It may be helpful in interpretation of these results to note that a S/N ratio of 20 dB in commercial TV yields an unacceptable picture as a result of excessive "snow" resulting from the additive noise, while a S/N ratio of 30 dB yields a fairly good picture.)

★ **4.7.13** Antenna noise consists of both externally caused noise (e.g., sky noise) and internally caused noise (e.g., attenuation from cables, etc.). In this chapter we have assumed that the latter is small compared to the former, and this is generally true in practice. Here we investigate what are the effects of the internally caused noise in the antenna. It is customary to refer all noise measurements to the input terminals of the preamplifier (or receiver, if no preamplifier is used). If α is the attenuation ($P_{\text{in}}/P_{\text{out}}$) and T_a is the

externally caused noise temperature, show that the effective noise temperature of the antenna at the preamplifier input is given by

$$T_{\text{eff}} = [T_a + (\alpha - 1)290]/\alpha .$$

[*Hint:* Write an expression for the noise at the input terminals of the preamplifier and then define a new equivalent temperature.]

★ **4.7.14** The minimum-loss resistive attenuator shown in Fig. P–4.7.14 can be used to match a 300-Ω antenna to a 75-Ω cable. Calculate the equivalent noise temperature and the noise figure of the attenuator. [*Hint:* Use the result of Example 4.7.6.]

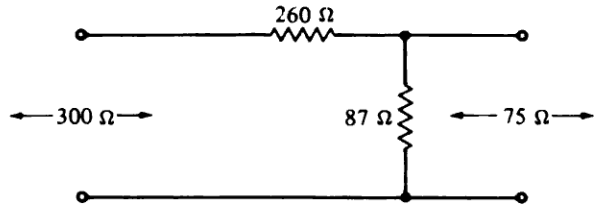


Figure P–4.7.14.

★ **4.7.15** A proposed receiving station for a space experiment has a receiver (e.g., a cooled low-noise amplifier) with a 0.2 dB noise figure and a center frequency of 3 GHz. The receiving antenna has a gain of 48 dB and the spacecraft antenna gain is 6 dB. If the transmitted power is 10 W and a minimum 14 dB S/N ratio is required at the receiver, make use of Fig. 4.16 to estimate the maximum system bandwidth for communication over a distance of 5×10^9 miles (i.e., approximately the extent of our planetary system).

★ **4.7.16** A given receiving system for satellite transmissions has a physical separation between antenna site and the main receiver. At the antenna site a low-noise amplifier (LNA) provides initial signal gain and a down converter shifts the frequency of transmission to a lower frequency to decrease the transmission line losses to the receiver. In this system, the LNA has a gain of 30 dB and $T_e = 150$ K. The down converter has a noise figure of 8 dB and a (conversion) gain of 3 dB. Transmission line losses between antenna site and receiver are 2 dB, and the noise figure of the main receiver is 12 dB.

a) Determine the overall noise figure, and the effective noise temperature, of the receiving system.
b) Repeat part (a) for the case in which the LNA has a gain of 20 dB, all other factors remaining the same.

√★ **4.7.17** Assume that antenna noise in the system in Problem 4.7.16 (including noise entering as a result of antenna sidelobes, etc.) is 80 K, and that the system yields a S/N ratio of 15 dB for the 30 dB LNA gain.

a) Make a plot of receiver S/N (in dB) versus LNA gain over the range 0–30 dB.
b) The LNA specifications cited are at 290 K. In the design specifications it is given that the LNA noise figure (in dB) increases linearly with temperature above 290 K, 0.2 dB for every 10 K increase. Determine the maximum temperature rise that yields a 1 dB degradation in overall S/N ratio as a result of this effect. (Assume all other factors remain constant.)
c) Suppose that the noise figure of the down converter is also a matter of concern in the system design. Calculate the degradation in estimated overall S/N ratio if the

noise figure of the down converter changes from 8 dB to 12 dB, all other factors remaining the same as previously. Calculate this effect for LNA gains of 20 dB and 30 dB. [*Note:* These types of estimates in sensitivities to parameter changes are typical in actual systems designs.]

★ **4.7.18** Determine an equation for $(S/N)_{dB}$ in terms of the receiver figure of merit $(G_r/T_r)_{dB}$ for receiving stations in each of the following geostationary satellite systems designs. In each case, assume a nominal path length between ground station and satellite of 25,000 miles (geostationary altitude is 22,240 miles to the satellite subpoint, which is located in the equatorial plane for fixed antenna pointing).

a) A television relay satellite, designed to relay TV programming material to cable television systems. Assume a bandwidth of 30 MHz, an EIRP of 36 dBW, and a downlink center frequency of 4 GHz. (The uplink frequency is 6 GHz, and the satellite merely acts as a relay.)

b) A direct broadcast satellite (DBS), designed to transmit television broadcasts directly to conventional home receivers. In this case, assume B = 6 MHz, EIRP = 55 dBW, and a center frequency of 12 GHz.

★ **4.7.19** The average current through a thermionic diode can be used to generate shot noise. The dc term is blocked by a capacitor and the output is used to simulate a thermal noise generator.

a) Derive a relation between the average diode current and the effective thermal noise temperature.

b) Assume that the diode current is passed through a resistor R whose value is matched to the input of an amplifier under test with an equivalent noise temperature T_e. When a given amount of diode current flows, it is observed that the output noise power of the amplifier is twice that when the diode current is zero. Derive a relationship for the equivalent noise temperature T_e on the basis of these measurements.

DESIGN PROBLEMS

Design Problem 4.1 Autocorrelation Waveform Design

In a given system used for location and ranging, a binary pulse waveform is transmitted and the receiver measures the arrival time by cross-correlating the input with a stored replica of what was sent. Without noise, this operation is that of autocorrelation. Suppose that we decide to transmit a binary pulse waveform composed of four successive digits, ± 1. Of the 16 different codes we could pick, determine a "good" code such that (1) the magnitude of the autocorrelation for nonzero shifts is as small as possible and, within this constraint, (2) the amplitude at the first shift from zero is nonpositive. [*Note:* If you wish to observe the autocorrelation on a graphics monitor, it is advantageous to add four zeros, use a 8-point FFT, magnitude-square the result, and then take an inverse FFT.]

Design Problem 4.2 Microwave Radiometer

An instrument designed to make measurements of antenna temperature T_a from a satellite at superhigh or extra-high frequencies (e.g., 37 GHz) is called a scanning microwave radiometer. Such an instrument, when pointed toward the earth from satellite orbit, can provide valuable information on distributions of moisture content at or above the earth's surface. Basically, the radiometer consists of an antenna, a waveguide connecting antenna to receiver, and a microwave receiver and power detector. Suppose that at the frequencies of interest the antenna has a loss (attenuation) of 0.5 dB, the waveguide loss is 0.4 dB, and the receiver noise figure is 5 dB.

a) Develop an equation for T_a as a function of the output power reading P_o. Assume that all attenuators are at 290 K.
b) The radiometer is a difficult instrument to calibrate accurately so a comparison-type measurement is preferred. Two antennas are mounted, and one is pointed out toward space to measure a "cold" reference temperature T_c.[†] A switch is placed at the receiver input to switch in one antenna/waveguide at a time. Develop an equation for T_a in terms of T_c and P_o. Assume a loss of 0.8 dB for the switch.
c) Are there other ways you can think of to improve the calibration accuracy of the system?

[†] If you are interested in reading about the calibration procedures for an actual scanning microwave radiometer, see, e.g., P. N. Swanson and A. L. Riley, "The Seasat Scanning Multichannel Microwave Radiometer (SMMR): Radiometric Calibration, Algorithm Development, and Performance," *IEEE Journal of Oceanic Engineering*, OE-5 (1980): 116–124.

CHAPTER

5 Amplitude Modulation

Up to this point, we have worked with signals wherever they happened to be positioned in frequency. In particular, we have not been very concerned with altering signals in time and frequency to accomplish desired objectives. We begin an investigation of methods to alter continuous high-frequency sinusoidal signals in response to given low-frequency signals. This is often referred to as *continuous-wave* (CW) *modulation*. We shall continue this discussion of CW modulation in Chapter 6.

Generally speaking, modulation is that process by which a property or a parameter of a signal is varied in proportion to a second signal. The precise dependence is determined by the type of modulation employed. In amplitude modulation, the amplitude of a sinusoidal signal, whose frequency and phase are fixed, is varied in proportion to a given signal. This alters the given signal by translating its frequency components to higher frequencies.

The use of amplitude modulation may be advantageous whenever a shift in the frequency components of a given signal is desired. This need may arise, for example, in designing filters for very demanding requirements. Or consider the possibility of transmitting a voice communication through space via electromagnetic waves. If the maximum voice frequency is 3 kHz, the minimum wavelength is 100,000 meters. Because antennas with dimensions less than one-quarter wavelength are inefficient, it is a distinct advantage to be able to raise the frequency by at least several orders of magnitude before transmission is attempted.

By proper choices, a large number of signals can be transmitted at the same time without mutual interference. Commercial radio and television stations, for example, are assigned differing frequencies on which to operate and this allows the simultaneous operation of many stations. Because these signals are spaced in frequency, each receiver can easily separate out the desired signal. We shall discuss this more in a later section.

5.1 AMPLITUDE MODULATION: SUPPRESSED CARRIER

The equation of a general sinusoidal signal can be written as

$$\phi(t) = a(t) \cos \theta(t), \tag{5.1}$$

where $a(t)$ is the time-varying amplitude and $\theta(t)$ is the time-varying angle. It is convenient to write $\theta(t) = \omega_c t + \gamma(t)$ so that

$$\phi(t) = a(t) \cos [\omega_c t + \gamma(t)]. \tag{5.2}$$

We assume that $a(t)$ and $\gamma(t)$ are slowly varying compared to $(\omega_c t)$. The term $a(t)$ is called the envelope of the signal $\phi(t)$ and the term ω_c is called the carrier frequency; $\gamma(t)$ is the phase modulation of $\phi(t)$.

In amplitude modulation, the phase term $\gamma(t)$ in Eq. (5.2) is zero (or a constant) and the envelope $a(t)$ is made proportional to the given signal $f(t)$. Letting the constant of proportionality be unity here, we have†

$$\phi(t) = f(t) \cos \omega_c t . \tag{5.3}$$

The term $\cos \omega_c t$ in Eq. (5.3) is called the carrier signal and $f(t)$ is called the modulating signal. The resultant signal, $\phi(t)$, is called the modulated signal.

Applying the modulation property of the Fourier transform to Eq. (5.3), we find the spectral density of $\phi(t)$ is

$$\Phi(\omega) = \tfrac{1}{2}F(\omega + \omega_c) + \tfrac{1}{2}F(\omega - \omega_c) . \tag{5.4}$$

Amplitude modulation therefore translates the frequency spectrum of a signal by $\pm\omega_c$ rad/sec but leaves the spectral shape unaltered, as shown in Fig. 5.1. This type of amplitude modulation is called *suppressed-carrier* because the spectral density of $\phi(t)$ has no identifiable carrier in it, although the spectrum is centered at the frequency ω_c.

From Fig. 5.1(f), we see that both the positive and the negative frequency content of $f(t)$ are now displayed as positive frequencies. This implies that the bandwidth of

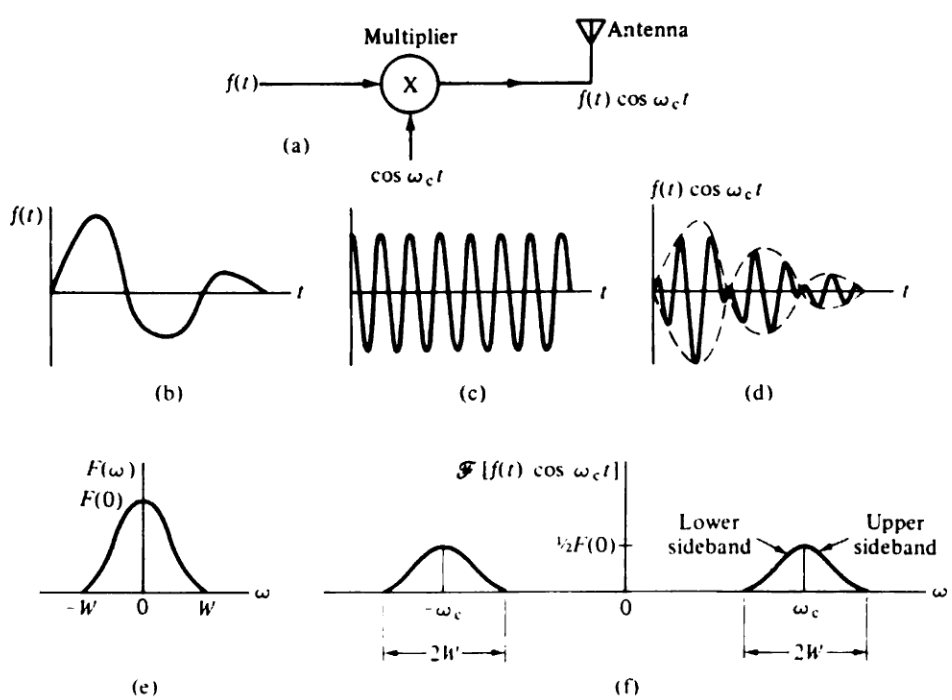


Figure 5.1 An amplitude modulation suppressed-carrier transmission system.

† Including a multiplier constant k_a, this can be written as $k_a f(t) \cos \omega_c t$.

$f(t)$ is doubled when this type of amplitude modulation is used, as we have indicated in Fig. 5.1(e), (f). In Fig. 5.1(f), the spectral content for positive frequencies above ω_c is called the *upper sideband* of $\phi(t)$ and the spectral content for positive frequencies below ω_c is called the *lower sideband* of $\phi(t)$. It is easily seen that, for positive frequencies, the upper sideband of $\phi(t)$ displays the positive frequency components of $f(t)$ and the lower sideband of $\phi(t)$ displays the negative frequency components of $f(t)$. Similar relations, though just reversed, hold for the negative frequencies in $\phi(t)$.

From this discussion, we conclude that amplitude modulation with suppressed carrier provides us with a convenient means to observe the complete frequency spectrum of a given signal $f(t)$. All we must do is translate the signal using a carrier frequency ω_c which is higher than the spectral bounds (bandwidth) of the signal. This principle is widely used in spectral analysis. The fact that we have two sidebands — an upper and a lower — and no separate carrier present in $\phi(t)$ suggests the following convenient designation for this type of modulation: *double-sideband, suppressed-carrier* — abbreviated DSB-SC.

Recovery of the original signal $f(t)$ from the DSB-SC signal $\phi(t)$ requires another translation in frequency to shift the spectrum to its original position. The process of retranslation of the spectrum to its original position in frequency is called demodulation or detection.

Because the modulation property of the Fourier transform proved useful in translating spectra for modulation, we try it again for demodulation. Assuming that $\phi(t) = f(t) \cos \omega_c t$ is the transmitted signal, we have

$$\begin{aligned} \phi(t) \cos \omega_c t &= f(t) \cos^2 \omega_c t \\ &= \tfrac{1}{2} f(t) + \tfrac{1}{2} f(t) \cos 2\omega_c t. \end{aligned} \tag{5.5}$$

Taking the Fourier transform of both sides of Eq. (5.5) and using the modulation property, we get

$$\mathcal{F}\{\phi(t) \cos \omega_c t\} = \tfrac{1}{2} F(\omega) + \tfrac{1}{4} F(\omega + 2\omega_c) + \tfrac{1}{4} F(\omega - 2\omega_c). \tag{5.6}$$

This result can also be obtained by recalling that the operation of multiplication in the time domain is equivalent to that of convolution in the frequency domain. Thus the spectrum of $\phi(t) \cos \omega_c t$ can be obtained by convolving the spectrum of the received signal $\phi(t)$ with that of $\cos \omega_c t$ (i.e., with impulses at $\pm\omega_c$). Both approaches yield the same result and the spectrum is shown in Fig. 5.2(e). A low-pass filter is required to separate out the double-frequency terms from the original spectral components, as indicated in Fig. 5.2. Obviously, we require that $\omega_c > W$ for proper signal recovery, a requirement easily satisfied in practice.

The principles developed above for DSB-SC modulation hold for any general modulating signal $f(t)$ as long as it is slowly varying with respect to the carrier frequency; that is, $W < \omega_c$. It is instructive to consider the special case of a single-frequency sinusoidal modulating signal to better understand the concepts. Several descriptive waveforms for this case are shown in Fig. 5.3.

It is evident from Fig. 5.3 that it is difficult for an observer to recognize $f(t)$ from an observation of the modulated signal $\phi(t)$ unless there is some knowledge of the

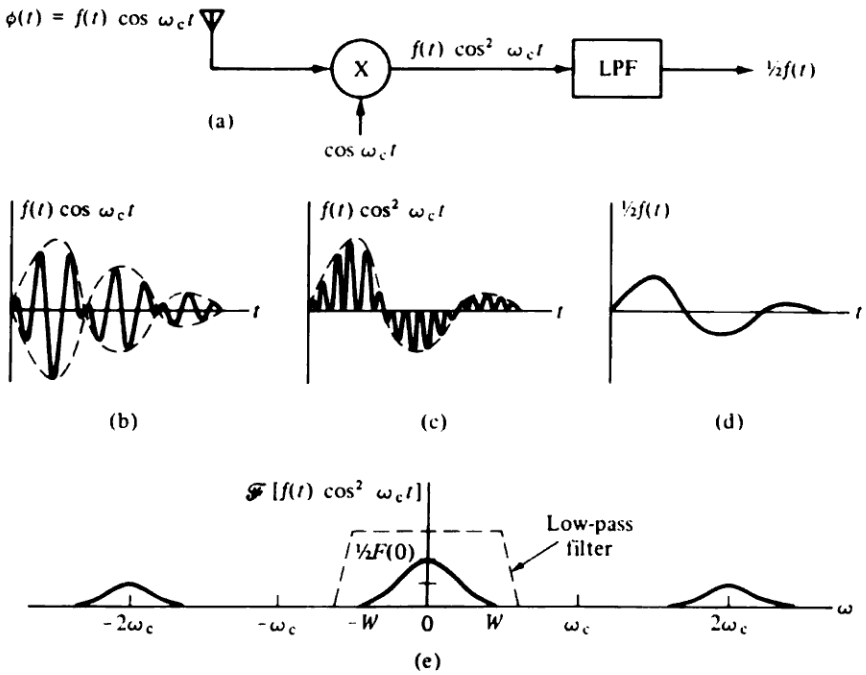


Figure 5.2 Demodulation of amplitude modulation suppressed-carrier signals.

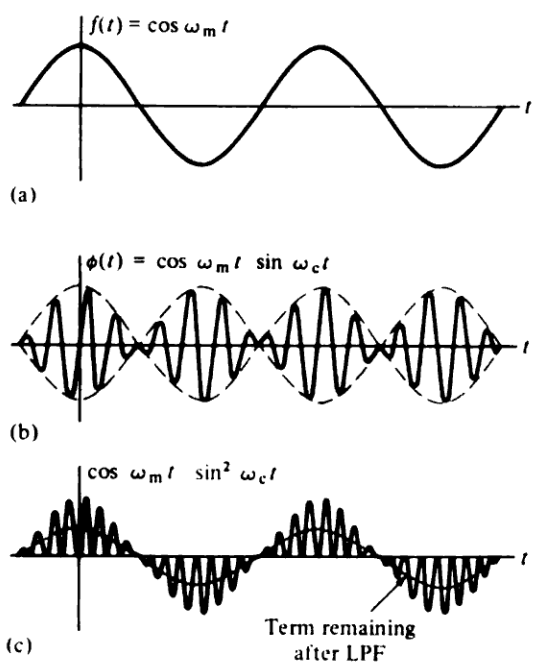


Figure 5.3 The DSB-SC modulation and demodulation of a sinusoidal signal.

phase relations involved. An observation of the amplitude alone is not sufficient. To be more precise, the relative phase of the modulated signal $\phi(t)$ conveys the sign of $f(t)$, as evidenced by the 180° phase reversals when $f(t)$ goes through zero. We conclude from this that both the correct phase and frequency must be known to correctly demodulate DSB-SC waveforms.

To confirm this conclusion, let us introduce a small frequency error, $\Delta\omega$, and a phase error, θ_0, in the locally generated carrier signal at the receiver. The receiver then forms the product,

$$\begin{aligned}\phi(t)\cos[(\omega_c + \Delta\omega)t + \theta_0] &= f(t)\cos\omega_c t\cos[(\omega_c + \Delta\omega)t + \theta_0]\\ &= \tfrac{1}{2}f(t)\cos[(\Delta\omega)t + \theta_0]\\ &\quad + \tfrac{1}{2}f(t)\cos[(2\omega_c + \Delta\omega)t + \theta_0]. \qquad (5.7)\end{aligned}$$

The second term on the right-hand side of Eq. (5.7) is centered at $\pm(2\omega_c + \Delta\omega)$ and is filtered out by the low-pass filter. The output of this filter will then be given by the remaining term in Eq. (5.7).

$$e_o(t) = \tfrac{1}{2}f(t)\cos[(\Delta\omega)t + \theta_0]. \qquad (5.8)$$

Thus the output signal is not $\frac{1}{2}f(t)$, as we had before, unless both $\Delta\omega$ and θ_0 are zero (i.e., no frequency or phase error).

We shall consider two special cases of phase and frequency errors here. In the first, let $\Delta\omega = 0$ so that there is only a phase error.

$$e_o(t) = \tfrac{1}{2}f(t)\cos\theta_0.$$

The phase error in the local carrier causes a variable gain factor in the output signal that is proportional to the cosine of the phase error. For small fixed phase errors, this is quite tolerable. For phase errors approaching ±90°, however, the received signal is wiped out. In some cases, the phase error varies randomly, resulting in unacceptable performance.

In the second special case, let $\theta_0 = 0$ so that there is only a frequency error. Equation (5.8) simplifies to

$$e_o(t) = \tfrac{1}{2}f(t)\cos(\Delta\omega)t.$$

Instead of recovering the original signal $f(t)$, we obtain the signal $f(t)$ multiplied by a low-frequency sinusoid. This results in undesirable and unacceptable distortion.

The effects of both frequency errors and random phase errors render this demodulation of the signal unsatisfactory. It is necessary, therefore, to have synchronization in both frequency and phase between the transmitter and the receiver when DSB-SC modulation is used. The synchronization of the carrier signals presents no major problem when the transmitter and the receiver are in close proximity. Methods for overcoming this limitation when the receiver is remotely located will be discussed in a later section.

Recovering the original signal $f(t)$ from the modulated signal $\phi(t)$ using a synchronized oscillator is called *synchronous* detection, or *coherent* detection. If only

one synchronized oscillator is used (as we have done), it is also called *homodyne* detection.†

EXAMPLE 5.1.1

Using the orthogonality of sines and cosines makes it possible to transmit and receive two different signals simultaneously on the same carrier frequency. A scheme for doing this, known as *quadrature multiplexing*, is shown in Fig. 5.4. Show that each signal can be recovered by synchronous detection of the received signal using carriers of the same frequency but in phase quadrature.

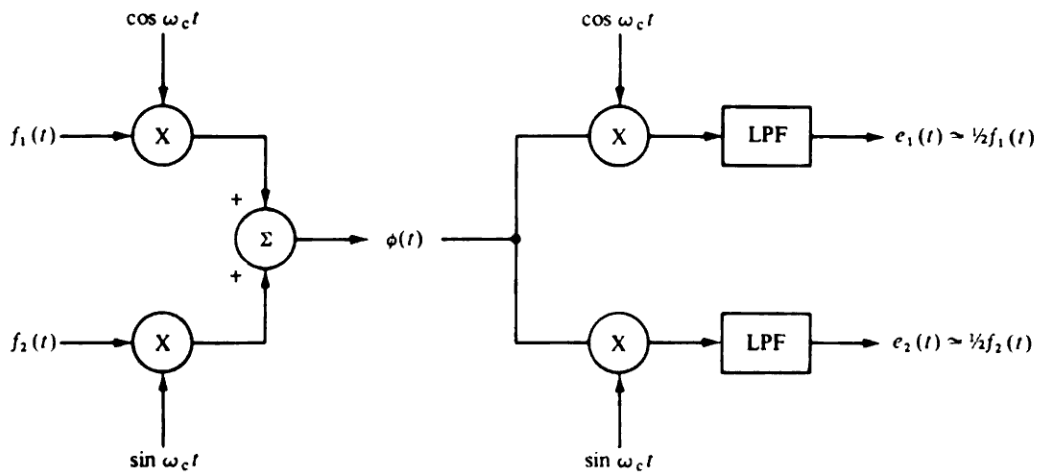


Figure 5.4 The quadrature multiplexing system described in Example 5.1.1.

Solution

$$\phi(t) = f_1(t)\cos\omega_c t + f_2(t)\sin\omega_c t,$$

$$\begin{aligned}\phi(t)\cos\omega_c t &= f_1(t)\cos^2\omega_c t + f_2(t)\sin\omega_c t\cos\omega_c t\\ &= \tfrac{1}{2}f_1(t) + \tfrac{1}{2}f_1(t)\cos 2\omega_c t + \tfrac{1}{2}f_2(t)\sin 2\omega_c t,\\ \phi(t)\sin\omega_c t &= f_1(t)\cos\omega_c t\sin\omega_c t + f_2(t)\sin^2\omega_c t\\ &= \tfrac{1}{2}f_1(t)\sin 2\omega_c t + \tfrac{1}{2}f_2(t) - \tfrac{1}{2}f_2(t)\cos 2\omega_c t.\end{aligned}$$

In the low-pass filter, all terms at $2\omega_c$ are attenuated, yielding

$$e_1(t) = \tfrac{1}{2}f_1(t),$$

$$e_2(t) = \tfrac{1}{2}f_2(t).$$

† A discussion of homodyne detection, with specific applications to measurement systems, can be found in: R. J. King, *Microwave Homodyne Systems*, Stevenage, Herts (England): Peter Peregrinus Ltd., 1978.

Quadrature multiplexing is an efficient method of transmitting two message signals within the same bandwidth. It requires precise phase synchronization of transmitter and receiver, as shown by the following drill problem. Quadrature multiplexing is used in the transmission of color (chroma) information signals in commercial television broadcasts. (See Appendix D for a description of commercial television signal transmissions.)

DRILL PROBLEM 5.1.1

Practical use of the quadrature multiplexing scheme of Example 5.1.1 is limited by the rigid demands on the phase synchronization required. For example, let the local carrier generated in the receiver in Fig. 5.4 have a phase error of θ_0 radians. Derive an expression for $e_1(t)$ if θ_0 is small; also estimate how small θ_0 must be to keep the interference caused by $f_2(t)$ at least -40 dB (1%) of $f_1(t)$ if $f_1(t)$, $f_2(t)$ are of equal magnitude.

ANSWER: $e_1(t) \approx \frac{1}{2}[f_1(t) \pm \theta_0 f_2(t)]$; $\pm 0.57°$.

DRILL PROBLEM 5.1.2

Show that $f_1(t)$ and $f_2(t)$ essentially control both the amplitude and the phase of $\phi(t)$ in the quadrature multiplexing system of Example 5.1.1. In other words, derive an expression for $\phi(t)$ which exhibits the envelope and phase angle in terms of $f_1(t)$ and $f_2(t)$. [*Hint:* Recall a Fourier series conversion.] Also find $\phi(t)$ for the special case where $f_1(t) = \cos\theta$, $f_2(t) = \sin\theta$.

ANSWER: $\phi(t) = \sqrt{f_1^2(t) + f_2^2(t)} \cos\{\omega_c t + \tan^{-1}[-f_2(t)/f_1(t)]\}$; $\cos(\omega_c t - \theta)$.

5.1.1 Generation of DSB-SC Signals

As a result of the modulation property of the Fourier transform, we have observed that the spectrum of any signal may be translated $\pm\omega_c$ rad/sec in frequency by multiplying it with a sinusoidal frequency of ω_c rad/sec. It is not necessary, however, that we multiply it with a waveform that is sinusoidal. In fact, we shall now demonstrate that a signal spectrum can be translated an amount $\pm\omega_c$ by multiplying the signal with *any* periodic waveform whose fundamental frequency is ω_c rad/sec. Our reasoning here will closely parallel that followed in the discussion of the sampling theorem in Chapter 3.

Any periodic signal $p_T(t)$ with finite average power can be represented by the Fourier series,

$$p_T(t) = \sum_{n=-\infty}^{\infty} P_n e^{jn\omega_0 t}, \tag{5.9}$$

where $\omega_0 = 2\pi/T$. Choosing $\omega_0 = \omega_c$ and multiplying Eq. (5.9) by $f(t)$ gives us

$$f(t)p_T(t) = \sum_{n=-\infty}^{\infty} P_n f(t) e^{jn\omega_c t}. \tag{5.10}$$

Applying the frequency translation property of the Fourier transform to Eq. (5.10), we have

$$\mathcal{F}\{f(t)p_T(t)\} = \sum_{n=-\infty}^{\infty} P_n F(\omega - n\omega_c). \tag{5.11}$$

From this, we see that the spectrum of $f(t)p_T(t)$ contains $F(\omega)$ and $F(\omega)$ translated by $\pm\omega_c, \pm 2\omega_c, \ldots, \pm n\omega_c, \ldots$. The amplitudes of these successive spectral replicas of $F(\omega)$ are scaled by the constants $P_0, P_1, P_2, \ldots, P_n, \ldots$. This is illustrated in Fig. 5.5 using a periodic square wave. These results can also be demonstrated using the frequency convolution property of the Fourier transform.

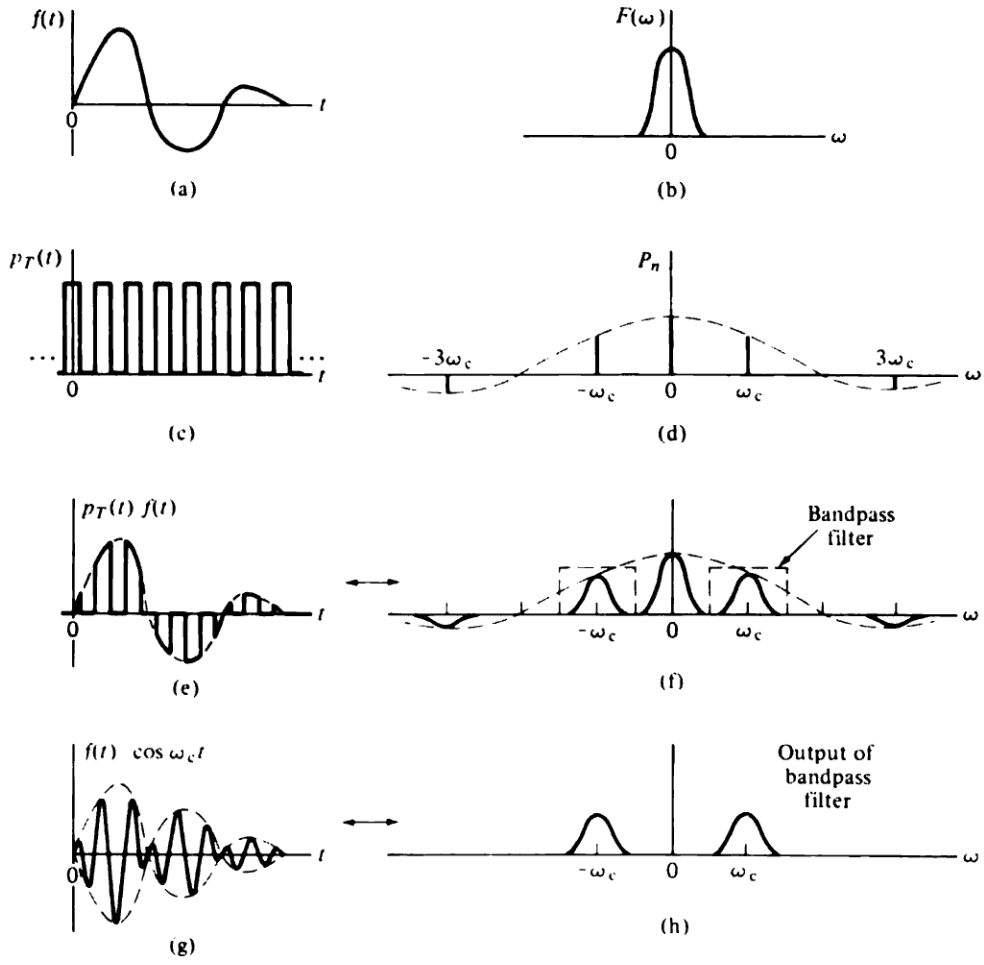


Figure 5.5 Amplitude modulation using multiplication by periodic functions.

In amplitude modulation, we are interested only in the spectrum centered around $\pm\omega_c$. This can be obtained by using a bandpass filter following the multiplication. The bandpass filter allows those frequency components centered at $\pm\omega_c$ to pass and attenuates all other frequency components. If ω_c is chosen much larger than W, the bandwidth of $f(t)$, the undesired frequency components are widely separated from the desired frequency components and simple filters can be used.

The process of frequency translation is also called *frequency conversion, frequency mixing,* and *heterodyning*. Systems that perform this operation are called frequency converters or frequency mixers. In order to translate a spectrum, such systems must produce new frequencies that are different from those present in the input signal. However, the response of linear time-invariant systems cannot contain frequencies other than those present in the input signal (cf. Section 2.12). Therefore we conclude that to produce modulation we must use systems that are either time-varying or nonlinear (or both).

This fact can be used to advantage in the testing of linear amplifiers. The sum of two sinusoids, say, at 400 Hz and 1 kHz, is applied to the input. Any output at frequencies other than 400 Hz or 1 kHz indicates the presence of nonlinearities in the amplifier. In particular, the output at 1400 Hz (or 600 Hz) is the amount of modulation present which in turn is a relative measure of the nonlinearities. This test is called an *intermodulation test* and is used in the testing of high-fidelity audio equipment.

EXAMPLE 5.1.2

A periodic signal consists of the exponentially decreasing waveform e^{-at}, $0 \le t < T$, repeated every T seconds. A given signal $f(t)$ is multiplied by this periodic signal. Determine an expression describing the spectrum and the time waveform of the resulting amplitude-modulated signal if all components except those centered at $\pm\omega_c$, $\omega_c = 2\pi/T$, are discarded.

Solution We can write a Fourier series,

$$p_T(t) = \sum_{n=-\infty}^{\infty} P_n e^{jn\omega_c t}$$

where

$$P_n = \frac{1}{T}\int_0^T e^{-at}e^{-jn\omega_c t}\,dt = \frac{1}{T}\frac{1 - e^{-aT}}{a + jn\omega_c}.$$

Using Eq. (5.11), the spectrum of the product $p_T(t)f(t)$ is

$$\frac{1}{T}(1 - e^{-aT})\left[\frac{1}{a + j\omega_c}F(\omega - \omega_c) + \frac{1}{a - j\omega_c}F(\omega + \omega_c)\right].$$

The corresponding terms in the Fourier series are

$$\frac{1}{T}(1 - e^{-aT})\left[\frac{1}{a + j\omega_c} f(t)e^{j\omega_c t} + \frac{1}{a - j\omega_c} f(t)e^{-j\omega_c t}\right].$$

Combining [or using Eqs. (2.57), (2.60), and (2.61)] yields the time waveform

$$\frac{2}{T}\frac{1 - e^{-aT}}{\sqrt{a^2 + \omega_c^2}} f(t) \cos(\omega_c t + \theta_0),$$

where

$$\theta_0 = \tan^{-1}(-\omega_c/a),$$

$$\omega_c = 2\pi/T.$$

DRILL PROBLEM 5.1.3

Repeat Example 5.1.2 if the periodic signal $p_T(t)$ is a sequence of unit impulse functions spaced T seconds apart with one impulse function at $t = 0$. Check your result using frequency convolution.

ANSWER: $\frac{1}{T}F(\omega - \omega_c) + \frac{1}{T}F(\omega + \omega_c)$; $\quad \frac{2}{T}f(t) \cos \omega_c t$.

The Chopper Modulator

As an example of a modulator that operates as a linear time-varying system, consider a system whose gain (i.e., transfer function) is varied with time. The simplest example of such a system is a switch that is opened and closed at a given rate or frequency, ω_c. This type of system is often descriptively called a "chopper" and is illustrated in Fig. 5.6(a).

The operation of the chopper type of modulator shown in Fig. 5.6(a) can be explained briefly as follows. The switch alternates between the open and closed states, driven by a source at a frequency of ω_c rad/sec. We shall assume that the transition time between contacts is negligible. For one-half the period, the switch allows the relatively slowly varying signal $f(t)$ to be applied to the filter input. During the next half-period, the input to the filter is grounded. Therefore the signal waveform at the filter input has literally been chopped at a frequency of ω_c rad/sec. This chopping operation is essentially a multiplication of the signal $f(t)$ with a periodic square wave $p_T(t)$ whose two levels are 0 and 1. As discussed earlier, such a chopped waveform contains the spectrum of $f(t)$ translated in frequency by multiples of ω_c. The desired amplitude modulated signal $\phi(t)$ may be obtained by passing this chopped signal through a bandpass filter centered at $\pm\omega_c$. Note that the sampling theorem must be satisfied to allow the proper filtering of the waveform.

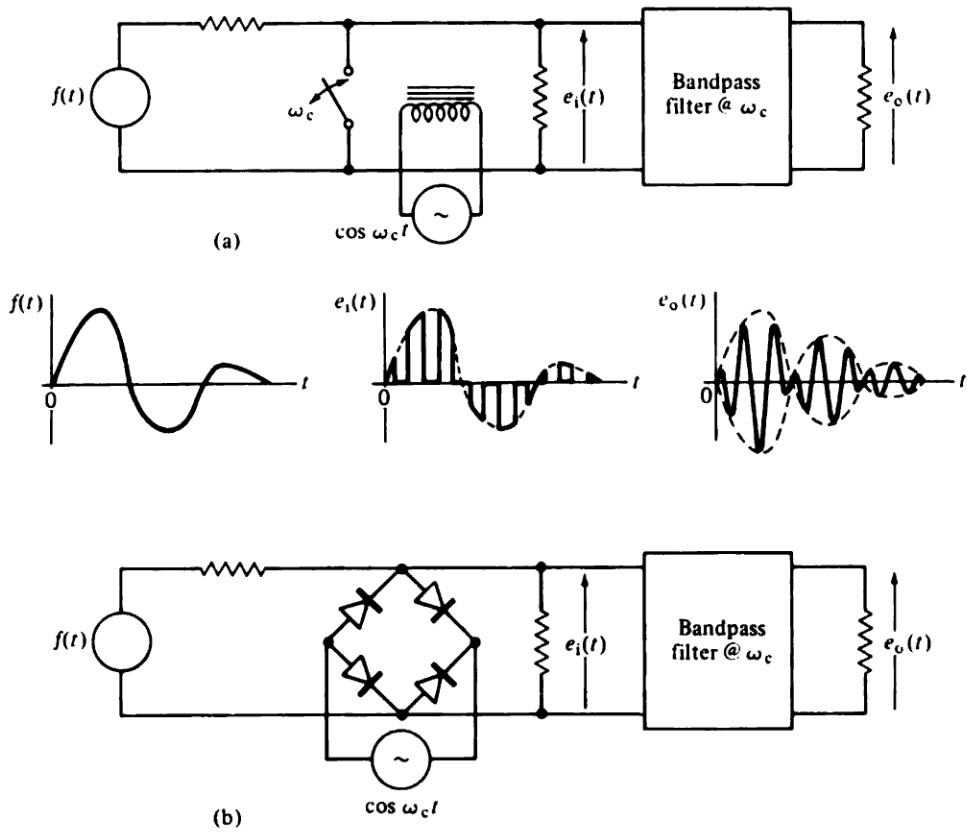


Figure 5.6 Schematic diagram of a chopper modulator: (a) using an electromechanical switch and (b) using diodes as switches.

The chopper modulator is widely used in practical DSB-SC modulation systems. At low frequencies, the chopper may be a magnetically driven switch. Photo diodes, FET's, and mechanically interrupted light beams are also used. At higher frequencies diodes are usually used as switches. A typical circuit using four diodes, sometimes known as a ring modulator, is shown in Fig. 5.6(b). It is easily seen that these diodes alternately switch the filter input between the input signal and ground at the frequency of ω_c rad/sec. This is the same action as that of the switch used above and the analysis follows in exactly the same manner. For best operation, the diodes should be identical (i.e., balanced) so that the bridge circuit is balanced when $f(t) = 0$ and there is no carrier output. Several configurations of the ring modulator are in common usage.

A popular configuration, known as the double-balanced (i.e., full-wave balanced) ring modulator, is shown in Fig. 5.7. Popularity of this modulator is attributable to the fact that it does not require ideal components, provided that their characteristics are matched, and the output is often usable without additional filtering.

The operation of the double-balanced ring modulator is briefly described as follows. First we assume that $e_i(t) = 0$. On the positive half-cycles of $\cos \omega_c t$, diodes D_1 and D_2 will conduct and point a is connected to the input through half of the

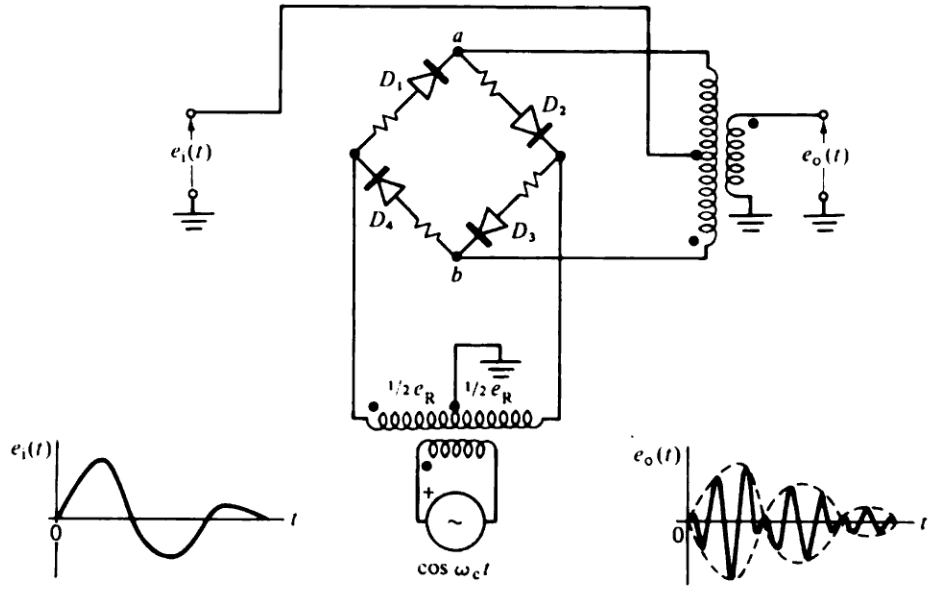


Figure 5.7 The double-balanced ring modulator.

output-transformer secondary. If the secondary of the reference transformer is accurately center-tapped and if the impedances of D_1 and D_2 are identical, no current will flow through the output transformer and no voltage will be developed at the output. During the next half-cycle, diodes D_3 and D_4 conduct and point b is connected to the input through the opposite half of the output-transformer secondary. Again, no current will flow and no output results. Thus we see that the carrier is indeed suppressed; i.e., the modulator is balanced.

Now let us apply a positive-polarity input signal whose peak amplitude is much smaller than that of the reference e_R. On the positive half-cycles of cos $\omega_c t$, point a is essentially at ground potential and a current will flow upward through half of the output-transformer secondary, inducing a positive output voltage. On the negative half-cycles of cos $\omega_c t$, point b is essentially at ground potential and a current will flow downward through the opposite half of the output-transformer secondary, inducing a negative output voltage. The peak positive and negative output voltages will be identical for a given fixed input-signal amplitude if the output transformer is also accurately center-tapped. Therefore we have developed a signal that alternates in sign at a rate determined by the carrier (reference) frequency and whose amplitude is proportional to the input signal amplitude. It can be seen that the phase of the output changes by 180° for negative-polarity input signals. The output signal, then, is an example of DSB-SC modulation.

The description above assumed a reference signal large enough to control the diode switching. Referring to Fig. 5.7, we see this requires that

$$[e_i(t)]_{\max} < \tfrac{1}{2}[e_R(t)]_{\max}.$$

In practice, an imbalance between diode characteristics and inaccuracies in transformer center-taps results in nonideal performance and some carrier does appear in

the output unless balancing controls are added. The transformer impedances are matched to the geometric mean of the forward and back resistances of the diodes for best performance. The use of hot-carrier diodes can result in good carrier suppression without a need for balancing controls.

Use of Nonlinear Devices

Amplitude modulation can also arise in nonlinear systems. A common nonlinearity in electrical devices is a nonlinearity between the voltage and the current. A semiconductor diode is a good example of such a device; tubes and transistors also exhibit similar characteristics under large-signal conditions.

A possible arrangement for the use of nonlinear devices to obtain DSB-SC modulation is shown in Fig. 5.8. Approximating the nonlinearity with a power series, we get†

$$i(t) = a_1 e(t) + a_2 e^2(t) + a_3 e^3(t) + \cdots .$$

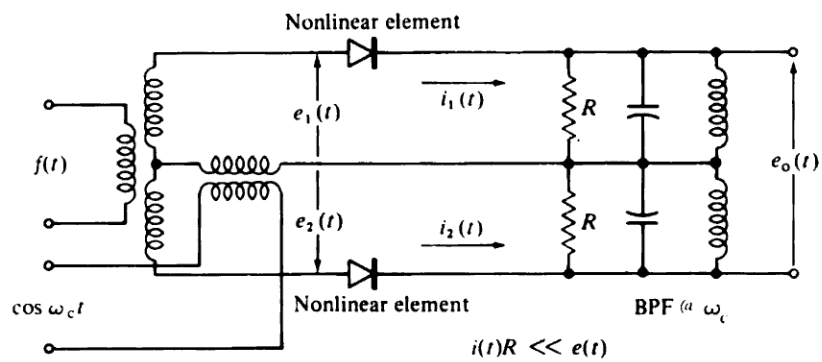


Figure 5.8 The balanced modulator using nonlinear devices.

Referring to Fig. 5.8, we find

$$e_1(t) = \cos \omega_c t + f(t),$$
$$e_2(t) = \cos \omega_c t - f(t).$$

Retaining the first two terms in the power series for each nonlinearity and assuming that the nonlinearities are identical gives us

$$i_1(t) = a_1[\cos \omega_c t + f(t)] + a_2[\cos \omega_c t + f(t)]^2,$$
$$i_2(t) = a_1[\cos \omega_c t - f(t)] + a_2[\cos \omega_c t - f(t)]^2.$$

† A discontinuity may exist if a piecewise linear model is chosen. The discontinuity prevents the use of a power-series expansion. In this case, the nonlinear characteristic can be approximated by a polynomial series following the methods used in Problem 2.5.4.

For a resistive load, the net voltage is $[i_1(t) - i_2(t)]R$, or

$$[i_1(t) - i_2(t)]R = 4a_2R\left[f(t)\cos\omega_c t + \frac{a_1}{2a_2}f(t)\right]. \tag{5.12}$$

The second term on the right-hand side of Eq. (5.12) can be filtered out by using a bandpass filter centered at $\pm\omega_c$, yielding the desired output.

Semiconductor diodes may be used for the nonlinear devices in such modulators, as shown in Fig. 5.8. The performance of this type of modulator, as for the ring modulator, is dependent on how close the characteristics of the diodes can be matched. All of the modulators discussed above generate DSB-SC signals and are known as *balanced modulators*.

DRILL PROBLEM 5.1.4

The model of a possible DSB-SC modulator is shown in Fig. 5.9. Determine the required value of the constant k if the bandpass filter has unity gain at $\pm\omega_c$.

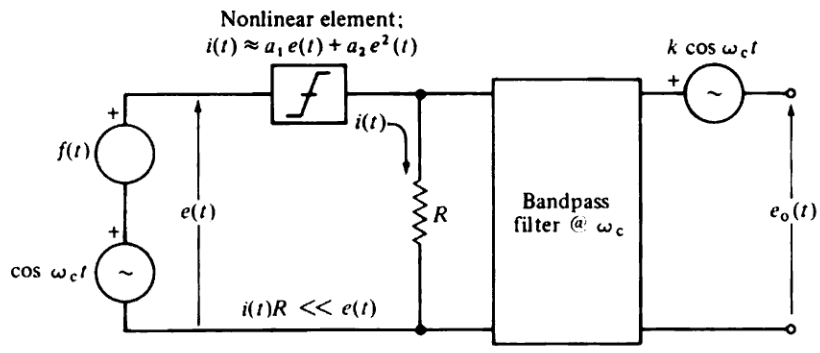


Figure 5.9 The modulator described in Drill Problem 5.1.4.

ANSWER: $k = a_1R$.

5.1.2 Demodulation (Detection) of DSB-SC Signals

To recover the original signal $f(t)$ from the modulated signal $\phi(t)$, we must translate the spectrum again. We have seen that this can be achieved by multiplying $\phi(t)$ by $\cos\omega_c t$ (i.e., synchronous detection). Therefore the same circuits as those used for modulation can be used for demodulation with only minor differences.

The first difference arises because the desired output spectrum of the modulator was centered about the carrier frequency. A bandpass filter centered at $\pm\omega_c$ was used to reject all other frequency components. In the demodulator, however, the desired

output spectrum is centered about $\omega = 0$ and therefore a low-pass filter is needed at the output.

A second difference arises from the fact that the oscillator in the demodulator must be synchronized to the oscillator in the modulator to achieve proper demodulation. This is usually accomplished by either a direct connection if the modulator and demodulator are in close proximity or by supplying a sinusoid displaced in frequency but related to the modulator-oscillator frequency. The sinusoid in the latter method is called a "pilot carrier." We shall have occasion to illustrate both methods.

The Chopper Amplifier

It is difficult to construct amplifiers with dc or very low frequency response yet with negligible drift and reasonable-size components. The principle of DSB-SC modulation can be used to advantage to shift the spectrum of the input signal to higher frequencies where it can be easily amplified. If a low-frequency output is desired, a synchronous demodulator can be used because the modulation oscillator signal is readily available. The chopper-type modulator is generally used for this application.

A typical chopper amplifier is shown in Fig. 5.10(a). The chopper grounds the input signal $f(t)$ for every half-cycle, producing the waveform and spectral density of Fig. 5.10(c). The bandpass amplifier amplifies those spectral components about $\pm\omega_c$, producing the output shown in Fig. 5.10(d). The chopper grounds the amplifier output every half-cycle, synchronously demodulating the output signal. As shown, the grounding half-cycles at the input and output are complementary. If a 180° phase reversal in the amplifier is assumed, the output signal is an amplified replica of the input. Note that the sampling theorem must be obeyed to allow adequate filtering.

The primary advantage of the chopper amplifier is that it allows one to build very stable high-gain amplifiers. As an example, a simplified diagram of a low-level, chopper-stabilized voltmeter is shown in Fig. 5.11. Separate choppers for input and output can be used if they are synchronized.

The use of DSB-SC modulation in the chopper amplifier is advantageous mainly because both the modulator and the demodulator are in close proximity and can operate from a common source. When the demodulator is remotely located, a problem arises in deriving the necessary phase and frequency synchronization. The use of a pilot carrier to facilitate this is discussed next.

Pilot Carrier Systems

A method commonly used in DSB-SC modulation to maintain synchronization between modulator and demodulator is to send a sinusoidal tone whose frequency and phase are related to the carrier frequency. This tone is generally sent outside the passband of the modulated signal so it will not alter the frequency response capability of the system. A tuned circuit in the receiver detects the tone, translates it to the proper frequency, and uses it to correctly demodulate the DSB-SC signal.

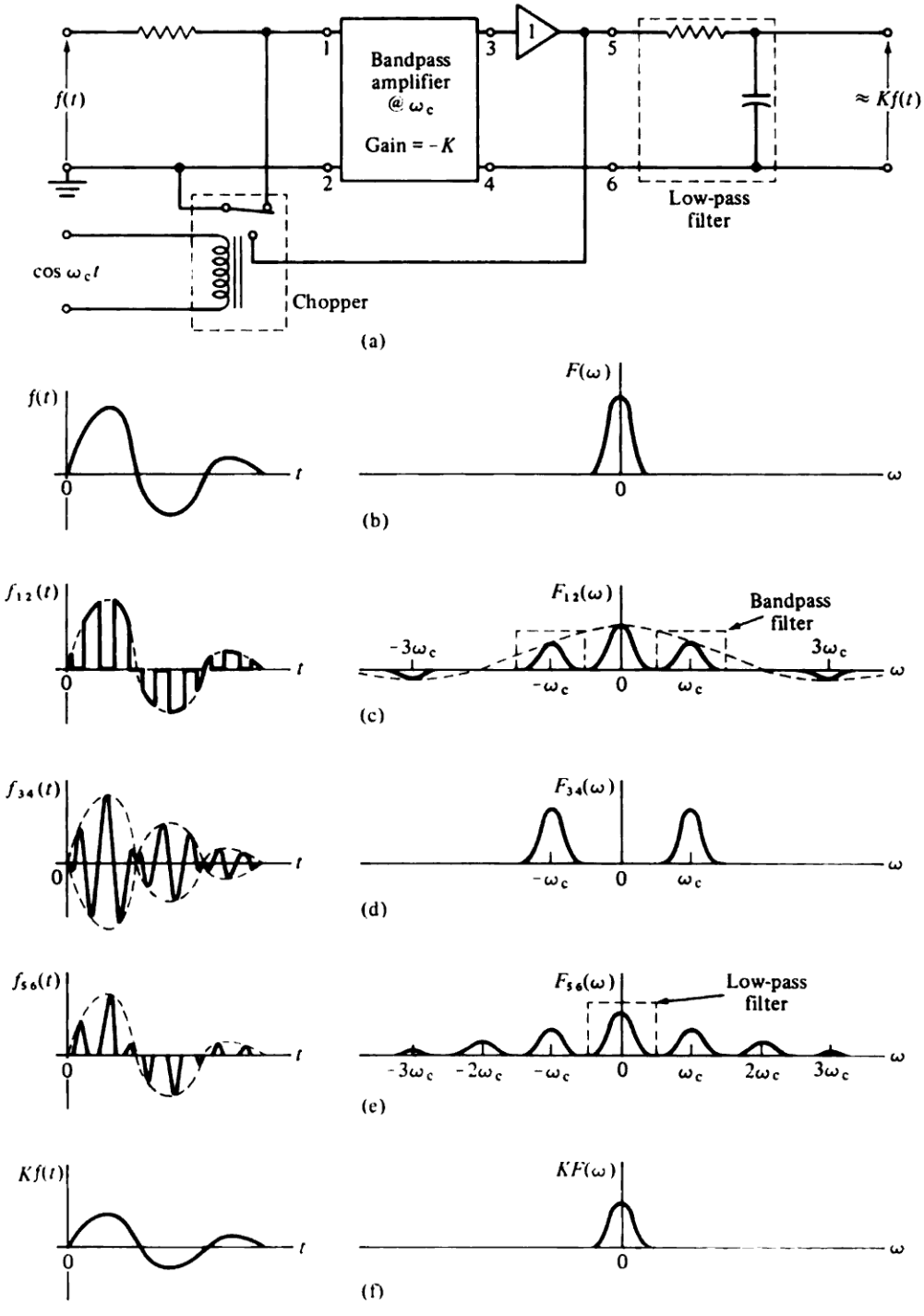


Figure 5.10 The chopper amplifier.

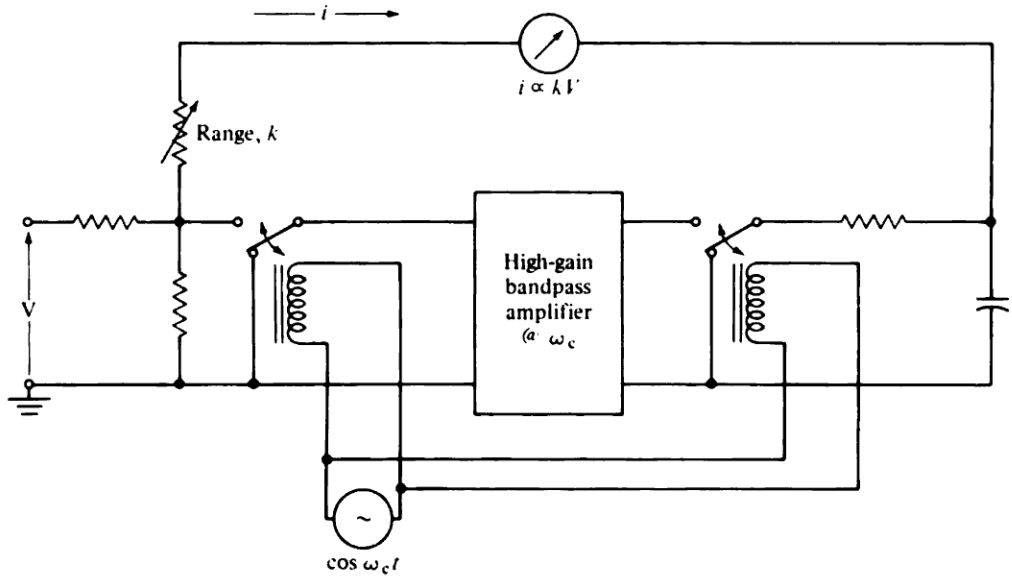


Figure 5.11 A chopper-stabilized voltmeter.

The stereo multiplexing used in commercial frequency modulation (FM) stations is an example of a pilot tone system.† An FM station is permitted the use of 75 kHz before using frequency modulation to transmit the composite signal. Upon reception, two demodulators are used in tandem, one for the FM, the other to separate the audio signals. We shall consider frequency modulation later; here we are concerned only with the stereo multiplexing.

In stereo broadcasting, it is necessary to transmit and receive both the left (L) and the right (R) audio channels while also furnishing the sum (L + R) to monophonic receivers. To serve both types of receivers, the (L + R) signal occupies the normal audio spectrum up to 15 kHz and the (L − R) signal, also in the same spectral range, is shifted up in frequency using DSB-SC modulation. A system designed to accomplish this is shown in Fig. 5.12(a). The carrier frequency used is 38 kHz and a sketch of the resulting composite spectrum is shown in Fig. 5.12(b). This leaves the frequency band 15–23 kHz unoccupied and a pilot carrier is placed at 19 kHz. In the receiver, a narrow bandpass filter centered at 19 kHz picks out the pilot carrier.‡ This frequency is doubled to 38 kHz and used to synchronize the demodulator. Finally, an addition and subtraction of the two signals (this is known as matrixing) yields the desired L and R audio signals.

5.1.3 Uses in Filter Design

Although the emphasis in our discussion of DSB-SC modulation thus far has been on signals, the same principles can be used to advantage in the design of specialized

† Frequency modulation is discussed in Chapter 6.

‡ Detection of this 19-kHz pilot carrier is used to switch on the stereo indicator light on the receiver panel.

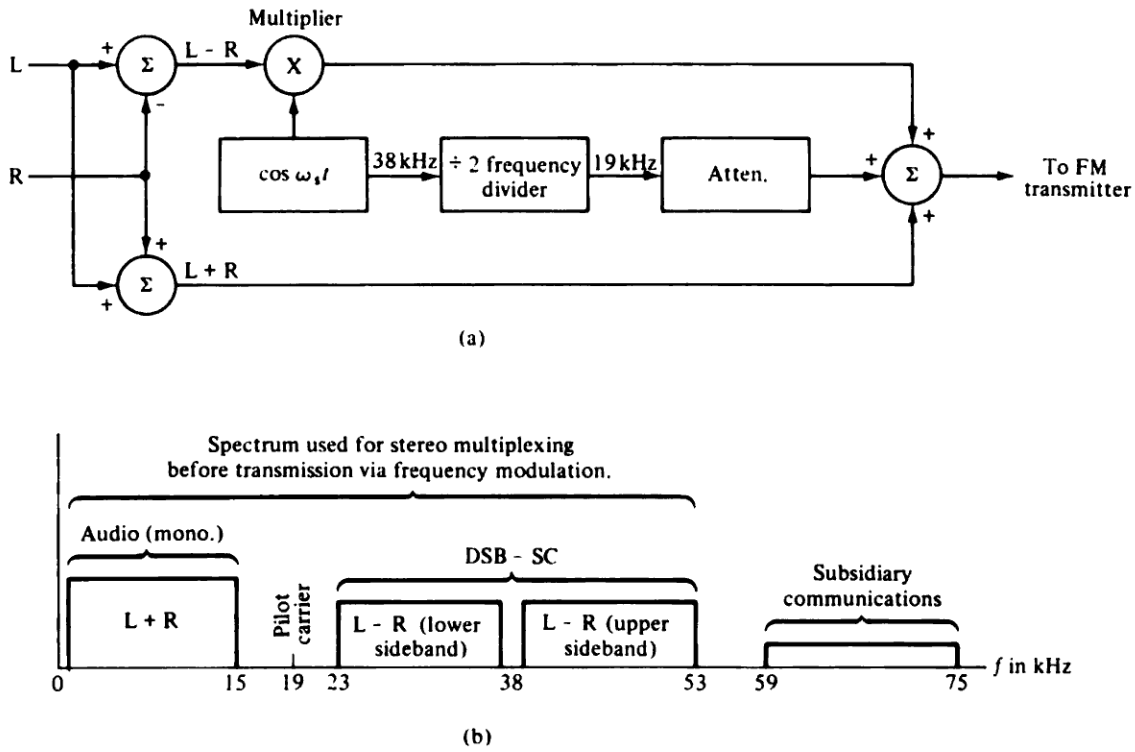


Figure 5.12 (a) Block diagram and (b) spectrum of a stereo-multiplex system.

filters. An important class of filters using these principles is that of the electrically controlled filter.

A problem arising in the design of some systems is that of a bandpass filter whose center frequency is movable and controllable. The design of such filters by the usual filter design methods often proves difficult. An attractive alternative is to use a fixed filter and to move the signal frequency components instead, using DSB-SC modulation. Let us examine two examples.

The Phase-Locked Loop

A question that arises in pilot tone systems is What is the best way to synchronize one sinusoid to another? Narrow bandpass filters are a possibility but they tend to drift in tuning and are not optimum in rejecting noise. A well-known solution to the synchronization problem is the so-called *phase-locked loop* (PLL).

The basic phase-locked loop consists of a signal multiplier (i.e., a balanced modulator), a low-pass filter, and a voltage-controlled oscillator (VCO), as shown in Fig. 5.13. Operation of the phase-locked loop when the VCO frequency is near the incoming frequency is as follows. The incoming sinusoidal signal, $\cos \omega_c t$, is multiplied by the output of the voltage-controlled oscillator (VCO). The low-frequency

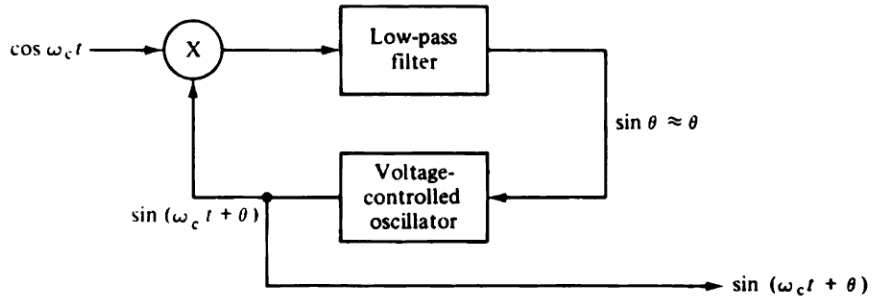


Figure 5.13 The phase-locked loop.

component of the output of the multiplier is a voltage whose magnitude and sign are proportional to the phase difference, for small differences, between the incoming sinusoid and the VCO. This voltage is used to control the VCO and the loop attempts to keep the phase difference small (ideally, zero) between the VCO signal and the incoming signal. The objective of synchronization between the two signals has thus been accomplished.

The bandwidth of the phase-locked loop is controlled by the bandwidth of the low-pass filter and can be made very small to reject noise. In practice, some compromises have to be made, however, because the loop may never find the signal if the bandwidth is too small. Generally the loop will pull into lock within one bandwidth. The operation of the phase-locked loop when it is not locked is very nonlinear, however, and we shall not continue farther into the subject.

The phase-locked loop is widely used in many types of measurement and communication equipment. It will be discussed in more detail in Chapter 6. A simplified block diagram of a phase-locked loop for the demodulation of stereo multiplexing is shown in Fig. 5.14. Note that the use of a divide-by-n circuit following the VCO in the feedback signal path results in a VCO frequency exactly n times the input frequency. This technique is often used in frequency synthesizers.[†]

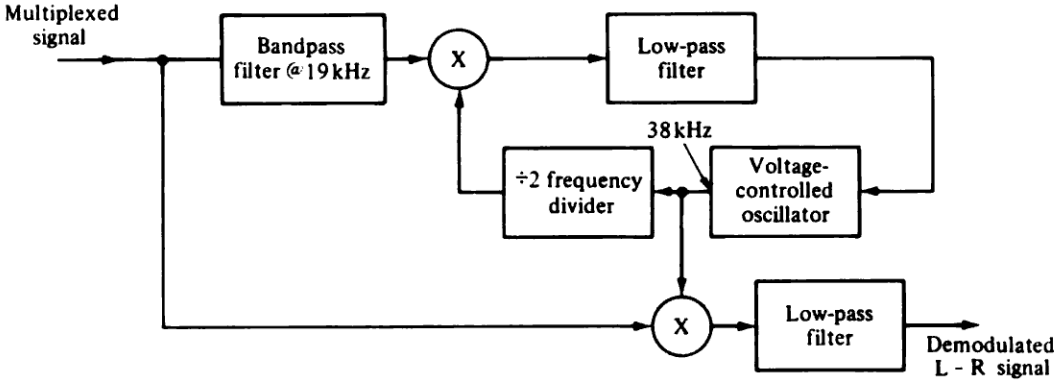


Figure 5.14 A simplified phase-locked loop stereo demodulator.

[†] A good treatment of the use of phase-locked loops for frequency synthesis can be found in U. Manassewitsch, *Frequency Synthesizers*, Third ed., Somerset, N.J.: Wiley-Interscience, 1987.

The Scanning Spectrum Analyzer

As a second example, suppose we wish to measure the magnitude of a power spectral density experimentally. Because power is proportional to area under the power spectral density, a very narrow bandpass filter is needed to slowly scan all frequencies of interest. This scanning filter would be difficult to build. Instead, it is much easier to move the spectrum and keep the filter fixed. This is normally accomplished in the following manner. A ramp voltage is generated and applied to a voltage-controlled oscillator (VCO), as shown in Fig. 5.15. This produces a sinusoid whose amplitude is constant and whose frequency is proportional to the input voltage. This sinusoid is used to multiply $f(t)$ in a balanced modulator, shifting the spectrum of $f(t)$. As the ramp progresses, the net effect is to move the signal spectrum past the (fixed) bandpass filter (rather than vice versa). The ramp voltage also serves to sweep the beam of a cathode-ray display horizontally and the magnitude of the bandpass-filter output is used to deflect the beam vertically. The resulting display is an approximation to the spectrum of $f(t)$. Accuracy of the approximation improves with decreasing filter widths and slower ramp-sweep rates. To permit adequate rise-time allowances, the rate of frequency change is usually no greater than the square of the filter bandwidth in Hz (see Problem 3.13.3).

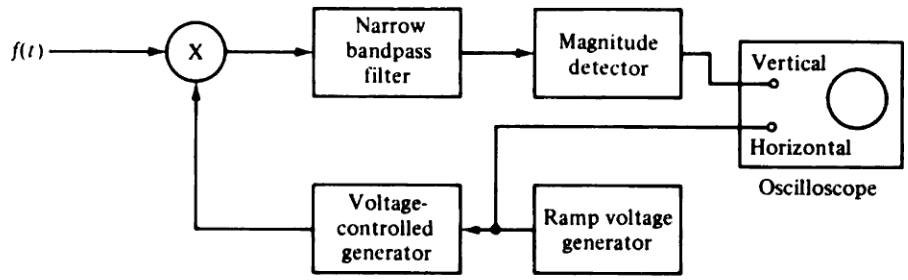


Figure 5.15 The scanning spectrum analyzer.

5.2 AMPLITUDE MODULATION: LARGE CARRIER (AM)

We have seen that the use of suppressed-carrier signal waveforms requires fairly complicated circuitry at a remotely located receiver in order to acquire and maintain the necessary phase synchronization. If we wish to use very inexpensive receivers, some alternative must be found to this synchronization problem even at the expense of a less efficient transmitter.

To this end, suppose we incorporate the carrier information as a part of the waveform being transmitted and in the same spectral width as the desired signal. In fact, it is convenient to let the amplitude of this carrier term be larger than any other part of the signal spectral density. But if we add this carrier, it will ruin the low-frequency response of the system. For some signals, e.g., audio signals, however, we really do

not need a frequency response down to zero. (We have seen that a dc, that is, $\omega = 0$, frequency response shows up in the DSB-SC signal as the spectral density at ω_c.) For systems not requiring a frequency response down to zero, then, we can place a large carrier term at ω_c. To differentiate this case from the previous case, we shall designate this as *double-sideband large-carrier* (DSB-LC). Because commercial broadcast stations use this method of transmission, it is commonly known as just *amplitude modulation* (AM).

The modulated waveform of a DSB-LC signal can be described mathematically simply by adding a carrier term, $A \cos \omega_c t$, to a DSB-SC signal:

$$\phi_{AM}(t) = f(t) \cos \omega_c t + A \cos \omega_c t. \tag{5.13}$$

The spectral density of $\phi_{AM}(t)$ is

$$\begin{aligned}\Phi_{AM}(\omega) = {} & \tfrac{1}{2}F(\omega + \omega_c) + \tfrac{1}{2}F(\omega - \omega_c) \\ & + \pi A\, \delta(\omega + \omega_c) + \pi A\, \delta(\omega - \omega_c).\end{aligned} \tag{5.14}$$

The spectral density of $\phi_{AM}(t)$ is the same as that of the DSB-SC signal $f(t) \cos \omega_c t$ with the addition of impulses at $\pm\omega_c$. This is illustrated in Fig. 5.16.

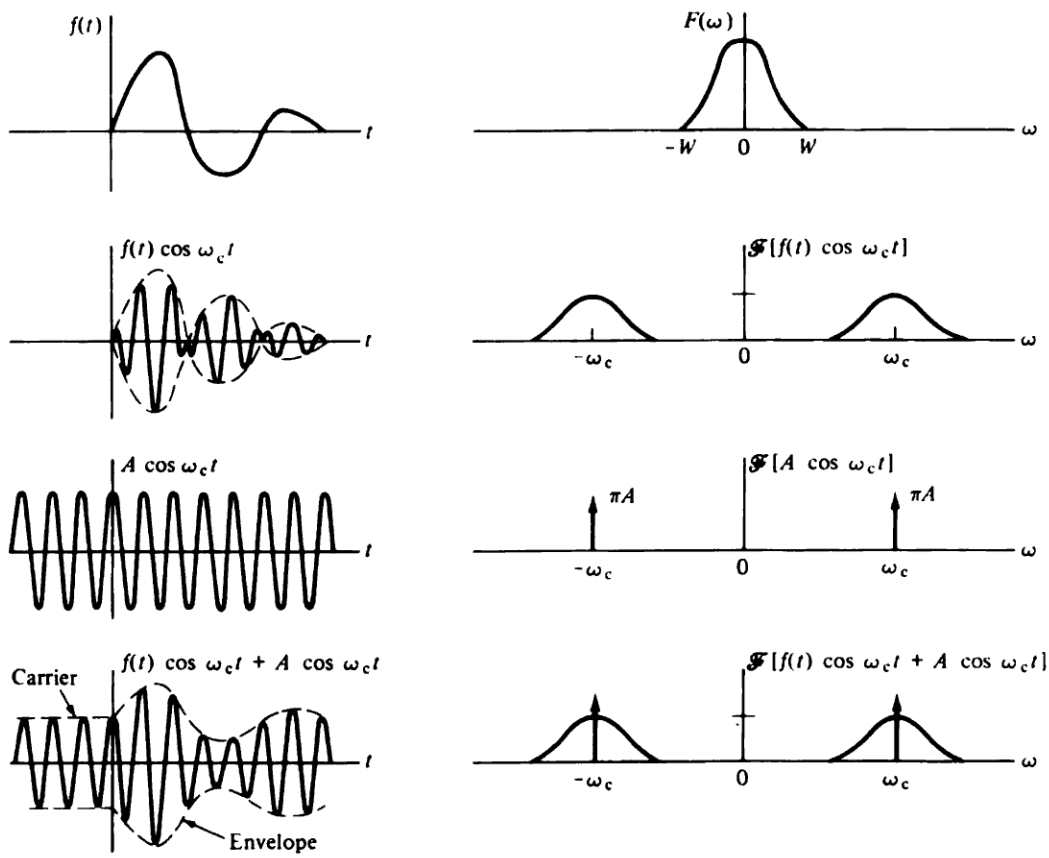


Figure 5.16 Addition of carrier to produce DSB-LC waveform.

The amplitude-modulated signal $\phi_{AM}(t)$ described by Eq. (5.13) can be rewritten in the form

$$\phi_{AM}(t) = [A + f(t)] \cos \omega_c t . \tag{5.15}$$

Written in this form, $\phi_{AM}(t)$ may be regarded as a carrier signal $\cos \omega_c t$ whose amplitude is given by the quantity $[A + f(t)]$. If A is large enough, the envelope (magnitude) of the modulated waveform will be proportional to $f(t)$. Demodulation in this case simply reduces to the detection of the envelope of a sinusoid, with no dependence on the exact phase or frequency of the sinusoid (i.e., of the carrier). If A is not large enough, however, the envelope of $\phi_{AM}(t)$ is not always proportional to $f(t)$, as shown in Fig. 5.17.

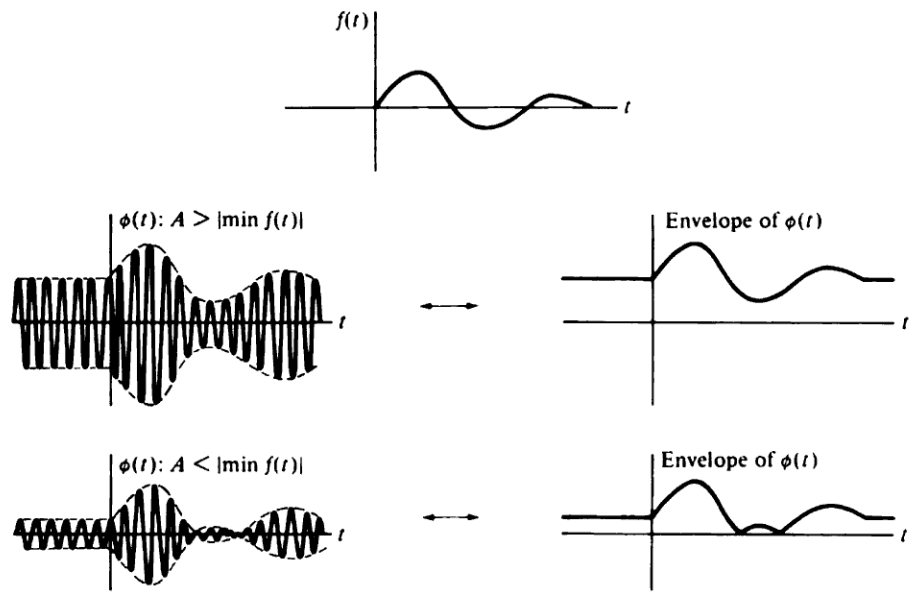


Figure 5.17 Importance of sufficient carrier in DSB-LC waveform.

The carrier amplitude A must be made large enough so that $[A + f(t)] \geq 0$ at all times, or

$$A \geq |\min \{f(t)\}| . \tag{5.16}$$

If Eq. (5.16) is not satisfied, then $f(t)$ cannot be recovered by the relatively simple process of envelope detection.† Synchronous detection, however, will still correctly demodulate such signals. Note that the dc response of the signal $f(t)$ has been lost in the demodulation as a result of the addition of the carrier.

To emphasize these concepts, we now return to the special case of a single-frequency sinusoidal tone, $\cos \omega_m t$, as the modulating signal. Because the relative

† Methods of envelope detection will be discussed in Section 5.2.3.

magnitudes of the sideband and carrier portion of the signal are variable, we define a dimensionless scale factor, m, to control the ratio of the sidebands to the carrier:

$$m = \frac{\text{peak DSB-SC amplitude}}{\text{peak carrier amplitude}}, \tag{5.17}$$

so that†

$$\begin{aligned}\phi_{AM}(t) &= A \cos \omega_c t + mA \cos \omega_m t \cos \omega_c t \\ &= A(1 + m \cos \omega_m t) \cos \omega_c t. \end{aligned} \tag{5.18}$$

A graph of each of the terms appearing in Eq. (5.18) and their addition for different values of the parameter m is shown in Fig. 5.18. Note that those portions of the DSB-SC waveform that are in phase [that is, $f(t) > 0$] with the carrier add while those portions that are 180° out of phase [that is, $f(t) < 0$] with the carrier subtract. The maxima in the composite waveform envelope are $(1 + m)A$ and the minima for $m \leq 1$ are $(1 - m)A$. As long as $m \leq 1$, the peak values of the modulated carrier trace out an envelope waveform which is proportional to the modulating tone, $\cos \omega_m t$. When $m = 0$, of course, the envelope reduces to the constant A, as it should.

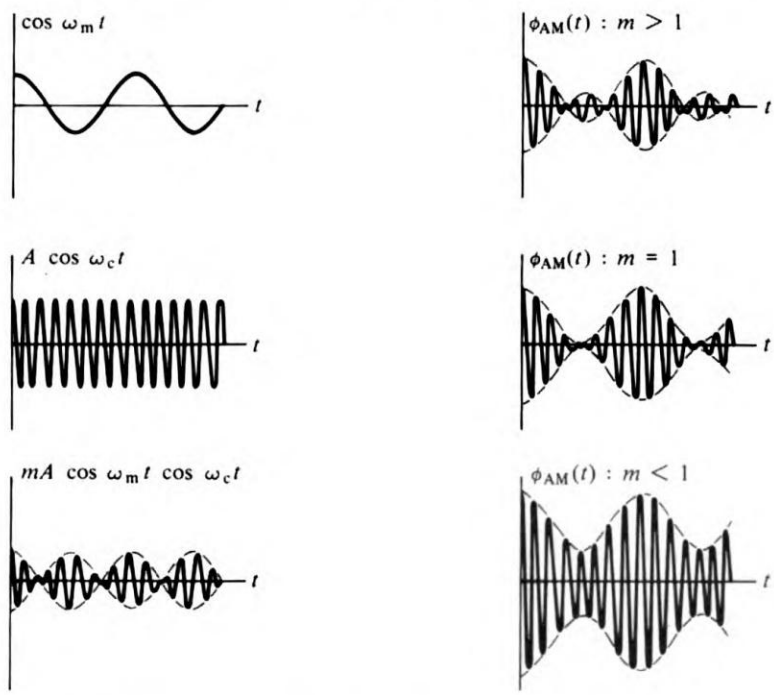


Figure 5.18 Effects of varying modulation indexes.

† Including a multiplier constant k_a, if $f(t) = a \cos \omega_m t$, then $m = ak_a$.

Often it is convenient to define a percentage of modulation for a DSB-LC signal with sinusoidal modulation as

$$\% \text{ mod.} = \frac{(\text{max. magnitude}) - (\text{min. magnitude})}{(\text{max. magnitude}) + (\text{min. magnitude})} \times 100\%$$

$$= \frac{(1 + m)A - (1 - m)A}{(1 + m)A + (1 - m)A} \times 100\% = m \times 100\% . \qquad (5.19)$$

The parameter m controlling the relative proportions of sideband to carrier is called the *modulation index* of the AM waveform. From Fig. 5.18 one can readily see that $m \leq 1$ is required for envelope detection to take place without severe distortion. If $m > 1$, the waveform is said to be overmodulated.

The student should beware not to confuse the waveform of the summation of two sinusoids with that of a sinusoidal amplitude-modulated waveform with a small modulation index. Both cases are depicted in Fig. 5.19. Note that their spectral densities are clearly very different.

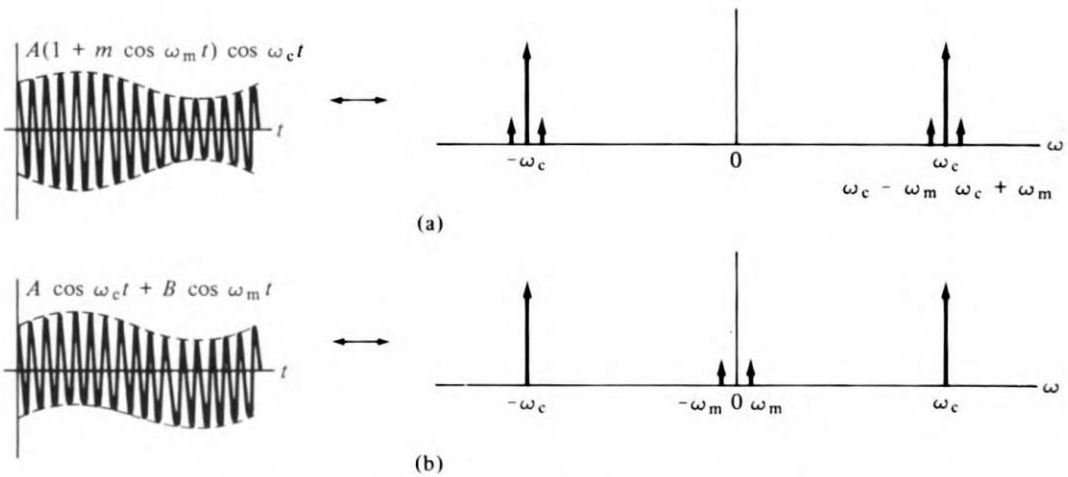


Figure 5.19 (a) Amplitude modulation and (b) addition for the sinusoidal case.

5.2.1 Carrier and Sideband Power in AM

In AM signal waveforms, the carrier term does not contain any information about the modulating signal $f(t)$. Therefore the power expended in this carrier is wasted for any transfer of information. It is the price one must be willing to pay in order to make cheap receivers available.

A general AM signal waveform can be described by

$$\phi_{AM}(t) = A \cos \omega_c t + f(t) \cos \omega_c t . \qquad (5.20)$$

For a 1-ohm load, the total average power is given by the mean-square value of Eq. (5.20),

$$\overline{\phi_{\mathrm{AM}}^2(t)} = A^2\,\overline{\cos^2\omega_c t} + \overline{f^2(t)\cos^2\omega_c t} + 2A\overline{f(t)\cos^2\omega_c t}\,, \tag{5.21}$$

where the bar indicates time averaging. We shall assume that $f(t)$ varies slowly with respect to $\cos\omega_c t$. If we also assume that the average value of $f(t)$ is zero (the usual case), then the last term in Eq. (5.21) is zero so that†

$$\begin{aligned}\overline{\phi_{\mathrm{AM}}^2(t)} &= A^2\,\overline{\cos^2\omega_c t} + \overline{f^2(t)}\;\overline{\cos^2\omega_c t}\\ &= A^2/2 + \overline{f^2(t)}/2\,.\end{aligned} \tag{5.22}$$

Thus the total power P_t can be expressed as the sum of a carrier power P_c and a sideband power P_s:

$$P_t = \tfrac{1}{2}A^2 + \tfrac{1}{2}\overline{f^2(t)} = P_c + P_s\,. \tag{5.23}$$

That fraction of the total power contained in the sidebands, μ, is given by

$$\mu = \frac{P_s}{P_t} = \frac{\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}\,. \tag{5.24}$$

Returning to the case in which $f(t)$ is a single sinusoid [cf. Eq. (5.18)],

$$\begin{aligned}\phi_{\mathrm{AM}}(t) &= A(1 + m\cos\omega_m t)\cos\omega_c t\\ &= A\cos\omega_c t + mA\cos\omega_m t\cos\omega_c t\,,\end{aligned}$$

we have

$$\overline{\phi_{\mathrm{AM}}^2(t)} = \tfrac{1}{2}A^2 + (\tfrac{1}{2})(\tfrac{1}{2})m^2A^2, \tag{5.25}$$

and

$$\mu = \frac{m^2}{2 + m^2}\,. \tag{5.26}$$

Because $m \le 1$, we see from Eq. (5.26) that the transmission efficiency of an AM (DSB-LC) system is at best 33%. Under the best condition, i.e., $m = 1$, 67% of the total power is expended in the carrier and represents wasted power as far as the transfer of information is concerned. For lower modulation indexes, the efficiency is less than 33%. For example, a modulation index of 0.50 yields an efficiency of 11.1%. In contrast, the transmission efficiency of a DSB-SC system is 100%. (If a pilot carrier is transmitted, the efficiency is slightly less than 100%.)

† If $\overline{f(t)}$ is not zero, then the first and last terms on the right-hand side of Eq. (5.21) can be combined in an effective carrier power. In other words, a nonzero average value of $f(t)$ shows up in the AM waveform as a carrier term.

EXAMPLE 5.2.1

A given AM (DSB-LC) broadcast station transmits an average carrier power output of 40 kW and uses a modulation index of 0.707 for sine-wave modulation. Calculate (a) the total average power output; (b) the transmission efficiency; and (c) the peak amplitude of the output if the antenna is represented by a 50-Ω resistive load.

Solution a) Using Eq. (5.25), we get

$$P_t = P_c(1 + m^2/2).$$

For $m = 0.707$,

$$P_t = (1 + \tfrac{1}{4})P_c = \tfrac{5}{4}P_c,$$

$$P_t = 50 \text{ kW}.$$

b) Using Eq. (5.26) gives us

$$\mu = \frac{(0.707)^2}{2 + (0.707)^2} = \frac{0.500}{2.500} = 20\%.$$

c) Then

$$P_c = \frac{A^2}{2R},$$

$$A^2 = 2RP_c = 4 \times 10^6,$$

$$(1 + m)A = 3414 \text{ V}.$$

DRILL PROBLEM 5.2.1

The amplitude of the modulating tone is decreased in the above example until the total power is 45 kW. Assuming that the carrier power remained constant, compute the new modulation index and transmission efficiency.

ANSWER: (a) 0.500; (b) 11.1%.

5.2.2 Generation of DSB-LC Signals

Conceptually, the easiest way to generate a DSB-LC signal is to first generate a DSB-SC signal and then add some carrier. This is shown in block diagram form in Fig. 5.20. It turns out, however, that DSB-LC signals are generally easier to generate directly so that the system shown in Fig. 5.20 has more analytical than practical value.

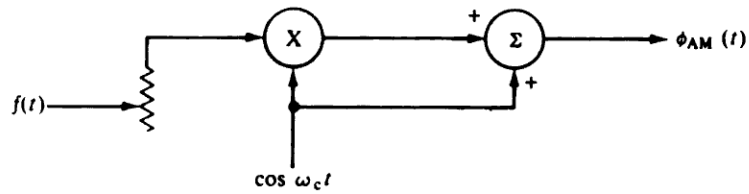


Figure 5.20 Block diagram of DSB-LC generation.

As in the case of DSB-SC signals, the generation of DSB-LC signals can be divided into the chopper (switch) type modulators and those modulators using nonlinear characteristics of devices.

The Chopper (Rectifier) Modulator

A straighforward extension of the chopper modulator for generating AM is to add a dc level to $f(t)$ before chopping. If this dc level A is large enough so that the condition $[A + f(t)] > 0$ holds, then it is easy to see that the output signal will be a DSB-LC signal.

Another possibility, which really amounts to the same thing, is to add some carrier to $f(t)$ before chopping. This is shown in Fig. 5.21(a). The chopper action may be viewed as a multiplication of the input signal with a periodic square wave $p_T(t)$ whose fundamental frequency is ω_c rad/sec. The spectral density of the resulting chopped signal can be found by using the modulation property of the Fourier transform or by taking the frequency convolution of the spectral density of $[f(t) + k \cos \omega_c t]$ with the spectral density of $p_T(t)$, as shown in Fig. 5.21. The result yields a term: $[\frac{1}{2}k \cos \omega_c t + (2/\pi)f(t) \cos \omega_c t]$ centered about $\pm\omega_c$. Unwanted frequency components at $\omega = 0$, $\pm 3\omega_c$, $\pm 5\omega_c$, etc., can be filtered out by proper choice of a bandpass filter.

It is not necessary that the chopper be driven by a separate source if the carrier amplitude is made much greater than $f(t)$. An ideal diode will act as a good switch and, if the carrier is large, will turn on and off at the carrier rate. A diagram of a modulator using a diode as a switch is shown in Fig. 5.21(b). The spectral characteristics are the same as for the chopper modulator described above. Because the diode eliminates the negative part of the composite signal $[f(t) + k \cos \omega_c t]$, it essentially performs a half-wave rectification of the input signal. For this reason, this type of modulator is also known as a rectifier-type modulator.

Amplitude modulation with the rectifier-type modulator is relatively easy to accomplish at high-power levels by using amplifiers operated in class C conditions. A parallel LC filter on the output performs the required bandpass filtering operation. If overmodulation is attempted with this type of modulator, no output will result instead of reversals of the carrier phase. This results in abrupt discontinuities in the modulated waveform and introduces much harmonic content in the spectrum. To prevent this, high-power AM broadcast transmitters operated in this mode are required to maintain a modulation index less than 100% on negative peaks in programming material.

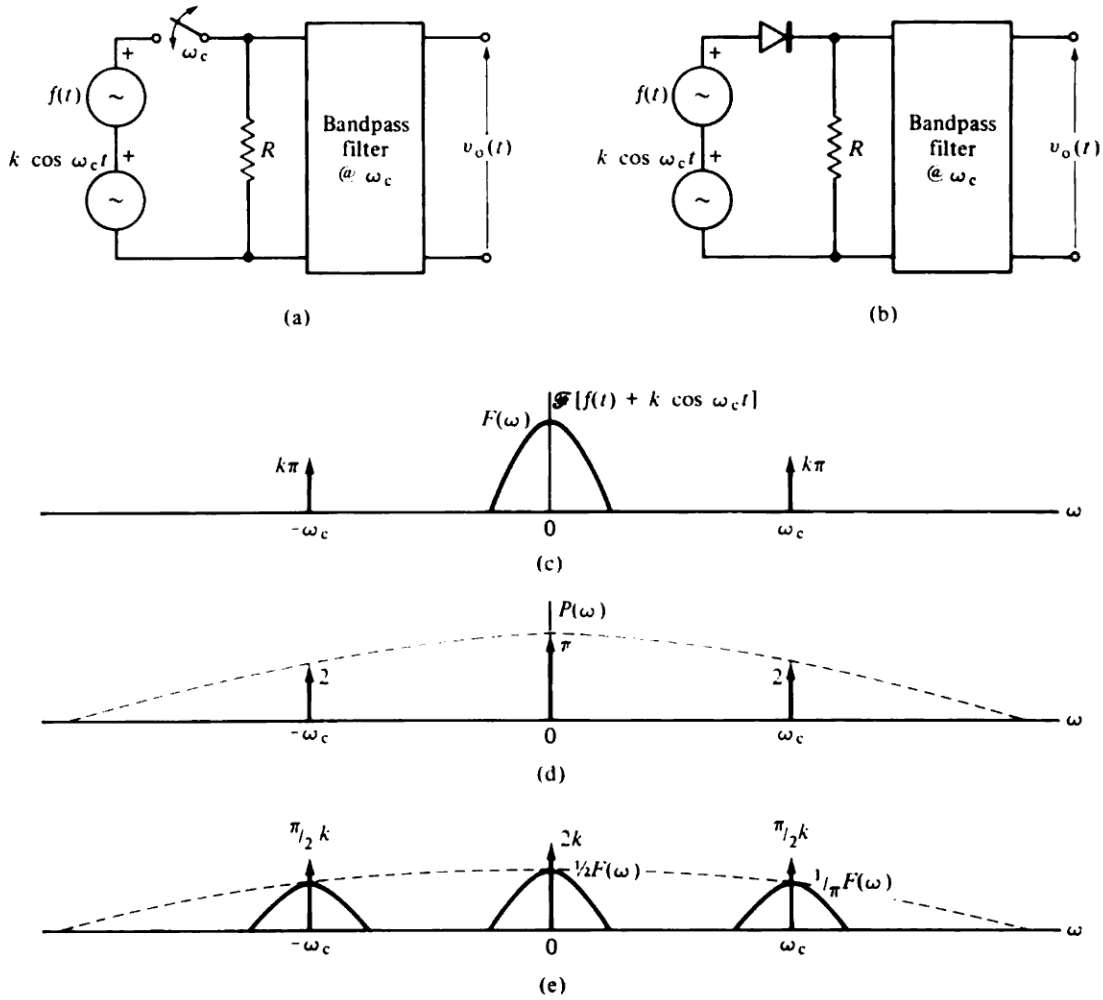


Figure 5.21 Generation of an AM signal using the chopper (rectifier) modulator.

Modulator Using Nonlinearities

The system shown in Fig. 5.21(b) can be used to generate AM even if the diode is not operated as an ideal switch. In this case, the nonlinearities in the diode characteristic may be approximated with a power series of the form

$$i(t) = a_1 e(t) + a_2 e^2(t) + \cdots, \qquad i(t)R \ll e(t). \tag{5.27}$$

Retaining only the first two terms, we find that the voltage at the input to the bandpass filter (neglecting effects of any finite input impedance) is

$$i(t)R = a_1 R[f(t) + k\cos\omega_c t] + a_2 R[f(t) + k\cos\omega_c t]^2 + \cdots. \tag{5.28}$$

Expanding and collecting terms at the carrier frequency, we have

$$v_o(t) = a_1 Rk\cos\omega_c t + 2a_2 Rkf(t)\cos\omega_c t. \tag{5.29}$$

Equation (5.29) is the desired result for a DSB-LC signal. A semiconductor diode operates as a combination of the rectifier modulator and the modulator using a nonlinear characteristic.

DRILL PROBLEM 5.2.2

Let $f(t)$ be equal to $\cos \omega_m t$. If $a_1 = 0.010$ and $a_2 = 0.001$ in Eq. (5.28), determine the modulation index of the resulting DSB-LC signal when all terms except those near the carrier frequency are filtered out.

ANSWER: $m = 0.20$.

5.2.3 Demodulation (Detection) of DSB-LC (AM) Signals

In DSB-LC (AM) signals, the desired signal waveform $f(t)$ is available in the envelope of the modulated signal. The use of synchronous detection will, of course, yield the desired waveform, but it is possible to demodulate AM signals by much simpler techniques. The simplest and most popular method is one that detects the envelope of the modulated waveform directly and is called the envelope detector.

The Envelope Detector

Any circuit whose output follows the envelope of the input signal waveform will serve as an envelope detector. The simplest form of an envelope detector is a nonlinear charging circuit with a fast charge time and a slow discharge time. It can easily be constructed using a diode in series with a capacitor, as shown in Fig. 5.22(b). A

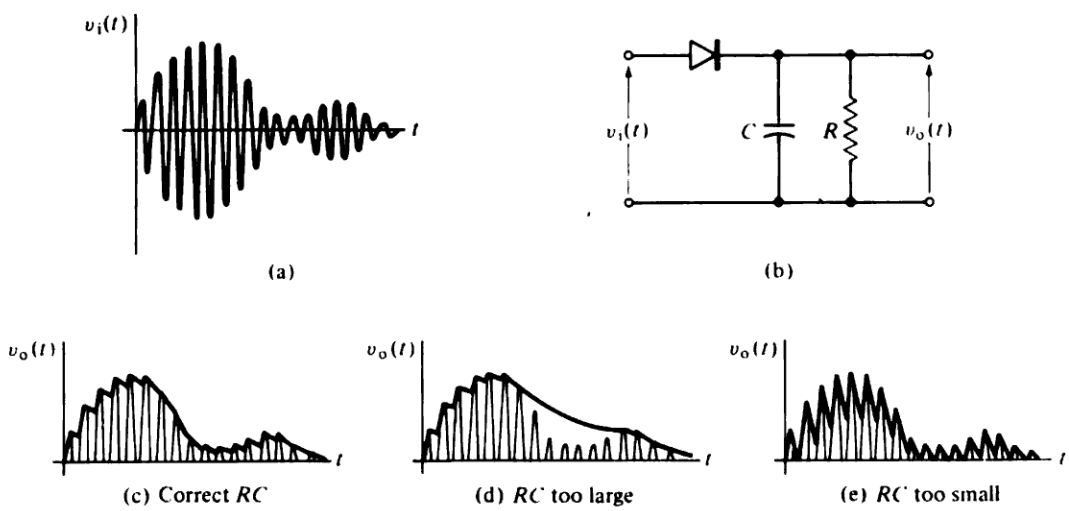


Figure 5.22 The envelope detector.

resistor placed across the capacitor controls the discharge time constant. Effects of various discharging time constants are shown in Fig. 5.22(c), (d), (e).

The operation of the envelope detector is as follows. The input signal waveform charges capacitor C to the waveform's maximum value during positive half-cycles of the input signal. As the input signal falls below its maximum, the diode turns off. This is followed by a slow discharge of the capacitor through the resistor until the next positive half-cycle when the input signal becomes greater than the capacitor voltage and the diode turns on again. The capacitor charges to the new peak value, etc.

For best operation, the discharge time constant, RC, should be adjusted so that the maximum negative rate of the envelope will never exceed the exponential discharge rate. If this time constant is too large, the envelope detector may miss some positive half-cycles of the carrier and thus will not reproduce the envelope faithfully. If the time constant is much too small, the envelope detector generates a very ragged waveform, losing some of its efficiency. The resultant detected signal is usually passed through a low-pass filter to eliminate the unwanted harmonic content. A coupling capacitor can be used to remove the dc level introduced by the carrier. As a result of the presence of this dc level, DSB-LC methods are not suitable if a frequency response for $f(t)$ down to $\omega = 0$ is desired.

The envelope detector is simple, efficient, and cheap to build. It is almost universally used for the purpose of detecting DSB-LC signals.

The ease and simplicity in using the envelope detector suggests that perhaps these techniques could also be applied to the detection of DSB-SC signals. Because there is no carrier term in such signals, enough carrier must be added in the receiver to make the envelope detection possible. This is sometimes done to simplify receiver design; this type of receiver is called an *injected-carrier* type receiver. A major problem, of course, is to keep the injected carrier at the correct frequency and thus such systems also have synchronization problems.

Other Methods of Detection

If a ground return is provided in the envelope detector circuit described above, the diode can be made to operate essentially like a switch, turning on and off at the carrier rate. Because the modulated waveform has been rectified by this action, this detector is called a rectifier detector. A low-pass filter removes the high-frequency terms after rectification, as shown in Fig. 5.23. The rectifier detector is not as efficient as the envelope detector and is rarely used in commercial receivers. It finds application in the amplitude detection of pulsed signals where the slow exponential discharge time of the envelope detector is undesirable.

The synchronous detection methods used for DSB-SC signal waveforms can, of course, also be used for DSB-LC signals. For example, a phase-locked loop can be used to synchronize an oscillator to the carrier frequency and this can be used to synchronously demodulate the signal. Such elaborate receiving techniques are rarely used, however, as a result of the required additional complexity in receiver design and operation.

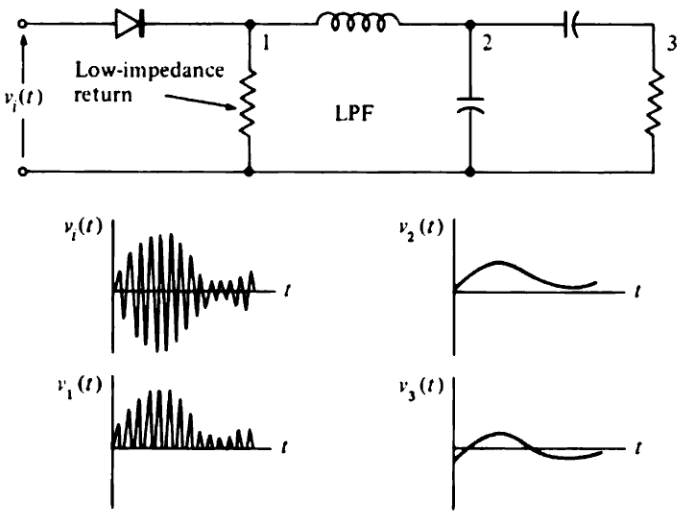


Figure 5.23 The rectifier detector.

5.3 FREQUENCY-DIVISION MULTIPLEXING (FDM)

From our discussion of amplitude modulation thus far, we see that it is possible to send several signals simultaneously by choosing a different carrier frequency for each signal. These carrier frequencies are chosen so that the signal spectra are not overlapping. This mode of transmission is called *frequency-division multiplexing* (FDM). Frequency-division multiplexing, then, is the positioning of signal spectra in frequency such that each signal spectrum can be separated out from all the others by filtering. We shall emphasize the use of amplitude modulation here, but frequency-division multiplexing does not preclude the use of other modulation methods.

To illustrate the principles of frequency-division multiplexing (FDM), suppose we wish to transmit several signals simultaneously, using DSB-LC or DSB-SC modulation. Let us assume that there are N different signals and, for convenience, that each signal is bandlimited to ω_m rad/sec. In order to separate these N signals in frequency, each is modulated with a carrier frequency $\omega_1, \omega_2, \ldots, \omega_N$. Thus the spectral density of every modulated signal has a bandwidth of $2\omega_m$ and each is centered at the various carrier frequencies $\omega_1, \omega_2, \ldots, \omega_N$. These carrier frequencies are chosen far enough apart such that each signal spectral density is separate from all the others. This requires that each carrier frequency be separated from an adjacent carrier frequency by at least $2\omega_m$. Figure 5.24(a), (b), (c) illustrates the case in which three signals are translated in frequency and transmitted simultaneously.

At the receiver, we consider two possibilities. In the first, the receiver processes the various signal spectra simultaneously, separating them in frequency with the appropriate bandpass filters and then demodulating. This is shown in Fig. 5.24(d). In practice, the composite signal formed by spacing several signals in frequency may, in turn, be modulated using another carrier frequency. To distinguish the choice of the first carrier frequencies in this case, the $\omega_1, \omega_2, \ldots, \omega_N$ are called *subcarriers*.

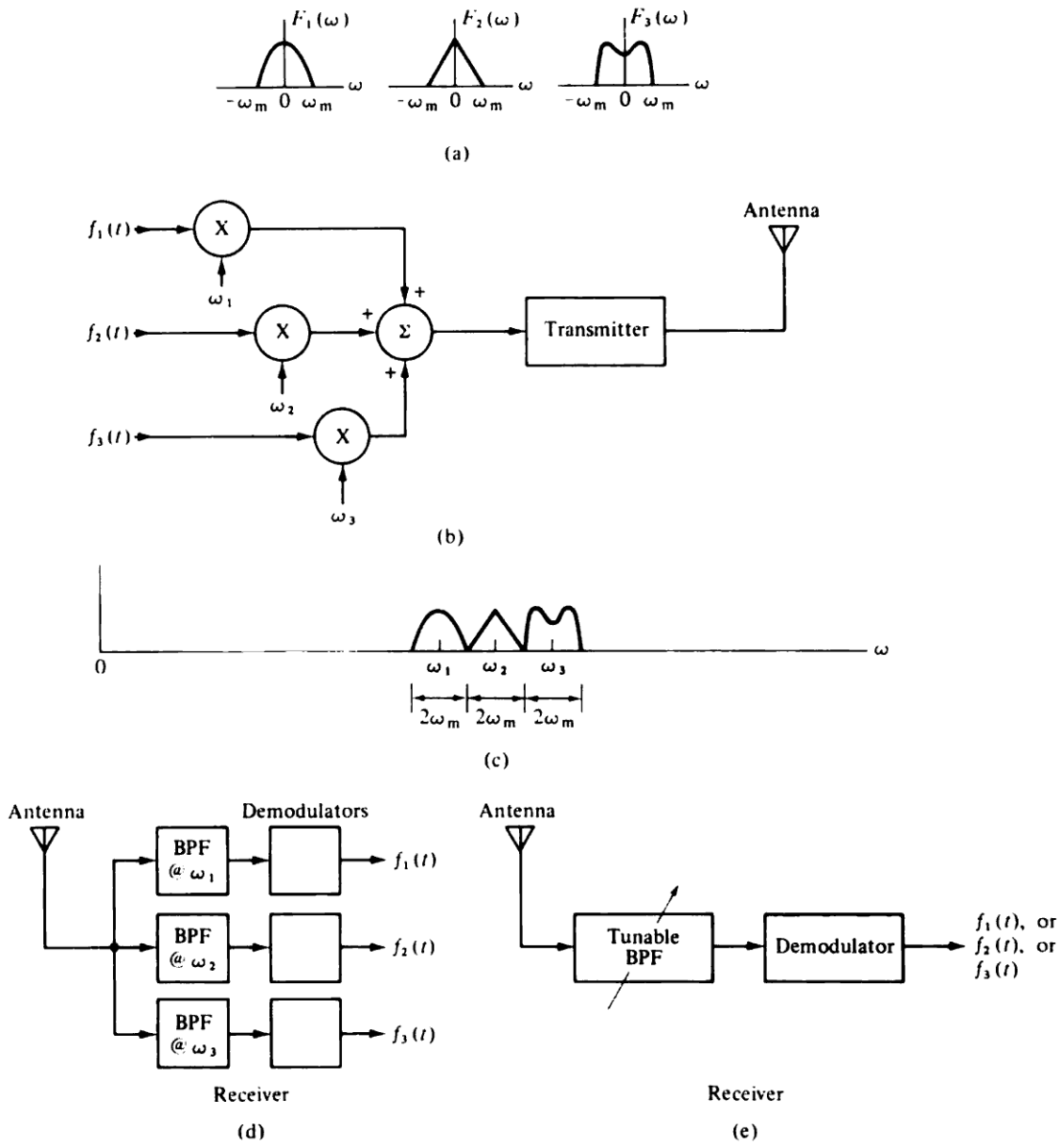


Figure 5.24 Frequency-division multiplexing (FDM).

The second possibility is to have each receiver select only one of all possible signals. This is accomplished by tuning a bandpass filter to the center frequency of the desired signal and then demodulating. A receiver diagram for this possibility is shown in Fig. 5.24(e).

Each of these possibilities has applications and the choice largely determines whether large-carrier or suppressed-carrier amplitude modulation is employed. In the first, several signals are processed simultaneously so that at most only one large carrier, or several pilot carriers, need be sent to the receiver. This allows the use of the

more efficient DSB-SC modulation. The stereo multiplexing discussed earlier is a familiar example of this. On the other hand, in the second method there is only one signal being received at a time and, to keep the receiver simple, DSB-LC is usually used for this type of operation. Commercial radio and television receivers provide familiar examples of this type of FDM. It is worthwhile considering the standard radio receiver to see how this is accomplished in practice.

Each commercial AM broadcast transmitter sends out a DSB-LC signal whose carrier frequency is separated from the carrier frequencies of other stations. Carrier frequencies are assigned at 10-kHz spacings from 540 kHz to 1600 kHz.[†] A receiver can pick up any one of these signals by tuning to the appropriate carrier frequency. The selected signal is then demodulated using an envelope or rectifier detector to produce the desired signal.

Early AM receivers performed these operations in exactly this manner. An antenna, an *LC* tuning arrangement for station selection, a diode for a detector, and a pair of headphones constituted the first receivers. Unless one lived near a broadcast station, the received signals at the antenna terminals were small and amplification and filtering were needed to improve the sensitivity (i.e., the ability to receive weaker signals) and the selectivity (i.e., the ability to separate the signals of different stations).

The next logical step in receiver improvement was the tuned-radio-frequency (TRF) receiver, shown in Fig. 5.25(a). All stages of amplification (usually about three) were tuned simultaneously to select stations. Varying three separate tuning controls, however, made changing stations a frustrating chore. Ganging all three together, on the other hand, created additional problems because all stages did not always change at the same rate (i.e., track) unless the bandwidth of each stage was broadened, ruining the sensitivity and selectivity. What was needed was the amplification of a narrow band of frequencies that would remain fixed regardless of which station was selected. This gave rise to the heterodyne receiver that is widely used today.

Heterodyning means the translating or shifting in frequency. In the heterodyne receiver the incoming modulated signal is translated in frequency, thus occupying an equal bandwidth centered about a new frequency. This new center frequency, known as an intermediate frequency (IF), is fixed and is not dependent on the received signal center frequency. The signal is amplified at the IF before demodulation. If this intermediate frequency is lower than the received carrier frequency but above the final output signal frequency, it is called a superheterodyne receiver. In most commercial AM broadcast receivers of the superheterodyne type, the intermediate frequency is 455 kHz. AM broadcast frequency assignments are from 540 kHz to 1600 kHz. The required frequency translation to the intermediate frequency is accomplished by mixing the incoming signal with a locally generated signal that differs from the incoming carrier by the intermediate frequency (in this case 455 kHz). The received signal, now translated to a fixed intermediate frequency, can easily be amplified, filtered,

[†] Frequency allocations in the United States are described in Appendix C.

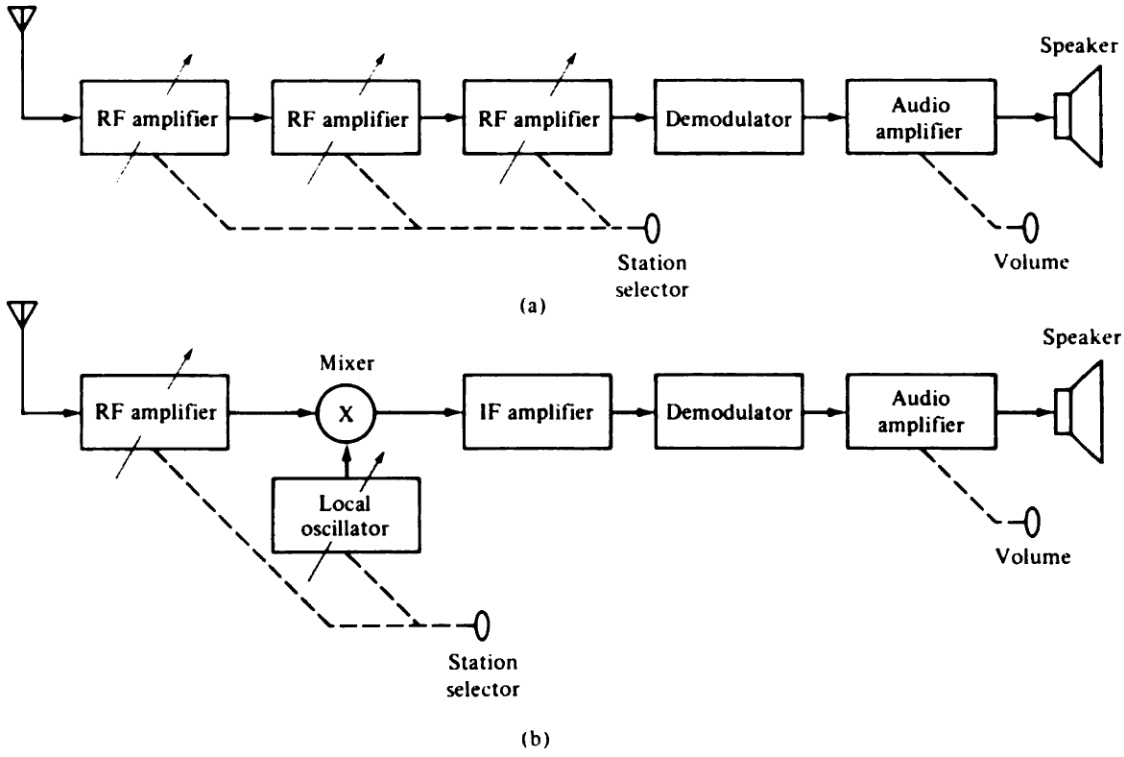


Figure 5.25 Radio receivers: (a) the TRF and (b) the superheterodyne.

and demodulated.† The primary advantage of the superheterodyne receiver over the TRF receiver is that the amplification and filtering is performed at a fixed frequency regardless of station selection. A block diagram of the superheterodyne receiver is shown in Fig. 5.25(b).

To translate the spectrum of the incoming signal to the fixed intermediate frequency, the locally generated carrier must be at a frequency either above or below the incoming carrier frequency by the intermediate frequency (455 kHz). In the AM superheterodyne receiver, the locally generated frequency is chosen to be 455 kHz higher than the incoming signal. The reason for this choice is simply that it is easier to build oscillators which are reasonably linear in tuning over a 1–2 MHz range than over a 0.1–1.1 MHz range.

The advantages of the superheterodyne receiver are not without disadvantages. Suppose, for example, that reception of an AM station at 600 kHz is desired. This means that the locally generated carrier, assuming it is 455 kHz above the station carrier, is at 1055 kHz. Now if there is another station at 1510 kHz, it also will be received (note that 1510 kHz − 1055 kHz = 455 kHz). This second frequency, 1510 kHz, is called the *image frequency* of the first, and after the heterodyning operation it is impossible to distinguish the two.

† For reception of higher frequencies, the translation to an intermediate frequency may be repeated a second time. Receivers using this method are called dual-conversion receivers.

It is instructive to use the modulation property (or the frequency-convolution property) of the Fourier transform to see how the image-frequency problem arises. The spectral densities are shown in Fig. 5.26 for cases in which the station center frequency is either above or below the local oscillator frequency by the amount of the intermediate frequency. If *both* station transmissions are present, both will be received equally well. One is called the *image* of the other. Note that the image frequency is displaced by twice the intermediate frequency from the desired station frequency.

There are two ways to minimize the image-frequency problem. Noting that the image frequency is separated from the desired signal by exactly twice the intermediate frequency, we may choose the intermediate frequency as high as possible and practical. The other way is to attenuate the image frequency before heterodyning. This is the primary function of a selective radio-frequency (RF) amplifier placed before the mixer. It must, of course, be tunable over the frequency range of the incoming signals.

Another disadvantage of the superheterodyne receiver is that the high-gain IF stages are tuned outside the assigned frequency band. Thus one must be careful that the intermediate frequency chosen is free from other strong transmissions or

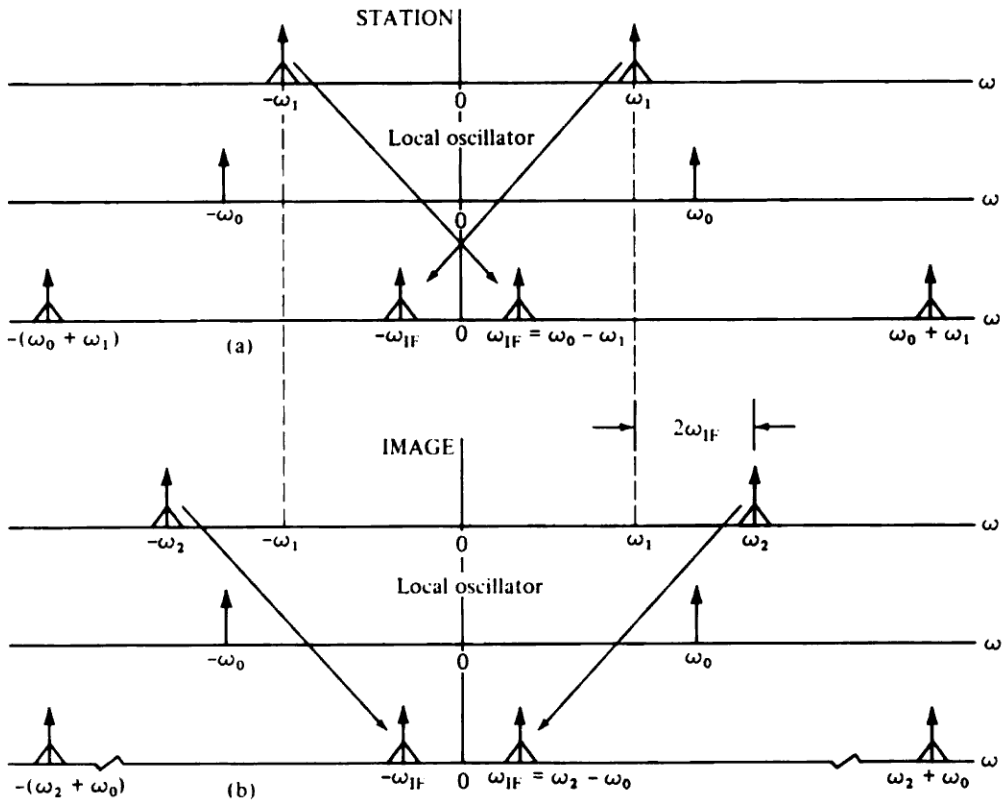


Figure 5.26 Frequency heterodyne with local oscillator above station carrier showing (a) translation of the received station signal to the intermediate frequency and (b) translation of the image frequency to the intermediate frequency.

otherwise the receiver will amplify these spurious signals as they leak into the high-gain IF stages. A combination of these considerations makes the choice of 455 kHz a popular one for standard AM broadcast receivers. Although the details of modulation differ, commercial FM and television receivers also use the superheterodyne principle with intermediate frequencies usually at 10.7 MHz and about 44 MHz, respectively.

Because the heterodyne receiver generates an internal carrier, it is also a miniature transmitter of energy at the local oscillator frequency. This can be demonstrated by placing two AM superheterodyne receivers in close proximity and tuning one 455 kHz below the second. When close to this value, some interference may be observed as the second receiver picks up a local oscillator signal from the first. This radiation of the heterodyne receiver has been used to advantage in electronic detection and countermeasures work. On the other hand, it can cause undesirable interference effects. These effects can be minimized by proper shielding in the offending receiver.

EXAMPLE 5.3.1

A given radar receiver operating at a frequency of 2.80 GHz and using the superheterodyne principle has a local oscillator frequency of 2.86 GHz. A second radar receiver operates at the image frequency of the first and interference results.

a) Determine the intermediate frequency of the first radar receiver.

b) What is the carrier frequency of the second receiver?

c) If you were to redesign the radar receiver, what is the minimum intermediate frequency you would choose to prevent image-frequency problems in the 2.80–3.00 GHz radar band?

Solution a) $f_{IF} = f_{LO} - f_c = 2.86 \text{ GHz} - 2.80 \text{ GHz} = 60 \text{ MHz}$

b) $f_{IMAGE} = f_c + 2f_{IF} = 2.80 \text{ GHz} + 0.12 \text{ GHz} = 2.92 \text{ GHz}$

c) $2f_{IF} \geq f_{MAX} - f_{MIN} = 3.00 \text{ GHz} - 2.80 \text{ GHz} = 0.20 \text{ GHz}; f_{IF} \geq 100 \text{ MHz}$

DRILL PROBLEM 5.3.1

A telemetry receiver is designed to receive satellite transmissions at 136 MHz. The receiver uses two heterodyne operations with intermediate frequencies of 30 MHz and 10 MHz. (This type of receiver, commonly known as a dual-conversion receiver, is shown in block diagram form in Fig. 5.27.) The first local oscillator is designed to operate below the incoming carrier frequency; the second, above the first (30 MHz) intermediate frequency. Determine all possible image frequencies. Do not assume filters are ideal. (In other words, what possible input frequencies could result in images for both the first and second frequency mixers if the filters were not ideal?)

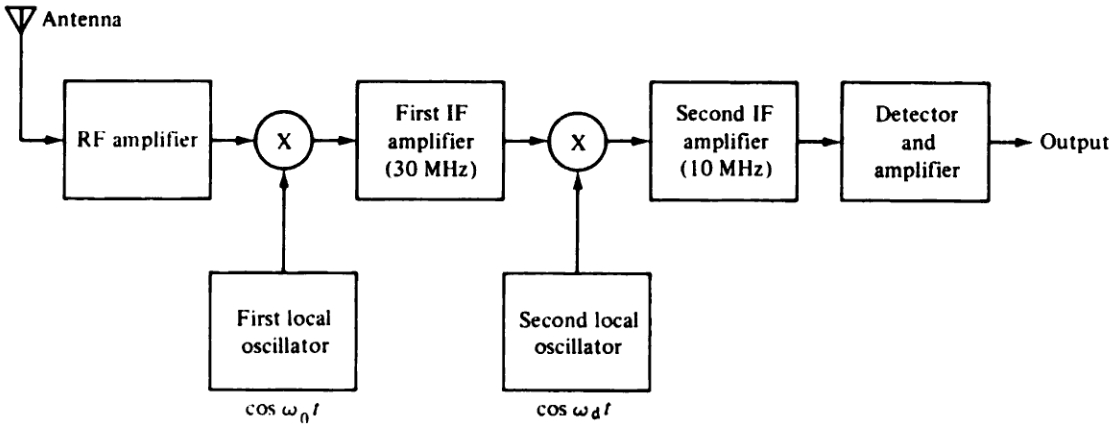


Figure 5.27 Block diagram of a dual-conversion receiver (cf. Drill Problem 5.3.1).

ANSWER: 56 MHz, 76 MHz, 156 MHz.

5.4 SINGLE-SIDEBAND (SSB) MODULATION

We have seen in previous sections that DSB modulation results in a doubling of the bandwidth of a given signal. However, this doubling of the bandwidth is a disadvantage when a given band of frequencies is crowded. Thus we wish to investigate whether this doubling of bandwidth is really necessary.

It will be helpful to first recall the spectral densities of DSB signals (we assume suppressed-carrier for convenience). The spectral density of any real-valued signal must exhibit the symmetry condition (cf. Fig. 5.28a)

$$F(-\omega) = F^*(\omega). \tag{5.30}$$

After multiplication by a sinusoid at ω_c rad/sec, half of this spectral density is translated up in frequency and centered about ω_c and half is translated down to $(-\omega_c)$, as shown in Fig. 5.28(b).

From Fig. 5.28(a), (b), it can be seen that what was originally positive frequency content in $F(\omega)$ now becomes the upper sideband for $\omega > 0$ and the lower sideband for $\omega < 0$. Conversely, that which was originally negative frequency content now becomes the lower sideband for $\omega > 0$ and the upper sideband for $\omega < 0$. Equation (5.30) has been satisfied throughout this process. Each pair of sidebands (i.e., upper or lower) contains the complete information of the original signal.

Now we entertain an intriguing thought: Why not transmit only the upper or the lower pair of sidebands since each pair contains all the information about the signal? We note that Eq. (5.30) will still be satisfied and that the original signal can be

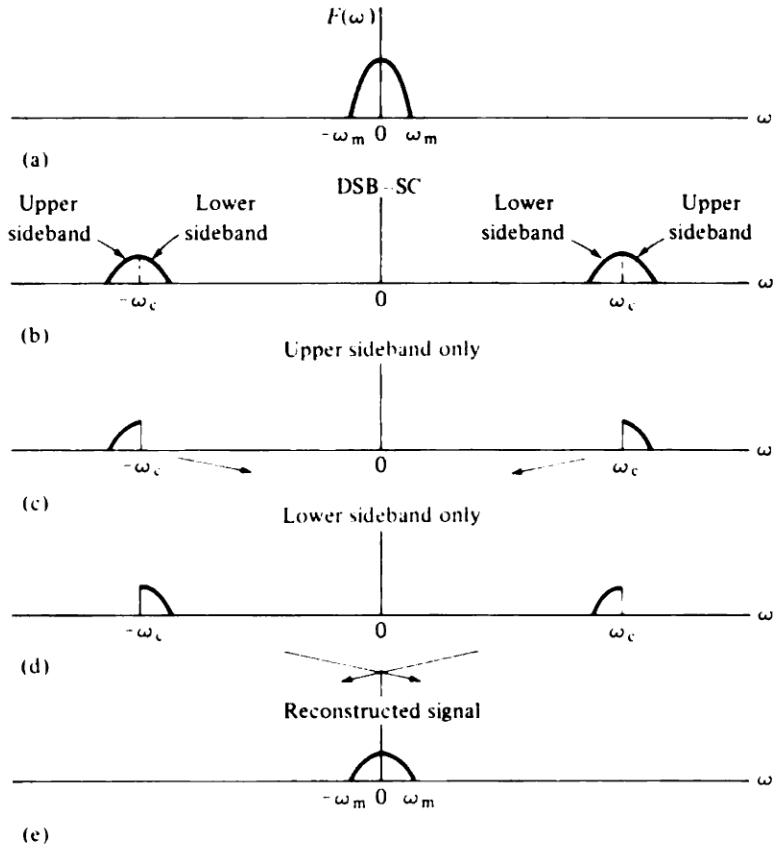


Figure 5.28 Spectra of DSB and SSB signals.

recovered again from either the upper or lower pair of sidebands by an appropriate frequency translation, as suggested in Fig. 5.28(e). This type of modulation is called *single-sideband* (SSB) modulation. Single-sideband modulation is efficient because it requires no more bandwidth than that of the original signal and only half that of the corresponding DSB signal. We shall investigate SSB modulation and demodulation methods in more detail.

5.4.1 Generation of SSB Signals

The most straightforward way conceptually to generate an SSB signal is to first generate a DSB signal and then suppress one of the sidebands by filtering. This procedure is shown in Fig. 5.29.

In practice, these operations are not as easy as they may appear and the primary source of difficulty is in meeting the filter requirements. The sideband filter required must have a very sharp cutoff characteristic at ω_c to reject all frequency components on one side of ω_c and to accept all frequency components on the other side of ω_c. Because it is impossible to obtain an ideal filter characteristic, we must settle for

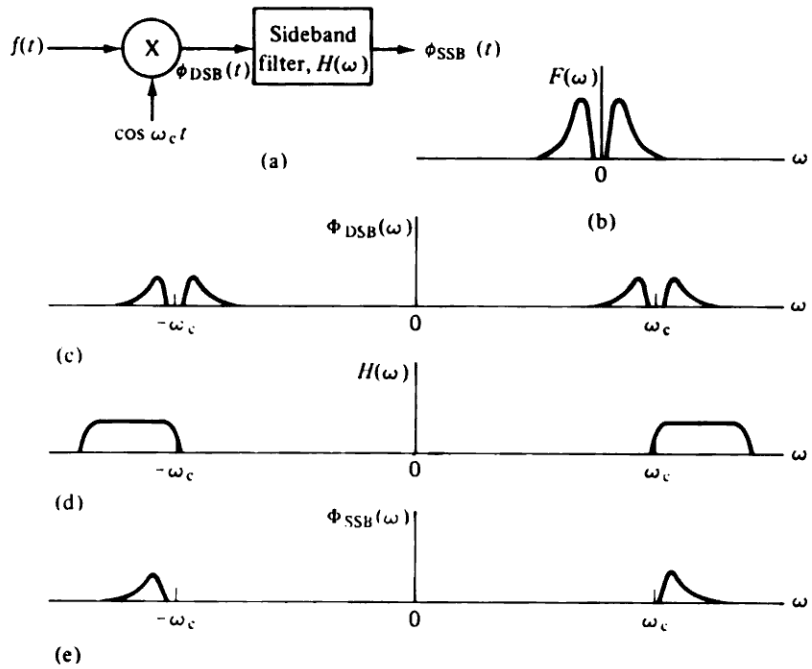


Figure 5.29 SSB modulator using filtering for the upper sideband.

some compromises. First we note that if the modulating signal $f(t)$ does not contain significant low-frequency components there will be no significant frequency components in the vicinity of ω_c after modulation. Therefore a response that does not extend down to zero frequency will allow the use of a filter with a less steep slope. Second, it is usually easier to build a sideband filter at a frequency dictated by the filter components and not by the transmitted frequency. Heterodyning can be employed to translate the spectrum, as desired, after the sideband filtering. Even with these provisions, the design of sideband filters is not easy. In practice, electromechanical-resonator filters are used in the .05–.5 MHz range; crystal-lattice filters are used in the 1–10 MHz range.

It is also possible to generate SSB by a proper phasing of signals. Because it affords some insight into the nature of SSB signals, we shall investigate this method in some detail, first with sinusoidal signals and then in more generality.

Consider now a single-frequency complex exponential signal. This signal has a one-sided spectral density and it is only in taking the real part that the spectral density becomes two-sided. To be specific, let the modulating signal $f(t)$ be $f(t) = e^{j\omega_m t}$ and let the carrier signal be $e^{j\omega_c t}$. The spectral densities of these signals are shown in Fig. 5.30. Multiplying, we get $f(t)e^{j\omega_c t} = e^{j\omega_m t}e^{j\omega_c t}$, shifting the spectral density up in frequency according to the frequency-translation property of the Fourier transform. Now taking the real part (cf. Appendix A), we have

$$\begin{aligned}\mathcal{R}e\{e^{j\omega_m t}e^{j\omega_c t}\} &= \mathcal{R}e\{e^{j\omega_m t}\}\mathcal{R}e\{e^{j\omega_c t}\} - \mathcal{I}m\{e^{j\omega_m t}\}\mathcal{I}m\{e^{j\omega_c t}\} \\ &= \cos\omega_m t\cos\omega_c t - \sin\omega_m t\sin\omega_c t.\end{aligned}$$

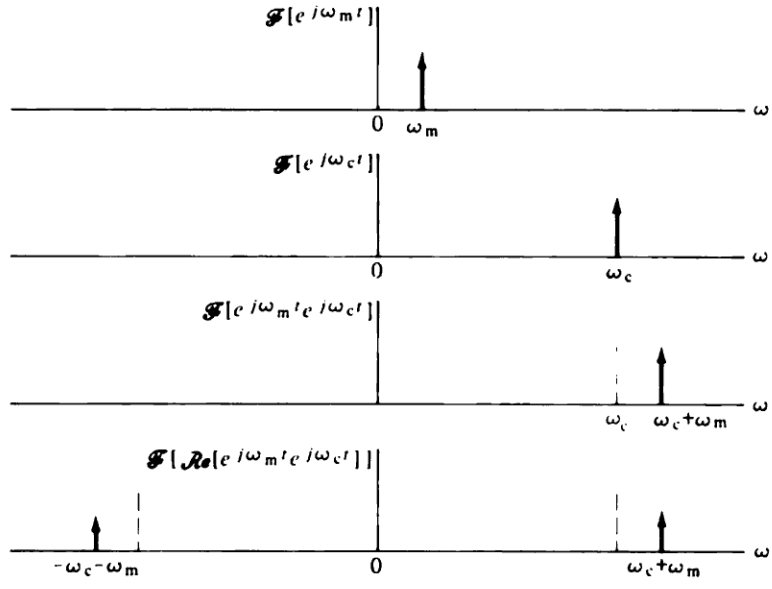


Figure 5.30 Analytical generation of SSB—upper sideband.

The procedure of taking the real part restores the two-sided characteristic to the spectral density so that this result can be realized using real-valued signals. Because this represents the upper sideband, we write

$$\phi_{\mathrm{SSB}_+}(t) = \cos \omega_m t \cos \omega_c t - \sin \omega_m t \sin \omega_c t . \tag{5.31}$$

It is left to the reader to verify that the lower sideband can be obtained in a similar manner by using $f(t) = e^{-j\omega_m t}$ yielding

$$\phi_{\mathrm{SSB}_-}(t) = \cos \omega_m t \cos \omega_c t + \sin \omega_m t \sin \omega_c t . \tag{5.32}$$

Although these results have been derived for the sinusoidal case, they hold true for the more general case and we can write

$$\phi_{\mathrm{SSB}_\mp}(t) = f(t) \cos \omega_c t \pm \hat{f}(t) \sin \omega_c t , \tag{5.33}$$

where $\hat{f}(t)$ is that signal obtained by shifting the phase of $f(t)$ by 90° at each frequency. We shall investigate this in more detail in the following section.

The signals described by Eqs. (5.31) and (5.32) can be produced by balanced modulators provided that both the modulating signal and the carrier are shifted in phase by 90° to form the second terms in these equations. This phase-shift method of generating SSB signals is diagrammed in Fig. 5.31. It is used in low-frequency SSB generation and in the digital generation of SSB signals. A major problem in the design of a phase-shift SSB system is the practical realization of the 90° phase-shift network for $f(t)$ because all frequency components must be shifted by exactly 90°. Thus such systems have very restricted bandwidths.

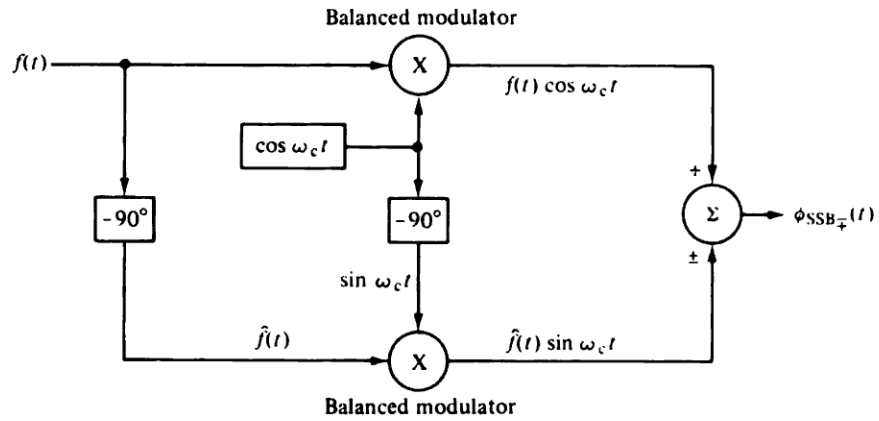


Figure 5.31 Phase-shift method of generating SSB.

DRILL PROBLEM 5.4.1

A DSB-LC signal is generated using a 1-kHz carrier and the input $f(t) = \cos 200\pi t$. The modulation index is 80%. The lower sideband is attenuated (assume ideal filtering). Find an expression for the resulting SSB-LC signal if it develops 0.58 watt across a 1-ohm resistive load.

ANSWER: $\cos(2000\pi t) + 0.4\cos(2200\pi t)$.

5.4.2 ★ Analytic Signals and the Hilbert Transform

The previous discussion on SSB signals demonstrated that real-valued bandpass signals can be represented in terms of complex-valued signals with one-sided spectral densities. Such signals are called *analytic signals* and are useful in describing the general behavior of signals. In other words, an analytic signal is a complex-valued signal whose spectral density is one-sided and whose real part is the original real-valued signal. Not all complex-valued signals are analytic signals but all analytic signals must be complex-valued.

Consider a given real-valued signal $f(t)$; let its analytic signal representation be $z(t)$. Maintaining $f(t)$ as the real part of $z(t)$, we can write

$$z(t) = f(t) + j\hat{f}(t), \tag{5.34}$$

where $\hat{f}(t)$ is yet to be determined. The Fourier transform of Eq. (5.34) is

$$Z(\omega) = F(\omega) + j\hat{F}(\omega). \tag{5.35}$$

To be an analytic signal, $Z(\omega)$ must be one-sided. Let us require that $Z(\omega) = 0$ for $\omega < 0$; a moment's reflection then shows that [recall that the real part of Eqs. (5.34) and (5.35) cannot be changed arbitrarily]

$$\hat{F}(\omega) = jF(\omega), \qquad \omega < 0. \tag{5.36}$$

To maintain a phase characteristic which is an odd function of frequency then requires that

$$\hat{F}(\omega) = -jF(\omega), \qquad \omega > 0. \tag{5.37}$$

Combining Eqs. (5.36) and (5.37), we have

$$\hat{F}(\omega) = \begin{Bmatrix} -jF(\omega) & \omega > 0 \\ jF(\omega) & \omega < 0 \end{Bmatrix} = -jF(\omega)\ \mathrm{sgn}\ (\omega). \tag{5.38}$$

Using Eq. (5.38) in Eq. (5.35) we can now verify that the spectral density of $z(t)$ is one-sided:

$$Z(\omega) = \begin{cases} 2F(\omega) & \omega > 0 \\ 0 & \omega < 0 \end{cases}. \tag{5.39}$$

It is left to he reader to verify that $\mathcal{F}\{z^*(t)\}$ is similar to Eq. (5.39) with the exception that the inequalities are reversed.

The function $\hat{f}(t)$ can by found be taking the inverse Fourier transform

$$\hat{f}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{F}(\omega)e^{j\omega t}\,d\omega. \tag{5.40}$$

The function $\hat{f}(t)$ is called the quadrature function of $f(t)$ because each frequency component of $\hat{f}(t)$ is in phase quadrature (i.e., 90°) with that of $f(t)$. We see now that this is a general relationship and not limited to the sinusoids used in the previous section.

The conditions described by Eq. (5.38) define what is known as the *Hilbert transform;* i.e., the function $\hat{f}(t)$ is the Hilbert transform of the function $f(t)$ if Eq. (5.38) holds true. In specific cases, Eqs. (5.38) and (5.40) can be used to find $\hat{f}(t)$ from $f(t)$. For example, let $f(t)$ be the rectangular pulse signal: $f(t) = \mathrm{rect}\ (t/\tau)$. Using (cf. No. 13, Table 3.1) $F(\omega) = \tau\mathrm{Sa}\ (\omega\tau/2)$, Eqs. (5.38) and (5.40) yield

$$\hat{f}(t) = \frac{1}{\pi}\ln\left|\frac{t + \tau/2}{t - \tau/2}\right|,$$

which is shown in Fig. 5.32. Note that the Hilbert transform becomes infinite at points of finite discontinuity in the signal $f(t)$. Thus a very large instantaneous peak power is required for the transmission of pulse data via SSB.

DRILL PROBLEM 5.4.2

Show that (a) $\hat{\hat{f}}(t) = -f(t)$; (b) $\overline{f^2(t)} = \overline{\hat{f}^2(t)}$.

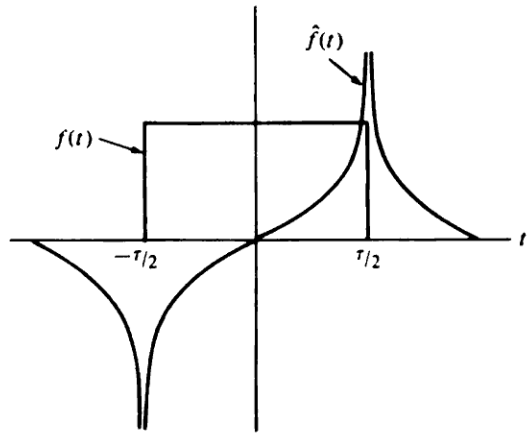


Figure 5.32 The Hilbert transform of a rectangular pulse.

We can also write relationships in the time domain by using (cf. transform pair No. 6, Table 3.1)

$$\operatorname{sgn}(\omega) \leftrightarrow \frac{j}{\pi t}.$$

Application of the time convolution property to Eq. (5.38) then yields

$$\begin{aligned}\hat{f}(t) &= f(t) \circledast \frac{1}{\pi t} \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(\tau)}{t-\tau}\,d\tau. \end{aligned} \tag{5.41}$$

Equation (5.41) is an alternate definition of the Hilbert transform, though not always as useful as the phase-shift interpretation as a result of the improper integral form.

Now let us return to the SSB signal. Given a signal $f(t)$, we form the analytic signal $z(t)$ by adding $\hat{f}(t)$ in quadrature, that is, $z(t) = f(t) + j\hat{f}(t)$. The spectral density of $Z(\omega)$ is one-sided, as shown in Fig. 5.33. Multiplying $z(t)$ by $e^{j\omega_c t}$ translates $Z(\omega)$ to $Z(\omega - \omega_c)$. Finally, the real part of $Z(\omega - \omega_c)$ places half of the spectral density, with reversed phase, symmetrically about $\omega = 0$ as shown in Fig. 5.33. The time domain equivalent is

$$\begin{aligned}\mathcal{R}e\{z(t)e^{j\omega_c t}\} &= \mathcal{R}e\{[f(t) + j\hat{f}(t)]e^{j\omega_c t}\} \\ &= f(t)\cos\omega_c t - \hat{f}(t)\sin\omega_c t,\end{aligned}$$

verifying Eq. (5.33). These relations are for the upper sideband case; it is left to the reader to verify that the use of $z^*(t)$ yields the lower sideband.

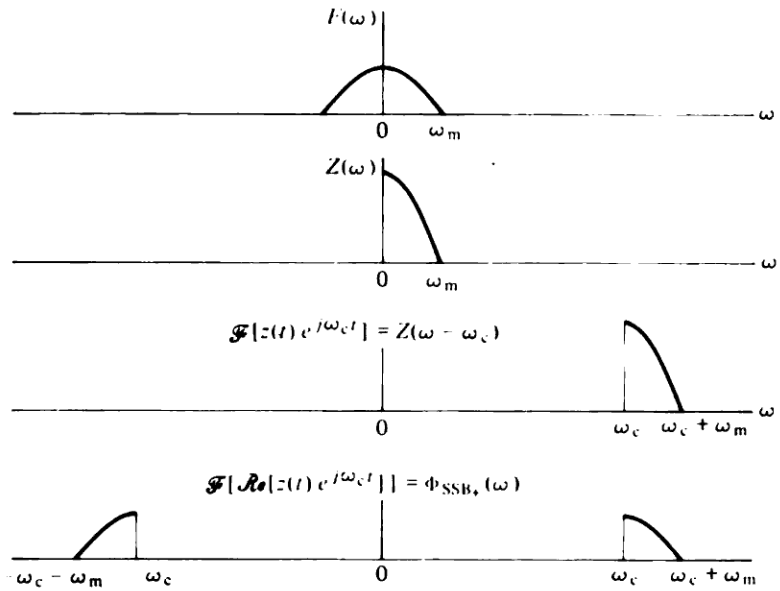


Figure 5.33 Generation of SSB using analytic signals.

5.4.3 Demodulation of SSB Signals

For demodulation, the spectral density of the SSB signal must be translated back to $\omega = 0$. As observed before, this operation can be achieved using synchronous detection. Multiplication of the SSB signal by $\cos \omega_c t$ translates half of each spectral density up in frequency by ω_c rad/sec and half down by the same amount, as shown in Fig. 5.34. That portion shifted up to a frequency of $2\omega_c$ can be filtered out with a low-pass filter. We conclude that the synchronous detector will properly demodulate SSB-SC signals.†

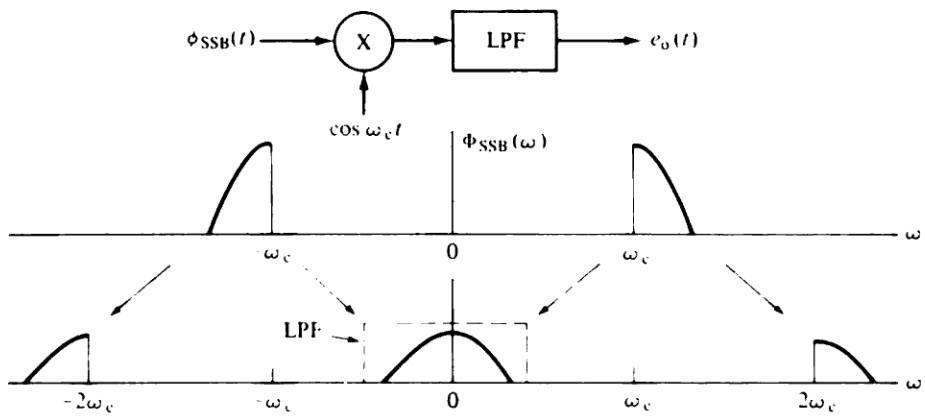


Figure 5.34 Synchronous demodulation of an SSB signal.

† Although the upper-sideband case is illustrated, a similar result holds for the lower-sideband case.

The above case, of course, assumed ideal conditions. Now let us reconsider when frequency and phase errors are present in the demodulation process. The incoming SSB-SC signal can be expressed as [cf. Eq. (5.33)]

$$\phi_{\mathrm{SSB}_{\mp}}(t) = f(t)\cos\omega_c t \pm \hat{f}(t)\sin\omega_c t\,.$$

Let the locally generated carrier signal be $\phi_d(t) = \cos[(\omega_c + \Delta\omega)t + \theta]$, where $(\Delta\omega)$ is the frequency error and θ is the phase error. Multiplying, we have

$$\begin{aligned}\phi_{\mathrm{SSB}_{\mp}}(t)\phi_d(t) &= [f(t)\cos\omega_c t \pm \hat{f}(t)\sin\omega_c t]\cos[(\omega_c + \Delta\omega)t + \theta] \\ &= \tfrac{1}{2}f(t)\{\cos[(\Delta\omega)t + \theta] + \cos[(2\omega_c + \Delta\omega)t + \theta]\} \\ &\quad \mp \tfrac{1}{2}\hat{f}(t)\{\sin[(\Delta\omega)t + \theta] - \sin[(2\omega_c + \Delta\omega)t + \theta]\}\,. \qquad (5.42)\end{aligned}$$

The double-carrier frequency terms can be eliminated using a low-pass filter. Designating the output of the low-pass filter by $e_o(t)$, we have[†]

$$e_o(t) = \tfrac{1}{2}f(t)\cos[(\Delta\omega)t + \theta] \mp \tfrac{1}{2}\hat{f}(t)\sin[(\Delta\omega)t + \theta]\,. \qquad (5.43)$$

Checking, if the frequency error $(\Delta\omega)$ and the phase error θ are both zero, the output is

$$e_o(t) = \tfrac{1}{2}f(t)\,.$$

Therefore a synchronous detector composed of a local oscillator, multiplier, and low-pass filter will correctly demodulate SSB-SC signals if it is correctly synchronized.

To investigate the effects of a phase error, we set $(\Delta\omega) = 0$ and Eq. (5.43) becomes

$$e_o(t) = \tfrac{1}{2}[f(t)\cos\theta \mp \hat{f}(t)\sin\theta]\,. \qquad (5.44)$$

The first term in Eq. (5.44) is the same as occurred for a phase error in the DSB-SC case. However, now there is also a second term present that cannot be filtered out. To investigate, let us rewrite Eq. (5.44) as (cf. Appendix A)

$$e_o(t) = \tfrac{1}{2}\,\mathcal{R}e\{[f(t) \pm j\hat{f}(t)]e^{j\theta}\}\,. \qquad (5.45)$$

Because both $f(t)$ and $\hat{f}(t)$ are present now in the output, we conclude that a phase error in the locally generated carrier will result in phase distortion in the output of the receiver.

It turns out that the human ear can interpret speech despite phase changes in signals, and so phase distortion in SSB-SC demodulation is quite tolerable for voice communications. Even slowly varying frequency errors are tolerable for voice, allowing speech to be understood but with an altered pitch in addition to the phase distortion. However, for other data signals, particularly pulse data, this type of distortion limits the use of SSB-SC systems.

Proceeding in the same way to see the effects of frequency errors, we set $\theta = 0$ and Eq. (5.43) becomes

$$e_o(t) = \tfrac{1}{2}[f(t)\cos(\Delta\omega)t \mp \hat{f}(t)\sin(\Delta\omega)t]\,, \qquad (5.46)$$

[†] Note that the upper set of signs refers to the lower-sideband case, the lower to the upper-sideband.

or

$$e_o(t) = \tfrac{1}{2}\mathcal{R}e\{[f(t) \pm j\hat{f}(t)]e^{j\Delta\omega t}\}. \tag{5.47}$$

Thus frequency errors give rise to spectral shifts as well as to phase distortion in the demodulated output. As long as these spectral shifts are small they can be tolerated in voice communications.

For values of frequency errors ($\Delta\omega$) approaching the bandwidth of the modulating signal $f(t)$, an interesting effect can occur, as demonstrated in Fig. 5.35. Here the frequency error is equal to the bandwidth of $f(t)$ and chosen in such a way that the spectra end up being exactly reversed! This system is a spectrum inverter, exchanging high-frequency spectral components for lows and vice versa. All the signal energy is present but it has been rearranged. This spectral inversion renders speech quite unintelligible and the principle is used in speech scramblers to ensure communication privacy. It is left to the reader to verify that a repetition of these operations will completely unscramble the signal again.

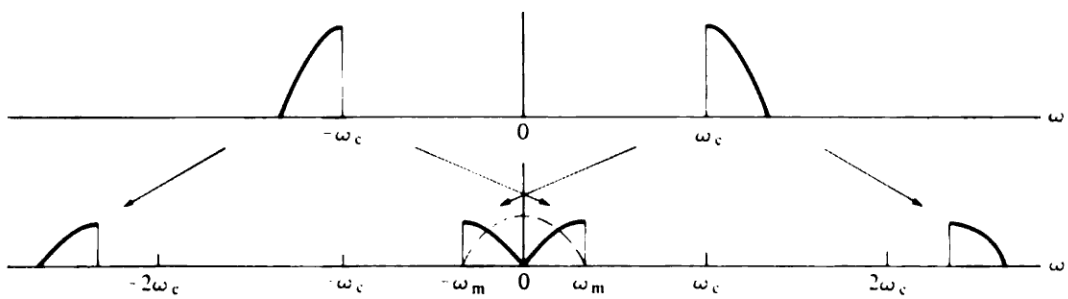


Figure 5.35 Some effects of frequency errors in demodulation of SSB signals.

As with DSB signals, the possibility exists here also of adding large amounts of carrier to SSB signals, which we designate as SSB-LC. An expression for such an SSB-LC signal is

$$\phi(t) = A\cos\omega_c t + f(t)\cos\omega_c t \mp \hat{f}(t)\sin\omega_c t. \tag{5.48}$$

Apart from a dc level introduced by the carrier term, the original signal $f(t)$ can always be recovered from $\phi(t)$ using synchronous detection. If the carrier is large, however, envelope detection can also be used. The envelope of Eq. (5.48) can be written as

$$r(t) = \sqrt{[A + f(t)]^2 + [\hat{f}(t)]^2}. \tag{5.49}$$

Equation (5.49) may be rewritten as

$$r(t) = A\sqrt{1 + \frac{2f(t)}{A} + \frac{f^2(t)}{A^2} + \frac{\hat{f}^2(t)}{A^2}}. \tag{5.50}$$

If the carrier is much larger than the SSB-SC envelope, the last two terms in Eq. (5.50) can be dropped to give

$$r(t) \simeq A\sqrt{1 + \frac{2f(t)}{A}}. \tag{5.51}$$

When we use the binomial expansion† and discard higher-order terms, Eq. (5.51) becomes

$$r(t) \approx A\left[1 + \frac{f(t)}{A}\right] = A + f(t). \tag{5.52}$$

When we discard the dc term introduced by the carrier, the SSB-LC signal can then be demodulated correctly, using an envelope detector.

We conclude that SSB-LC signals can be demodulated by using an envelope detector. Such signals could be received on commercial AM receivers and would require one-half the bandwidth of present AM transmissions. However, the proportionate amount of carrier required is considerably more than that used for DSB-LC.

Now we pose an interesting question. Does a modulation method exist that affords the spectral economy of SSB and also provides a signal that can be detected using envelope detection without a large carrier term? The answer, surprisingly, is—within some restrictions—yes, providing that a combination of amplitude modulation and phase modulation is used. Although the theory is not easy, a modulation method that approximately achieves the above objectives has the form‡

$$\phi(t) = a(t)\cos\left[\omega_c t + k_p \widehat{\log a(t)}\right] \tag{5.53}$$

where

$$a(t) = \sqrt{f(t)},$$

$$k_p = \text{phase modulation constant (cf. Section 6.1).}$$

The spectral density of the signal described by Eq. (5.53) is not strictly band-limited to one sideband, and there is some spectral spillover. However, it can be demodulated by using an envelope detector, preferably one with a square-law dependence, and without a large carrier term. Note that the generation of such a signal requires the nonlinear processing of the input signal $f(t)$. Signals that permit the use of envelope detection without using bandwidth in excess of that of the modulating signal are sometimes called *compatible single-sideband* (CSSB) signals.

† See Appendix A.

‡ See, e.g., K. H. Powers, "The Compatibility Problem in Single-Sideband Transmission," *Proceedings of the IRE*, vol. 48 (August 1960): 1431; or B. F. Logan and M. R. Schroeder, "A Solution to the Problem of Compatible Single-Sideband Transmission," *IRE Transactions on Information Theory*, vol. IT-8 (September 1962): S252.

DRILL PROBLEM 5.4.3

A given DSB-SC signal is generated using a 20.000 MHz carrier and a 0–3.0 kHz input signal. One sideband is attenuated using an SSB filter (assume ideal). The resulting SSB-SC signal is to be demodulated using a superheterodyne receiver whose local oscillator operates at a frequency f_o (which is above the incoming signal). The IF amplifier passes all frequency components within the frequency range 10.000–10.003 MHz (assume ideal). Following the IF amplifier is a product detector operating at a frequency f_d followed by a low-pass filter. Determine the numerical values of the frequencies f_c, f_d such that this receiver will correctly demodulate the SSB signal if (a) it is an upper-sideband transmission; (b) it is a lower-sideband transmission.

[The small frequency shifts required are often obtained by heterodyning with an audio frequency oscillator, known as a *beat frequency oscillator* (BFO).]

ANSWER: (a) 30.003; 10.003; (b) 30.000; 10.000 MHz.

DRILL PROBLEM 5.4.4

Determine and plot the envelope of each of the following large-carrier signals and compare (a) DSB-LC; (b) SSB-LC. Assume an input of the form $f(t) = \cos \omega_m t$ and normalize your answer such that the unmodulated carrier magnitude is unity and the peak envelope magnitude at $t = 0$ is 1.30.

ANSWER: (a) $1.00 + 0.30 \cos \omega_m t$; (b) $\sqrt{1.09 + 0.60 \cos \omega_m t}$; the plots will differ slightly.

5.5 VESTIGIAL-SIDEBAND (VSB) MODULATION

The generation of SSB signals may be quite difficult when the modulating signal bandwidth is wide or where one cannot disregard the low-frequency components. To conserve spectrum space, a compromise can be made between SSB and DSB in what is known as *vestigial-sideband* (VSB) modulation. In VSB modulation, only a portion of one sideband is transmitted in such a way that the demodulation process reproduces the original signal. The partial suppression of one sideband reduces the required bandwidth from that required for DSB but does not match the spectrum efficiency of SSB. If a large carrier is also transmitted, the desired signal can be recovered using an envelope detector. If no carrier is sent, the signal can be recovered using a synchronous detector or the injected carrier method discussed in the preceding section.

This filtering operation can be represented by a filter $H_V(\omega)$ that passes some of the lower (or upper) sideband and most of the upper (or lower) sideband. The magni-

tude characteristic of such a filter is shown in Fig. 5.36. The spectral density of the resulting vestigial-sideband signal is

$$\Phi_{VSB}(\omega) = [\tfrac{1}{2}F(\omega - \omega_c) + \tfrac{1}{2}F(\omega + \omega_c)]H_V(\omega)\,. \tag{5.54}$$

We wish to determine the requirements on this filter such that a synchronous detector will yield $f(t)$ without distortion.

Figure 5.36 VSB filtering.

The output of the synchronous detector is

$$e_o(t) = [\phi_{VSB}(t)\cos\omega_c t]_{LP}\,,$$

or

$$E_o(\omega) = \tfrac{1}{4}F(\omega)H_V(\omega + \omega_c) + \tfrac{1}{4}F(\omega)H_V(\omega - \omega_c)\,. \tag{5.55}$$

For faithful reproduction of $f(t)$, we then require that

$$[H_V(\omega - \omega_c) + H_V(\omega + \omega_c)]_{LP} = \text{constant}, \qquad |\omega| < \omega_m\,. \tag{5.56}$$

The frequency translations indicated in Eq. (5.56) are illustrated in Fig. 5.37. Note that, at least on a magnitude basis, Eq. (5.56) is satisfied if $|H_V(\omega)|$ is antisymmetric about the carrier frequency ω_c. Motivated by this observation, we let the constant in Eq. (5.56) be $2H_V(\omega_c)$. Under this condition Eq. (5.56) becomes

$$[H_V(\omega - \omega_c) - H_V(\omega_c)] = -[H_V(\omega + \omega_c) - H_V(\omega_c)]\,, \tag{5.57}$$

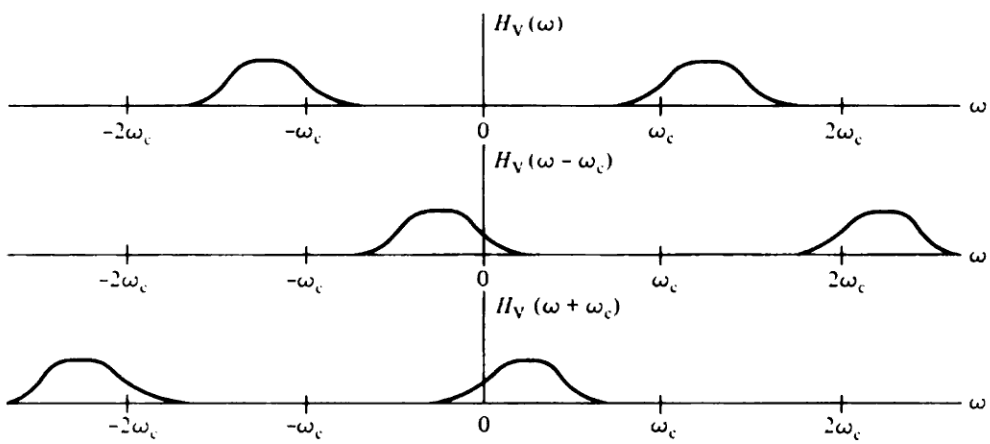


Figure 5.37 Characteristics of VSB filtering.

showing the required antisymmetry about the carrier frequency. Therefore the required filtering for a VSB signal is a complementary filter symmetry about the carrier frequency, as shown in Fig. 5.36. The required phase relationships are not as easy to satisfy and VSB is generally used in cases where the correct phase response is not of prime concern. Although synchronous detection is assumed here, these same principles hold also when a large carrier is present and an envelope detector is used.

Vestigial-sideband modulation is used to advantage in the video portion of public television systems. In a television transmission, 525 lines of video (picture) information are sent each $\frac{1}{30}$ of a second (that is, 15,750 lines per second—the horizontal trace frequency).[†] Allowing time for retrace and synchronization, this requires a minimum video bandwidth of 4 MHz to transmit an array of picture elements. (The spectral density of typical video, however, is not flat but heavily weighted toward the lower frequencies.) Use of DSB modulation for video transmission would require a frequency allocation of 8 MHz per channel. This is judged excessive and a version of VSB modulation is used to decrease the required video bandwidth to about 5 MHz. The spectrum of a public television channel is illustrated in Fig. 5.38.[‡] The filter characteristic is positioned not about the carrier in this case but 1 MHz below the carrier. The receiver filtering is used to complete the vestigial-sideband characteristic, as indicated in Fig. 5.38.

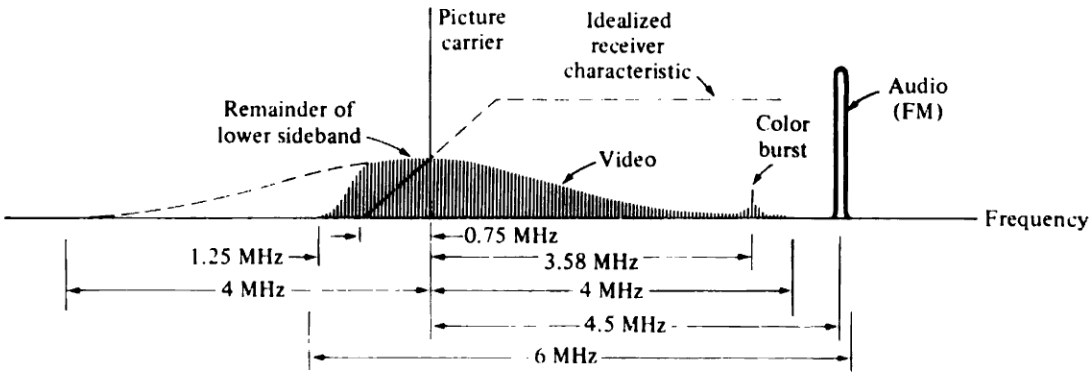


Figure 5.38 Spectrum of a public television transmission.

5.6 A TIME-REPRESENTATION OF BANDPASS NOISE

At this point, we return to a discussion of bandpass noise. We have not attempted to develop a description of noise in terms of time functions because the noise is random.

[†] Standards for North America; a standard of 625 lines sent each $\frac{1}{25}$ of a second is generally used where 50-Hz power is used. Commercial television signal transmissions are described in Appendix D.

[‡] Note that the video spectrum appears about the 15.75-kHz sampling lines. Here the (horizontal) sampling frequency is much lower than the signal bandwidth. The sampling theorem does *not* state that the signal cannot be recovered but merely that the signal cannot be recovered by filtering alone. Thus a television receiver must utilize time synchronization techniques to recover the original video information.

Yet for bandpass systems, which limit the bandwidth of the noise, some of the noise fluctuations are restricted and it is possible to write some rather general time relations.

As the bandwidth of the noise becomes small compared to the center frequency, it becomes possible to approximate it with a phasor representation. Such a phasor representation is shown in Fig. 5.39. Here $n_c(t)$ is that portion of the noise which is the in-phase component and $n_s(t)$ is the quadrature component in the phasor representation. Both $n_c(t)$ and $n_s(t)$ are random and thus the phasor sum has a random amplitude a_n and phase θ_n.

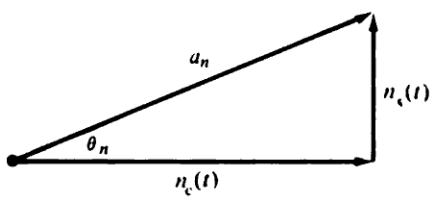


Figure 5.39 A phasor representation of narrowband random noise.

We can write out the complete phasor representation of the narrowband random noise using complex notation,

$$[n_c(t) + jn_s(t)]e^{j\omega_0 t}, \tag{5.58}$$

where ω_0 is the center frequency. Both $n_c(t)$ and $n_s(t)$ are slowly varying compared to $e^{j\omega_0 t}$ as a result of the narrow bandwidth restriction. Taking the real part of Eq. (5.58) to obtain an expression for $n(t)$, we have

$$n(t) = \mathscr{R}e\{[n_c(t) + jn_s(t)]e^{j\omega_0 t}\}.$$

Using an identity from complex algebra (cf. Appendix A) gives us[†]

$$n(t) = n_c(t) \cos \omega_0 t - n_s(t) \sin \omega_0 t. \tag{5.59}$$

Equation (5.59) is the desired representation of $n(t)$. Although derived for the narrowband case here, this representation can be extended to the more general bandpass case by properly defining $n_c(t)$ and $n_s(t)$. Therefore Eq. (5.59) is known as the bandpass representation of noise. The terms $n_c(t)$ and $n_s(t)$ are low-pass voltages whose fluctuations are limited by the bandwidth of the bandpass noise.

We are interested here in the result of bandpass noise applied to the synchronous detector, shown in Fig. 5.40. Multiplying both sides of Eq. (5.59) by $\cos \omega_0 t$:

$$n(t) \cos \omega_0 t = n_c(t) \cos^2 \omega_0 t - n_s(t) \sin \omega_0 t \cos \omega_0 t. \tag{5.60}$$

If we use the trigonometric identities $2 \cos^2 A = 1 + \cos 2A$, $2 \sin A \cos A = \sin 2A$, and if we retain only the low-pass (LP) terms, Eq. (5.60) becomes

$$e_o(t) = [n(t) \cos \omega_0 t]_{LP} = (\tfrac{1}{2}) n_c(t). \tag{5.61}$$

[†] Given two complex quantities z_1, z_2; then

$$\mathscr{R}e\{z_1 z_2\} = \mathscr{R}e\{z_1\} \mathscr{R}e\{z_2\} - \mathscr{I}m\{z_1\} \mathscr{I}m\{z_2\}.$$

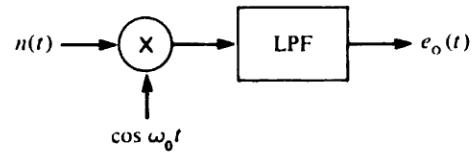


Figure 5.40 Synchronous detection of a waveform $n(t)$.

The power spectral density of Eq. (5.61) is, using Eqs. (3.45) and (4.17),

$$S_{n_c}(\omega) = \left\{ \lim_{T \to \infty} \left[\frac{|N_T(\omega - \omega_0) + N_T(\omega + \omega_0)|^2}{T} \right] \right\}_{\text{LP}} . \tag{5.62}$$

For random noise, the average of the cross products in Eq. (5.62) goes to zero, giving the result

$$S_{n_c}(\omega) = [S_n(\omega - \omega_0) + S_n(\omega + \omega_0)]_{\text{LP}} . \tag{5.63}$$

Thus, the power spectral density of the product of two independent signals is given by the convolution of their power spectral densities. In more general cases, however, the presence of the cross-product term must be considered. (The subject of independence is considered in Chapter 8.)

The case for $n_s(t)$ follows in the same manner by first multiplying both sides of Eq. (5.59) by $\sin \omega_0 t$, yielding

$$S_{n_s}(\omega) = [S_n(\omega - \omega_0) + S_n(\omega + \omega_0)]_{\text{LP}} . \tag{5.64}$$

Combining Eqs. (5.63) and (5.64), we have

$$S_{n_c}(\omega) = S_{n_s}(\omega) = [S_n(\omega - \omega_0) + S_n(\omega + \omega_0)]_{\text{LP}} . \tag{5.65}$$

A sketch of a given bandpass noise spectral density and the resulting low-pass $S_{n_c}(\omega)$ and $S_{n_s}(\omega)$ is shown in Fig. 5.41.

It can be seen from Fig. 5.41 that the area under $S_n(\omega)$ is equal to the area under $S_{n_c}(\omega)$ or $S_{n_s}(\omega)$. Therefore it follows that their mean-square values are equal; that is,

$$\overline{n^2(t)} = \overline{n_c^2(t)} = \overline{n_s^2(t)} . \tag{5.66}$$

These results apply to bandpass *random* noise. It is the random nature of the noise that tends to distribute the noise components over both the cosine and sine terms. This is in direct contrast to a deterministic signal in which we can control the phase to give only cosine or sine terms.

The mean-square value of bandpass random noise, using Eq. (5.59), is

$$\overline{n^2(t)} = \tfrac{1}{2}\overline{n_c^2(t)} + \tfrac{1}{2}\overline{n_s^2(t)} . \tag{5.67}$$

Our conclusion, based on Eqs. (5.66) and (5.67), is that the mean noise power is divided equally between the cosine and sine terms.

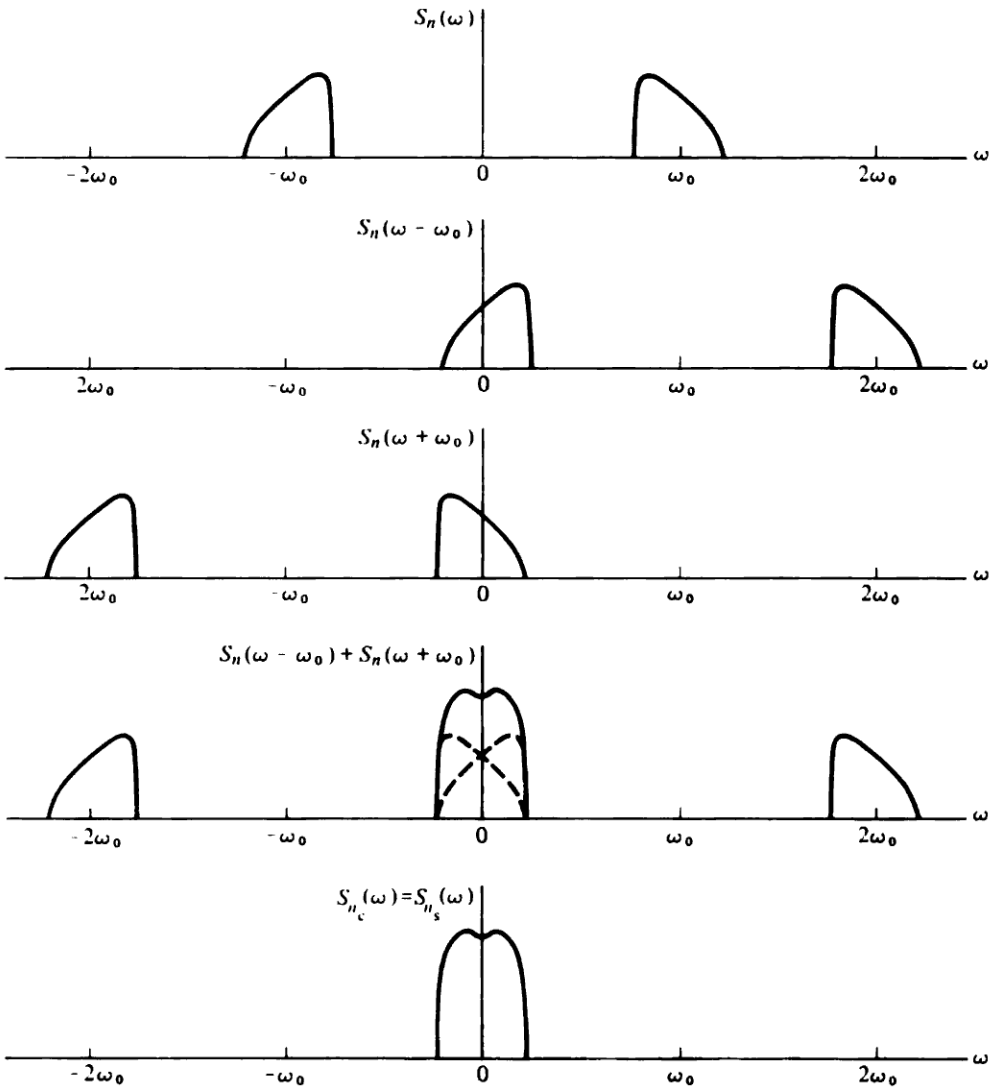


Figure 5.41 In-phase and quadrature noise spectral densities for bandpass noise.

5.7 SIGNAL-TO-NOISE RATIOS IN AM RECEPTION

In order to discuss and compare various modulation systems, we need a standard basis for the measurement of performance. A convenient quantity is the ratio of the mean-squared signal to the mean-squared noise, where both quantities are measured at the same point in the system. This signal-to-noise (power) ratio is designated by the ratio of the capital letters S/N, with the appropriate subscripts.

We assume that the signals that are transmitted and received are bandpass signals centered about a carrier frequency, f_c. This allows us to make use of the in-phase and quadrature representations of bandpass noise that are described in the preceding section.

5.7.1 DSB-SC

For DSB-SC signals, the receiver consists of a synchronous detector—a sinusoidal signal generator at $f = f_c$, a multiplier, and a low-pass filter. The input signal to the detector is $f(t) \cos \omega_c t$, so that

$$S_i = \overline{[f(t) \cos \omega_c t]^2} = \tfrac{1}{2}\overline{f^2(t)}. \tag{5.68}$$

The useful output signal is $\frac{1}{2}f(t)$ [the $\frac{1}{2}$ arises from $\overline{\cos^2 \omega_c t}$] so that

$$S_o = \overline{[\tfrac{1}{2}f(t)]^2} = \tfrac{1}{4}\overline{f^2(t)} = \tfrac{1}{2}S_i. \tag{5.69}$$

The noise output of the detector is $n_i(t) \cos \omega_c t$, filtered by the low-pass filter. Letting $n_d(t)$ be the output of the detector and $n_o(t)$ be the output of the low-pass filter, and using the bandpass noise representation [cf. Eq. (5.59)] gives us

$$n_d(t) = \tfrac{1}{2}n_c(t) + \tfrac{1}{2}n_c(t) \cos 2\omega_c t - \tfrac{1}{2}n_s(t) \sin 2\omega_c t, \tag{5.70}$$

$$n_o(t) = \tfrac{1}{2}n_c(t). \tag{5.71}$$

Defining

$$\overline{n_i^2(t)} = N_i,$$

we have

$$N_o = \overline{n_o^2(t)} = \tfrac{1}{4}\overline{n_c^2(t)} = \tfrac{1}{4}\overline{n_i^2(t)} = \tfrac{1}{4}N_i, \tag{5.72}$$

where use has been made of Eq. (5.66). Combining Eqs. (5.69) and (5.72) yields

$$\frac{S_o}{N_o} = 2\frac{S_i}{N_i} \quad \text{for DSB-SC.} \tag{5.73}$$

Thus the detector improves the signal-to-noise ratio in a DSB-SC system by a factor of 2. This improvement results from the fact that the coherent (synchronous) detector rejects the out-of-phase (quadrature) noise component in the input noise, thereby halving the mean-square noise power.

The use of a synchronous detector for the large-carrier case can be derived easily from the above result by noting that $f(t)$ is replaced by $[A + f(t)]$, so that $S_i = \frac{1}{2}A^2 + \frac{1}{2}\overline{f^2(t)}$ [assuming that $\overline{f(t)}$ is zero] and

$$\frac{S_o}{N_o} = \frac{2\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}\frac{S_i}{N_i}. \tag{5.74}$$

Because $|f(t)| < A$ for large-carrier systems, the signal-to-noise ratio is somewhat poorer than for a suppressed-carrier system.

5.7.2 SSB-SC

Using the general representation for an SSB-SC signal [cf. Eq. (5.33)],

$$\phi(t) = f(t) \cos \omega_c t \pm \hat{f}(t) \sin \omega_c t,$$

we have

$$S_i = \overline{\phi^2(t)} = \tfrac{1}{2}\overline{f^2(t)} + \tfrac{1}{2}\overline{\hat{f}^2(t)}. \tag{5.75}$$

However, we also have [because $\hat{F}(\omega)$ is only shifted in phase from $F(\omega)$]

$$|F(\omega)|^2 = |\hat{F}(\omega)|^2. \tag{5.76}$$

If we integrate and use Parseval's theorem, Eq. (5.76) yields

$$\overline{f^2(t)} = \overline{\hat{f}^2(t)}. \tag{5.77}$$

Equation (5.75) then becomes

$$S_i = \overline{f^2(t)}. \tag{5.78}$$

The useful output signal is $\frac{1}{2}f(t)$ so that

$$S_o = \overline{[\tfrac{1}{2}f(t)]^2} = \tfrac{1}{4}\overline{f^2(t)} = \tfrac{1}{4}S_i. \tag{5.79}$$

Combining Eqs. (5.72) and (5.79) we get

$$\frac{S_o}{N_o} = \frac{S_i}{N_i} \quad \text{for SSB-SC.} \tag{5.80}$$

Because the process of SSB detection is a straightforward frequency translation, all of the signal and noise components are transferred, unmodified, to low frequencies. Therefore the relationship between the signal and the noise remains unchanged by SSB detection and the predetection and postdetection S/N ratios are equal.

Is the DSB-SC system then superior to the SSB system? Not where noise power is proportional to bandwidth because the DSB-SC system requires twice the bandwidth of the SSB system and therefore has twice the noise power. We conclude that *the performance of SSB-SC is identical to that of DSB-SC from the point of view of noise improvement in the presence of white noise.*

5.7.3 DSB-LC: The Envelope Detector

The synchronous detector used in the SSB-SC and DSB-SC systems above could also be used here, but to utilize the full advantages of AM the envelope detector is more commonly used. However, the envelope detector is nonlinear and at best we shall only be able to establish some bounds on its performance.

Using the bandpass noise representation, the input signal and noise can be written as

$$s_i(t) + n_i(t) = [A + f(t)]\cos\omega_c t + n_c(t)\cos\omega_c t - n_s(t)\sin\omega_c t. \tag{5.81}$$

The envelope of this signal is (see Fig. 5.42)

$$r(t) = \sqrt{\{[A + f(t)] + n_c(t)\}^2 + \{n_s(t)\}^2}. \tag{5.82}$$

For high input signal-to-noise, this can be approximated by use of the binomial expansion to give

$$r(t) \approx A + f(t) + n_c(t). \tag{5.83}$$

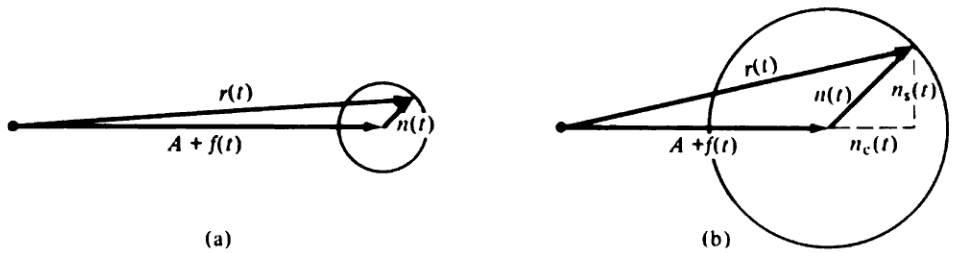


Figure 5.42 Phasor diagram of DSB-LC signals for (a) large S/N and (b) small S/N ratios.

The useful signal here is $f(t)$, while the noise is $n_c(t)$. The detector output gives

$$S_o = \overline{f^2(t)}, \tag{5.84}$$

$$N_o = \overline{n_c^2(t)} = \overline{n_i^2(t)} = N_i; \tag{5.85}$$

also

$$S_i = \overline{\{[A + f(t)] \cos \omega_c t\}^2} = \tfrac{1}{2}A^2 + \tfrac{1}{2}\overline{f^2(t)}. \tag{5.86}$$

Combining Eqs. (5.84) and (5.86), we get

$$\frac{S_o}{N_o} = \frac{2\overline{f^2(t)}}{A^2 + \overline{f^2(t)}} \frac{S_i}{N_i}, \quad \text{for AM, large signal-to-noise.} \tag{5.87}$$

Note that Eq. (5.87) is identical to a previous result using synchronous detection, as expressed in Eq. (5.74). Thus for the large signal-to-noise case, envelope demodulation in the presence of noise has the same performance quality as synchronous demodulation.

In the particular case of sinusoidal modulation, $f(t) = mA \cos \omega_m t$ and Eq. (5.87) can be rewritten as

$$\frac{S_o}{N_o} = \frac{2m^2}{2 + m^2} \frac{S_i}{N_i}, \quad \text{for AM, large signal-to-noise, sinusoidal modulation.} \tag{5.88}$$

The maximum improvement in the signal-to-noise ratio is $\frac{2}{3}$ at 100% modulation.

Sometimes it is convenient to refer the output signal-to-noise ratio to the *carrier-to-noise* ratio (CNR) at the output of the IF amplifiers. Noting that the mean-square value of the carrier is $S_c = A^2/2$ and that the mean-square value of the noise at the output of the IF amplifiers is $N_c = N_i$, we find Eqs. (5.84) and (5.85) give

$$\frac{S_o}{N_o} = \frac{2\overline{f^2(t)}}{A^2} \frac{S_c}{N_c}. \tag{5.89}$$

In the particular case of sinusoidal modulation, $f(t) = mA \cos \omega_m t$ and Eq. (5.89) can be rewritten as

$$\frac{S_o}{N_o} = m^2 \frac{S_c}{N_c}. \tag{5.90}$$

In other words, the maximum output signal-to-noise ratio of an AM system is equal to the carrier-to-noise ratio (CNR) at the output of the IF amplifiers. This form proves convenient in comparing the performance of AM with other types of modulation.

For the large-noise case, we can no longer make the approximation of Eq. (5.83). Rewriting the expression for the signal envelope of Eq. (5.82), we have

$$r(t) = \left\{A^2\left[1 + \frac{2f(t)}{A} + \frac{2n_c(t)}{A} + \frac{2n_c(t)f(t)}{A^2} + \frac{n_c^2(t)}{A^2} + \frac{f^2(t)}{A^2}\right] + n_s^2(t)\right\}^{1/2}. \tag{5.91}$$

This result shows that not only are there additive noise terms present but the desired signal is multiplied by noise. Thus for large-noise conditions the desired signal is severely distorted by the envelope detector.

This effect may be visualized with the aid of Fig. 5.42. The random noise $n(t)$ adds to the modulated carrier signal with a random phase and magnitude. The envelope $r(t)$ is controlled primarily by the modulated carrier for high S/N conditions. When the magnitude of the noise (on the average) is on the same order as that of the modulated carrier signal, however, this predominance in the control of the envelope vanishes. In fact, for $S/N < 1$, the noise takes over the control of the envelope.

We can combine some of the preceding results for the case of a white input noise power-spectral density of $\eta/2$ W/Hz and a modulating bandwidth of f_m Hz. For this case, we have $N_i = 2\eta f_m$ for DSB systems and $N_i = \eta f_m$ for SSB systems. Thus Eqs. (5.73) and (5.80) can be written as

$$\frac{S_o}{N_o} = \frac{S_i}{\eta f_m} \tag{5.92}$$

for both systems using synchronous detection. Note that for synchronous detection, the S/N improvement is independent of the input S/N ratio and that the effects of filtering before detection (predetection filtering) and after detection (postdetection filtering) are identical. A graph of Eq. (5.92) is shown in Fig. 5.43.

The output S/N ratio for envelope detection for DSB and SSB systems is shown also in Fig. 5.43.[†] For high input S/N ratio, the output S/N ratio for envelope detection approaches that for synchronous detection. For lower input S/N ratios, however, the output S/N decreases more rapidly than for the synchronous detection case with an S/N ratio that is less than about 10–15 dB. However, this S/N degradation effect is not very abrupt in AM systems.

For good voice communications, intelligibility and fidelity demand a high signal-to-noise ratio (e.g., 30 dB) and systems designs usually stay well above the signal-to-noise ratios at which the S/N degradation effects become important. In television

[†] The S/N analysis for envelope detection becomes quite involved; see, among others, W. D. Gregg, *Analog and Digital Communications*, New York: Wiley, 1977, Chapter 7.

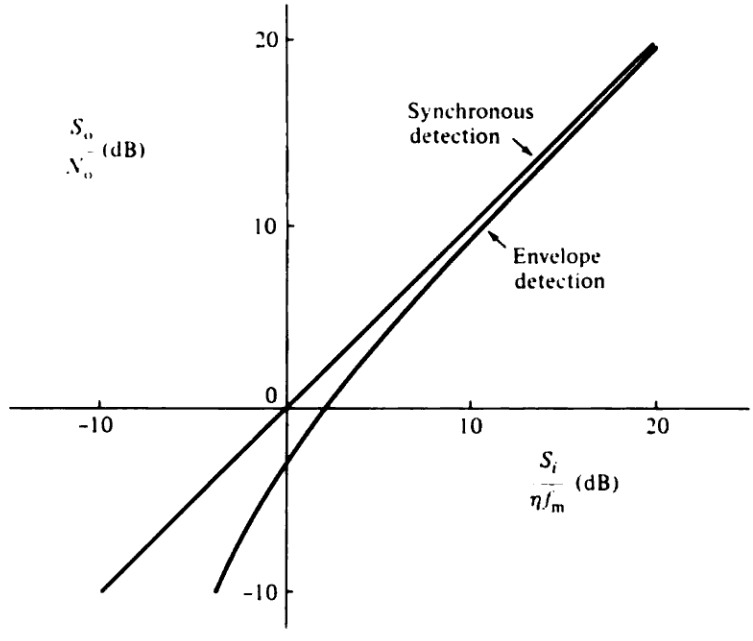


Figure 5.43 Signal-to-noise performance of AM systems.

systems, it has been found that the reception of video information using envelope detection requires a $S/N > 33$ dB for good picture quality, with 28–33 dB in the "passable" category, 23–28 dB in the "marginal" category, and 17–23 dB in the "inferior" category. Picture quality for a $S/N < 17$ dB in commercial television systems is judged unusable.[†]

The degradation of performance with low S/N is a property of envelope detectors; we observed no such effect with synchronous detectors. With synchronous detection, the output signal and noise always remain additive and the signal-to-noise improvement holds under all noise conditions.

Therefore, a comparison of envelope detectors with synchronous detectors in varying S/N conditions shows that performance of the two detectors is comparable when the S/N ratio is high. However, the nonlinear behavior of the envelope detector makes it increasingly unacceptable as the relative amount of input noise increases, and the S/N performance declines.

Finally, if we normalize all of the above AM detection methods to equal sideband input power and do not include the carrier power after detection, the output signal-to-noise ratios are all the same for large input signal-to-noise conditions. Synchronous detection methods continue to give this performance for low signal-to-noise conditions whereas the output of the envelope detector becomes unusable.

[†] See, for example, K. B. Benson (ed.), *Television Engineering Handbook*, New York: McGraw-Hill, 1986, Chapter 13.

5.8 PROPAGATION EFFECTS

Communication systems frequently use electromagnetic propagation to convey information from point to point. Because the choice of modulation may be influenced by the propagation, we briefly digress to investigate some of these effects.

Electromagnetic radiation through space has its average power divided equally between an electric field and a magnetic field. In free space the directions of the electric field, the magnetic field, and the direction of propagation are all at right angles. The polarization of an electromagnetic wave is taken to be in the direction of the electric field. Hence a wave is said to be vertically polarized if its electric field is in the vertical direction. In the presence of a ground plane, vertical polarization is usually preferred when antennas are within a wavelength of the ground, while horizontal is preferred when antennas are several wavelengths above ground.

As a result of electromagnetic theory, the minimum dimension of an antenna for efficient radiation of radio frequency energy is one-half the wavelength of the radio frequency. At low frequencies, this length becomes excessive and for vertical polarization the earth is used as one half of the antenna, as shown in Fig. 5.44. This type of antenna is used, for example, in the commercial AM broadcast band. At higher frequencies (i.e., shorter wavelengths), horizontally polarized half-wave antennas are commonly used, often in arrays to improve their performance. At very high frequencies, physical dimensions of the antenna become large compared to a wavelength, and optical approximations can be used.

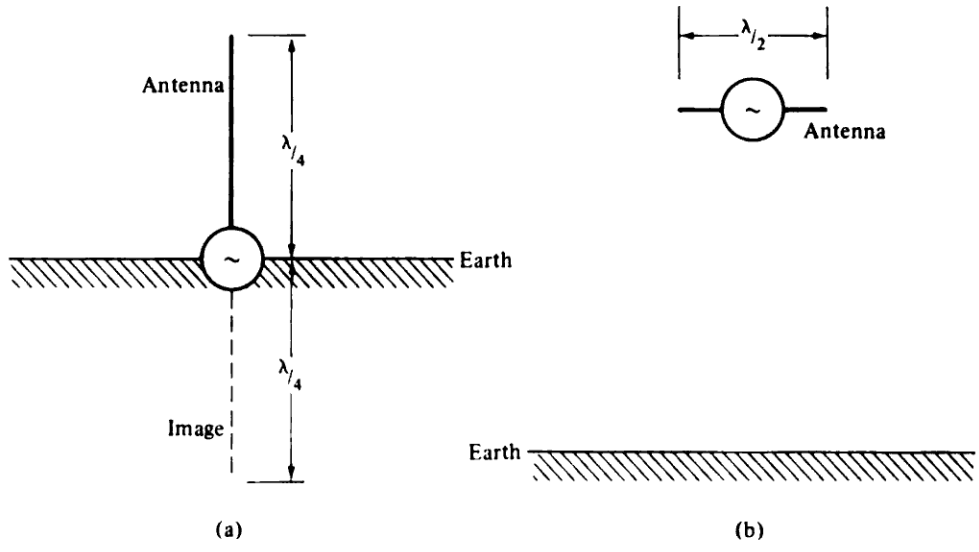


Figure 5.44 The basic half-wave antenna: (a) vertical polarization, low frequency; (b) horizontal polarization, high frequency.

In general, an electromagnetic wave radiated from an antenna is composed of a ground wave that arises from currents induced in the earth and travels along the earth's surface and a sky wave that is propagated through space. At very low frequencies, the ground wave can be used to communicate reliably over long distances while sky-wave propagation is generally used at high frequencies.

The ground wave loses energy as a result of dispersion along the earth's surface and energy dissipated in the earth. These losses vary directly with frequency and the resistance of the earth. The attenuation above 10 MHz is so high that ground wave propagation is of little practical value at these frequencies. Because the tangential component of the electric field cannot exist along the ground interface, vertical polarization is used where propagation via ground wave is desired.

The sky wave leaves the antenna and travels essentially in a straight path. Some of the sky wave's energy leaves the antenna at an angle to the horizontal and travels upward until it enters an ionized layer, known as the ionosphere, about 30–70 miles above the earth's surface. In this region the path of the wave is altered by refractive effects that are dependent on the intensity of ionization and the wave frequency. Depending on the conditions, the path of the sky wave may be bent downward enough to cause the radiated energy to return to earth, as illustrated in Fig. 5.45. The losses in this mechanism are small and the signal strength of the sky wave may be very strong. As a result, long-distance communications can be accomplished by use of the sky wave.

The ionization within the ionosphere varies with solar activity so that these effects vary with the time of day and season of the year. These effects are more pronounced at night and it is not unusual to receive strong signals in the AM broadcast band over a thousand-mile distance at night while the ground wave signal strength is weak after a few hundred miles. To avoid undue interference from these effects, some AM stations are permitted to operate only during daylight hours.

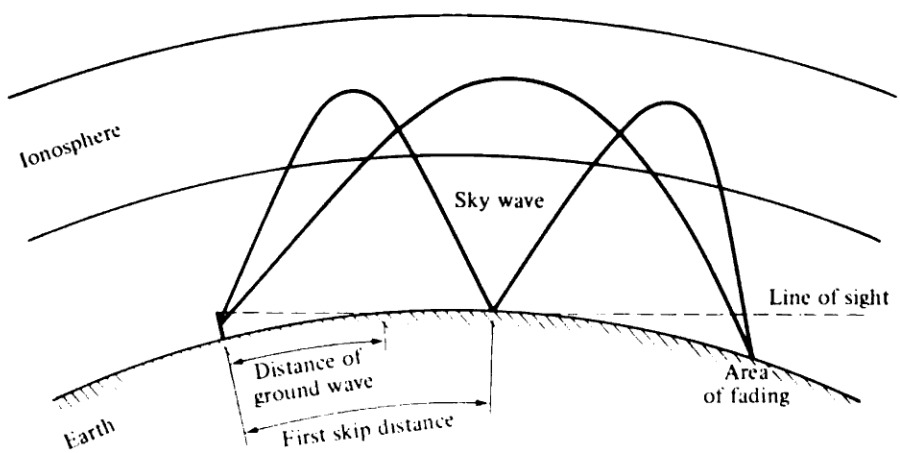


Figure 5.45 Propagation of ground wave and sky wave from an antenna. (Not drawn to scale.)

When two or more radio waves arrive at a point of reception along different paths, their phase relations may not be the same. The results of this multipath propagation can cause wide variations in signal intensity at the point of reception. Cancellation of one wave by another results in a loss of signal strength and is called *fading*. If the fading is highly sensitive to the frequency of propagation, different parts of the spectrum of a signal may undergo different amounts of fading. This effect is called *selective fading*. The effects of fading can be minimized by proper choices of carrier frequency and modulation, the use of several different carrier frequencies (frequency diversity), and the use of several different antennas (space diversity) or highly selective antenna arrays.

The direction of the sky wave above a frequency of about 30 MHz is not altered enough to return to earth and propagation above this frequency is predominantly straight-line from transmitter to receiver. Multipath propagation is still present as a result of reflected waves from the earth and man-made structures but is not as frequency dependent (e.g., the "ghosts" in television video). The distance of communication is governed predominantly by the height of the transmitting and receiving antennas. Some refraction or bending of the path does occur for low elevation angles, and communications beyond the line of sight are still possible at very high frequencies, although not as efficient as at the lower frequencies. Multipath propagation and fading may vary as a function of meteorological conditions in this mode of propagation.

For line-of-sight propagation between terrestial stations the signal path must be above the horizon for a clear path. Actually, electromagnetic waves do not propagate in exactly straight lines even above 30 MHz because the refractive index of the atmosphere decreases with altitude, causing some known bending of the wave fronts. This effect can be compensated in line-of-sight path calculations by using an effective earth radius that is four-thirds that of the actual radius. The point of tangency to this four-thirds earth model, as illustrated in Fig. 5.46, is called the radio horizon. Using this geometry, we can write

$$d^2 + r^2 = (r + h)^2,$$

$$d = \sqrt{2rh + h^2}.$$

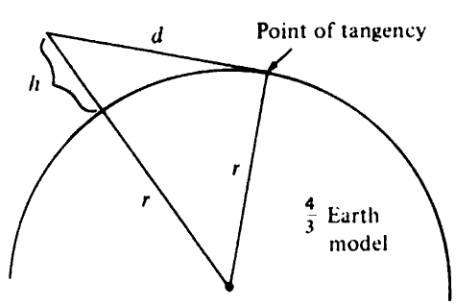


Figure 5.46 Line-of-sight distance to the radio horizon.

Assuming that $h^2 \ll 2rh$, and using 3960 miles for the radius of the earth, $r = (4/3)(3960) = 5280$, and a convenient approximation for the distance d in miles to the radio horizon is

$$d \approx \sqrt{2h}\,, \tag{5.93}$$

where h is the antenna height in feet. Thus a television station transmitting from a 1000-ft tower has an effective listening area out to about $\sqrt{2000} = 44.7$ miles, and this distance can be increased by additional tower height. It is for this reason that television (and FM) stations prefer to position antennas on tall buildings and towers to extend their effective listening area.

At vhf and higher frequencies, multipath propagation continues to cause problems. Because a major portion of the long-distance terrestial communication systems operate at these frequencies, multipath propagation at these frequencies has been studied extensively. In addition to the reflection and refraction of signals off the ground and man-made objects, and the gradual decrease of the refractive index with altitude, there may also be a layering effect of the lower atmosphere that causes sharp refractive index gradients. The result is that the transmitted signal is received via several paths with different characteristics, and these characteristics can change with time.

Recent work on multipath propagation at microwave frequencies has centered on fitting measured statistical data to mathematical models. One approach is to fit data to a two-path transfer function of the form:

$$H(\omega) = a[1 - be^{-j(\omega-\omega_0)\tau}]\,, \tag{5.94}$$

where a, b are scale and shape factors, respectively, ω_0 is the frequency of a minimum in $|H(\omega)|$, and τ is the channel difference between the two paths. (In a similar manner, one could postulate a three-path model, etc.) It has been found that this two-path model, when fitted against statistical propagation data, yields a usable channel transfer function for a relatively constant value of τ. Thus there is some advantage in building a filter that will compensate for this multipath. Such compensation, purposely introduced in the system, is called *equalization*.

From an overall perspective, the transfer function in Eq. (5.94) can be considered as the product of a frequency-selective term (the bracketed term in the equation) and a "flat-fade" term specified by the parameter a. Performance of angle modulation systems at vhf and higher frequencies can be characterized fairly well by a flat-fade margin. For example, the designer of an aircraft-to-aircraft or aircraft-to-ground FM system would specify an extra multipath margin (e.g., 20 dB) in the systems design to account for multipath propagation.

Digital transmissions (as well as AM) are more sensitive to frequency-selective fading.[†] Frequency-selective fading causes severe amplitude and delay distortion which degrades the reliability of such systems beyond that expected from a flat-fade margin alone.

[†] Digital modulation methods are covered in Chapters 7–10. For information on effects of multipath on digital communications, see, e.g., C. A. Siller, Jr., "Multipath Propagation," *IEEE Communications Magazine*, vol. 22 (February 1984): 6–15.

While both frequency diversity and space diversity are used in digital as well as analog transmission systems, the effects of multipath propagation are generally severe enough in digital systems to warrant additional countermeasures. These include amplitude equalization and channel equalization. Amplitude equalization methods attempt to restore the correct magnitude transfer function by adjustment of the parameters of the model (e.g., the two-path model). It has been found that amplitude equalization alone is of limited value for digital systems, but amplitude equalization coupled with some space diversity can be quite effective. However, multipath propagation corrupts both amplitude and phase of the transmitted signal and therefore a better solution is to equalize a channel for both amplitude and delay distortion. This leads to the design of adaptive transversal equalizers, a topic reserved for consideration in Chapter 9.

5.9 COMPARISON OF VARIOUS AM SYSTEMS

Various amplitude-modulated systems have been discussed, and it is interesting to compare their performance characteristics. First we compare large-carrier and suppressed-carrier systems.

Large-carrier systems have the advantage of simpler detectors at the receiver and thus are a natural choice where the emphasis is on inexpensive receivers. It also turns out that high-power modulators are easier to build for large-carrier systems because the carrier terms arising in generation do not have to be balanced or filtered out.

In contrast, suppressed-carrier systems require much less power to transmit the same information. Receivers for suppressed-carrier transmissions are more complicated because they must generate a carrier of the correct frequency and phase. Where communication efficiency is a prime consideration, however, the added receiver complexity may be justified. Suppressed-carrier systems are less susceptible to the effects of selective fading because a fade of the carrier can result in severe distortion in an envelope detector.

In comparing double-sideband and single-sideband systems, we note that the required receiver operations are identical. Double-sideband systems have an advantage in ease of generation, not requiring sideband suppressions. Double-sideband systems can be used to transmit information frequencies down to zero frequency with good fidelity.

In contrast, single-sideband systems require half the bandwidth of double-sideband systems, thus affording good spectrum efficiency. Because selective fading may disturb the phase relationships between sidebands, single-sideband systems are affected less by selective fading. Some of the advantages of both single-sideband and double-sideband can be realized using vestigial sideband systems.

For SSB-SC and DSB-SC signals, all of the transmitted power is in the sidebands. For equal-input sideband power, the output S/N ratio is the same for both on an average power basis. In addition, of course, the SSB signal occupies only half the

bandwidth. For sinusoidal modulation, DSB-LC results in at most only one-third the output S/N ratio of the corresponding DSB-SC signal [e.g., compare Eqs. (5.73) and (5.88)]. For typical voice modulation, the difference is greater because the peak-to-rms ratio is about 1/8. Taking the peak to be equal to 1 (for 100% modulation), the S/N ratio of DSB-LC for voice is more like about 1/65 that for a comparable DSB-SC signal on an average-power basis.

Another comparison can be made on a peak-power basis rather than on an average-power basis. The ratio of the peak amplitude in a DSB-LC signal compared to the corresponding DSB-SC signal, for 100% modulation on the peaks, is 2 [e.g., see Eq. (5.18)]. Thus for the same peak power limitation the DSB-SC signal has four times more average sideband power than the corresponding DSB-LC signal and therefore a 6-dB S/N advantage. Comparisons between SSB-SC and DSB-SC on a peak-power basis for general signals is more difficult. For sinusoidal modulation, the SSB-SC signal has an average power twice that of the DSB-SC signal for equal peak amplitudes. For sinusoidal modulation, then, the SSB-SC signal has a 3-dB S/N advantage over DSB-SC for the same peak power. However, this possible advantage varies with the type of modulating signal used.

In conclusion, we see that each type of amplitude modulation has its advantages and disadvantages. Each circumstance must be evaluated on its own merits and the choice of modulation tailored to that situation. We will continue this discussion when S/N ratios of angle-modulation systems are discussed in Chapter 6.

5.10 SUMMARY

Multiplication of a waveform with a sinusoid translates its spectral density in frequency. This effect, known as the modulation property, can be used to generate amplitude-modulated waveforms and to demodulate them again. Because such waveforms have no separate carrier frequency and are symmetic about a center frequency, they are called double-sideband suppressed-carrier (DSB-SC) waveforms. Proper demodulation of DSB-SC waveforms demands an exact knowledge of the frequency and phase used in modulation.

The addition of a carrier frequency is used to aid in waveform demodulation. If the added carrier amplitude is large, an envelope detector can correctly demodulate the waveform. This principle is used in commercial AM broadcasting.

Choice of different carrier frequencies allows several message spectra to be placed side by side in frequency and transmitted simultaneously. This is called frequency-division multiplexing (FDM).

To conserve bandwidth, one sideband of a DSB waveform can be suppressed; this type of modulation is called single-sideband (SSB). This demands exacting filter design or phase balance. An economical way to gain some of the advantage of SSB is to partially suppress one sideband. This is known as vestigial-sideband modulation.

Bandpass random noise may be described in terms of random amplitudes and phases of sines and cosines at the center frequency. In random noise, the average noise power is divided equally between the two components.

For a given input sideband power and noise spectral density, the output S/N ratios for amplitude modulation are all the same as long as the S/N ratios are large. Synchronous detection may be used for low S/N ratios whereas envelope detection becomes unusable for low S/N.

While large-carrier systems allow the use of inexpensive envelope detectors, the performance of these systems is not optimum in terms of total power or the signal-to-noise ratios required for good communication. The use of synchronous demodulation is preferable in low signal-to-noise conditions.

The phases of signals arriving by various paths at the receiver cause the signal strength to vary. Cancellation of these signals results in loss of signal strength or fading. Fading that is sensitive to frequency is called selective fading. The effects of selective fading are least detrimental to amplitude-modulated systems when the bandwidth of the signal is narrow and the maintenance of a large carrier is not necessary, making SSB-SC an obvious choice.

Selected References for Further Reading

1. R. L. Shrader. *Electronic Communication*. Fourth ed. New York: McGraw-Hill, 1980.
   A fairly detailed discussion of the circuits and operating characteristics of practical communication systems; easy reading. Chapter 20 has a brief discussion of antennas and propagation.
2. B. P. Lathi. *Communication Systems*. New York: John Wiley & Sons, 1968.
   Chapters 3 and 7 are a parallel treatment of amplitude modulation, stressing the use of general signal notation and convolution.
3. A. B. Carlson. *Communication Systems*. Third ed. New York: McGraw-Hill, 1986.
   Chapter 6 uses phasor representations to analyze amplitude modulation, particularly envelope detection.
4. H. Taub and D. L. Schilling. *Principles of Communication Systems*. Second ed. New York: McGraw-Hill, 1986.
   This book has a good discussion of signal-to-noise calculations in Chapters 7 and 8.
5. K. K. Clarke and D. T. Hess. *Communication Circuits: Analysis and Design*. Reading, Mass: Addison-Wesley, 1971.
   Chapters 7–10 present an advanced treatment of the circuits and theory used for amplitude modulation and demodulation.
6. D. R. Smith. *Digital Transmission Systems*. New York: Van Nostrand Reinhold, 1985.
   A major portion of Chapter 9 deals with propagation and multipath fading.
7. M. Schwartz, W. T. Bennett, and S. Stein. *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.
   Chapter 9 includes an extensive treatment of fading.
8. W. D. Gregg. *Analog and Digital Communication*. New York: John Wiley & Sons, 1977.
   Chapters 2, 4, 7 present an advanced analysis of AM generation and detection with examples of current practice.

PROBLEMS

5.1.1 In some systems it is possible to multiply a signal by the complex-valued function $e^{j\omega_1 t}$ after the DSB-SC modulator, as shown in Fig. P–5.1.1. [For example, this effect can be accomplished at microwave frequencies by a dielectric cylinder with a spiral conductor rotating inside a cylindrical waveguide, and rotated at ω_1 rad/sec. As a second example, it could be caused by a Doppler shift introduced as a result of relative motion.]

a) Examine the results of this second multiplication on the spectral density of the modulator output, $g(t)$.

b) Suppose $g(t)$ is demodulated using a synchronous detector with a frequency error $\Delta\omega$. Compare the effects of ω_1 and $\Delta\omega$ on (1) the magnitude and (2) the phase of the detector output.

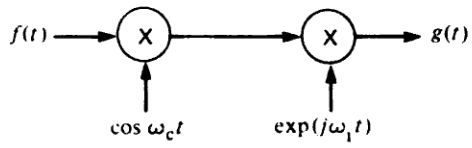


Figure P–5.1.1.

5.1.2 A modulating signal $f(t)$ [with Fourier transform $F(\omega)$] is applied to a double-sideband suppressed-carrier modulator operating at a carrier frequency of 200 Hz. Sketch the spectral density of the resulting DSB-SC waveform, identifying the upper and lower sidebands, for each of the following cases.

a) $f(t) = \cos 100\pi t$;

b) $F(\omega) = \begin{cases} [1 + \cos(\omega/200)]/2 & |\omega| < 200\pi, \\ 0 & \text{elsewhere.} \end{cases}$

5.1.3 We wish to examine the frequency and phase of one low-frequency oscillator relative to a second. We decide to multiply the outputs of the oscillators to make the comparison. However, as a result of difficulties in building a low-frequency multiplier we instead choose to use DSB-SC modulators and a bandpass multiplier, as shown in Fig. P–5.1.3.

a) Determine expressions for the DSB-SC signals.

b) Determine an expression for $g(t)$.

c) Sketch a magnitude spectrum of $g(t)$.

d) Under what conditions does $g(t)$ represent the desired product of the two input waveforms?

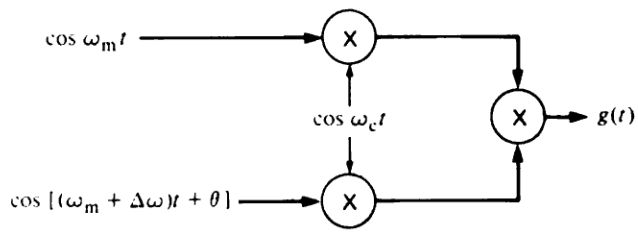


Figure P–5.1.3.

5.1.4 A sinusoidal signal $f(t) = \cos 2000\pi t$ is multiplied by a (periodic) symmetric triangular waveform (cf. Table 2.2) with unit peak amplitude and $T = 100\ \mu$sec. The output of the multiplier is applied to a LPF with unity gain within the pass band.

a) Determine the minimum and maximum bandwidth of the LPF if the output is to be a DSB-SC waveform corresponding to $f(t)$.
b) Determine an expression for the output of the LPF under the above conditions.
c) Can this system be represented by the DSB-SC modulator shown in Fig. 5.1(a) under the above conditions? If your answer is yes, then determine the two inputs to the modulator.

5.1.5 a) The input signals $x(t) = A \cos \omega_1 t + B \cos \omega_2 t$ and $y(t) = C \cos \omega_3 t$, $\omega_3 \gg \omega_1$, ω_2, are applied to the two inputs of a signal multiplier. Sketch a spectral density of the output, and show that linearity in $x(t)$ has been maintained.
b) The input signal $x(t)$ above is applied to both inputs of the multiplier. Sketch the spectral density of the output, and show that linearity in $x(t)$ has not been maintained under these conditions.

5.1.6 When the input to a given audio amplifier is $(4 \cos 800\pi t + \cos 2000\pi t)$ mV, the measured frequency component at 1 kHz in the output is 1 V and the ratio of the frequency component at 600 Hz to that at 1 kHz is 0.002. Represent the amplifier output-input characteristic by

$$e_o(t) = a_1 e_i(t) + a_2[e_i(t)]^2.$$

a) Evaluate the numerical values of a_1, a_2, from the test data given. (This type of test is called an intermodulation distortion test.)
b) What would you expect to be the magnitudes of the frequency component at 800 Hz and at 1400 Hz?

5.1.7 Define an error in Drill Problem 5.1.1 as the difference between $e_1(t)$ with and without a phase error.

a) Derive an expression for the autocorrelation function of the error in terms of θ_0 and the autocorrelation functions of $f_1(t)$ and $f_2(t)$.
b) Find an expression for the rms error if the mean-square values of $f_1(t)$, $f_2(t)$ are equal, and $f_1(t)$, $f_2(t)$ are uncorrelated. Plot your result over $0 \le \theta_0 < \pi/6$.

5.1.8 Consider the system shown in Fig. P–5.1.8. Assume that the comparator outputs are ± 1, dependent on the polarity of each input, and the EXCLUSIVE-NOR gate is an equivalence operation, yielding $+1$ if inputs have the same polarity and -1 if inputs have disimilar polarity.

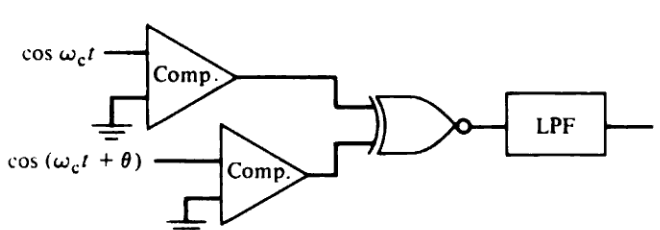


Figure P–5.1.8.

a) Show that the output voltage contains a term proportional to cos θ (and thus that this system could be used as a phase detector).
b) What is the maximum bandwidth of the LPF (assume ideal)?
c) Repeat part (b) if one of the comparator outputs is not quite symmetrical.
d) Repeat part (b) if both of the comparator outputs are not quite symmetrical.

5.1.9 Three possible sources of error in the system shown in Fig. 5.8 are (1) the transformer is not accurately center-tapped; (2) the linear gain coefficients (a_1) are not matched in both devices; (3) the quadratic coefficients (a_2) are not matched. Investigate to find which error is the most critical in obtaining a good DSB-SC waveform at the output of the BPF. State any assumptions that you make in your analysis.

5.1.10 Two measurement systems are shown in Fig. P–5.1.10. Find the output of each system if $e_1(t) = \cos \omega_1 t$. What type of modulation is present in each and how should it be detected?

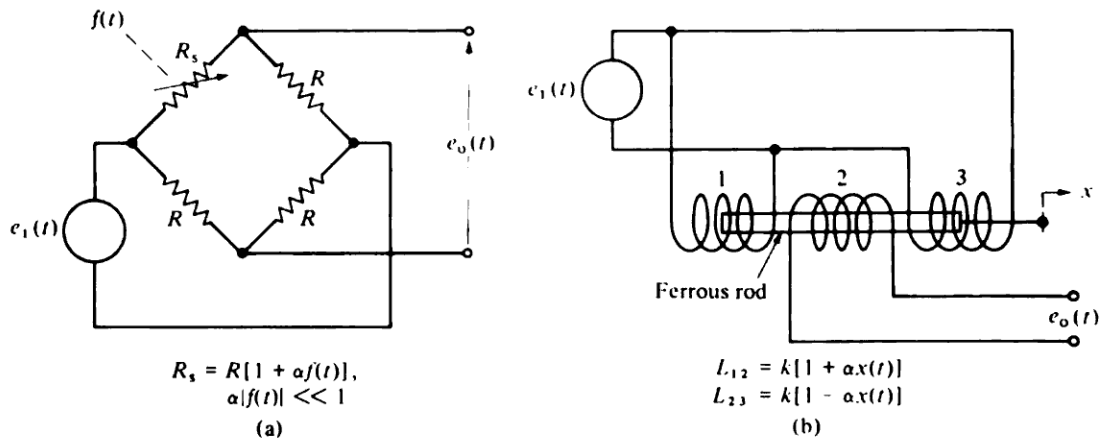


Figure P–5.1.10 (a) Bridge circuit; (b) linear motion transducer.

5.1.11 Show that the chopper system shown in Fig. P–5.1.11 will produce essentially the correct signals for FM stereo transmission.

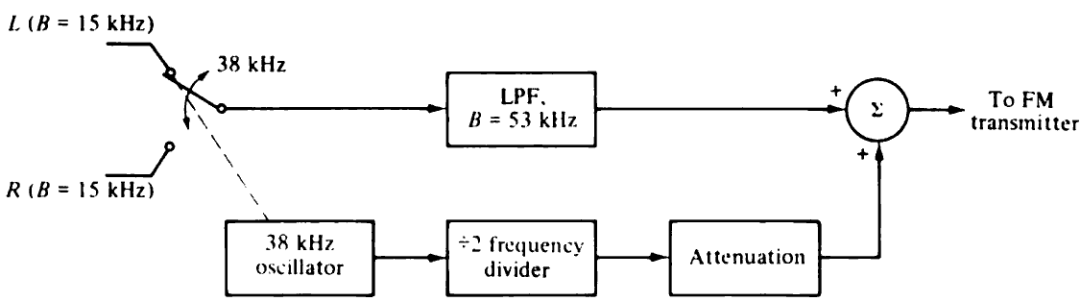


Figure P–5.1.11.

5.1.12 A proposed frequency synthesizer to control the output frequency of a stable oscillator digitally is shown in Fig. P–5.1.12. The stable oscillator operates at f_c = 100 kHz and the integer values of N_1, N_2 are selectable between 1 and 10.

a) What is the output frequency when $N_1 = 4$, $N_2 = 2$?
b) Determine the frequency range and minimum frequency increment of the synthesizer.

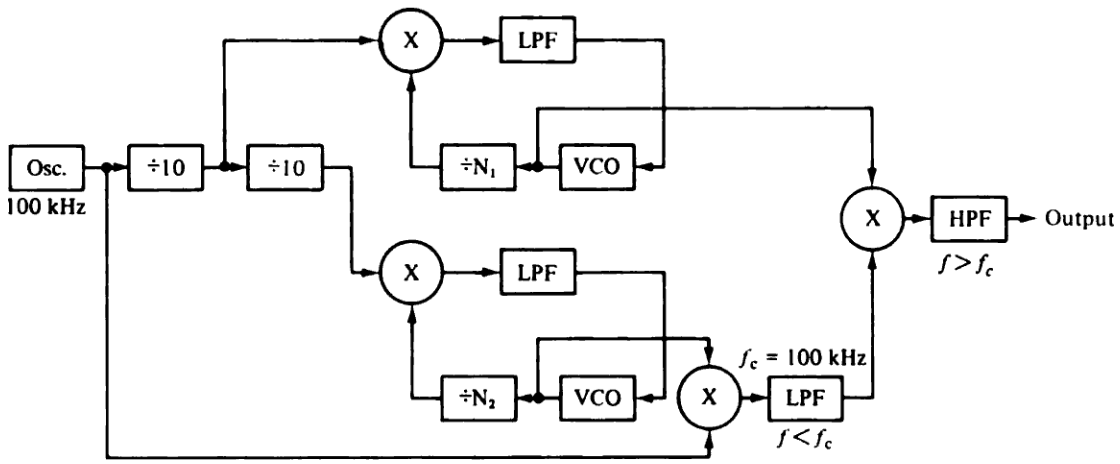


Figure P–5.1.12.

5.1.13 Repeat Problem 5.1.12, with the exception that the output in Fig. P–5.1.12 is applied as the input to the mixer in Fig. P–5.1.13.

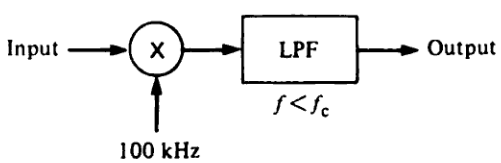


Figure P–5.1.13.

√ **5.1.14** Run a 128-point FFT of the discrete data

$$f(k) = \begin{cases} 10 & k = 62, 66 \\ 0 & \text{elsewhere.} \end{cases}$$

Plot and explain your answer.

5.2.1 Two AM systems are compared. In System 1 the signal $f_1(t) = \cos \omega_m t$ is applied to the input of a DSB-SC modulator and then a carrier term, $A \cos \omega_c t$, is added after the modulation. In System 2 the signal $f_2(t) = a + f_1(t)$ is applied to the input of the DSB-SC modulator.

a) Compare the results. Can a dc level in the input be distinguished from carrier in the AM waveform?
b) A difficulty in System 1 is that there may be some phase shift θ between the added carrier $A \cos \omega_c t$ and the carrier frequency of the DSB-SC modulator. This phase shift limits the modulation index that can be achieved with this method. Determine the highest modulation index that can be obtained for a given value of θ.

5.2.2 Identify the type of amplitude modulation present at the output of the modulator shown in Fig. P–5.2.2 (also see Problem 5.1.8) for the conditions listed below; if it is DSB-LC, then compute the modulation index. Also state what effect (if any) the polarity of E has on the output waveform.

a) Digital circuitry operates on a 0, 1 binary system.
b) Digital circuitry operates on a -1, $+1$ binary system.

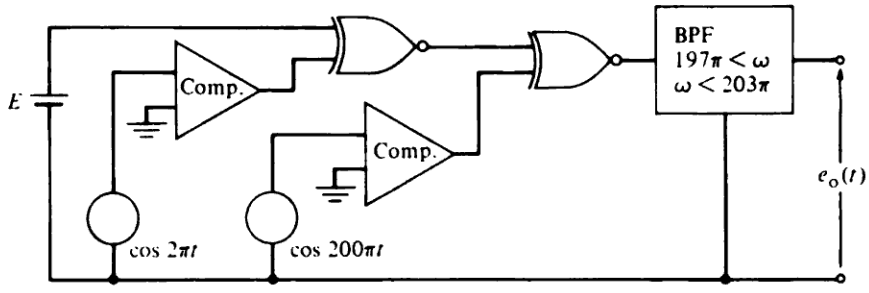


Figure P–5.2.2.

5.2.3 A type of power rating useful in transmitter specifications is that of peak envelope power (PEP). The PEP of a signal $\phi(t)$ is the average power of a sinusoid with the same peak amplitude as that of $\phi(t)$. Consider two amplitude-modulated signals $\phi_1(t)$, $\phi_2(t)$,

$$\phi_1(t) = A(1 + m \cos \omega_m t) \cos \omega_c t$$

$$\phi_2(t) = B \cos \omega_m t \cos \omega_c t$$

where $\omega_c \gg \omega_m$. Let the average power in $\phi_1(t)$, $\phi_2(t)$ be P_1, P_2, respectively.

a) Calculate and plot the ratio P_1/P_2 as a function of m if the PEP is equal in the two waveforms.
b) Sketch the power (line) spectrum of each signal for the above condition and $m = \frac{1}{2}$.

5.2.4 For the sinusoidally modulated DSB-LC waveform shown in Fig. P–5.2.4:

a) Find the modulation index.
b) Write an expression for the waveform in the form of Eq. (5.18).
c) Sketch a line spectrum of the waveform.
d) Show that the sum of the two sideband lines in part (c), divided by the carrier line, yields the modulation index.
e) Determine the amplitude and phase of the additional carrier which must be added to the waveform shown to attain a modulation index of 20%.
f) Repeat part (e) to attain a modulation index of 80%.

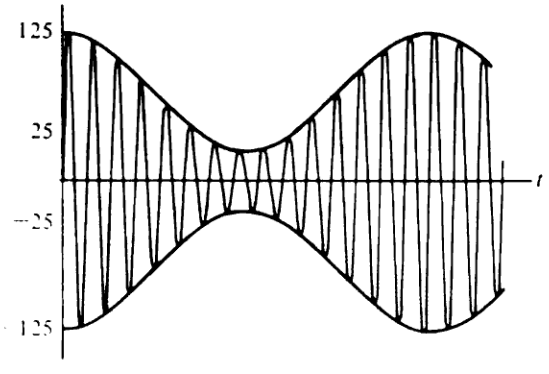


Figure P–5.2.4.

√ **5.2.5** One method of measuring the modulation index of a DSB-LC waveform is to apply the modulated waveform to the vertical deflection system of an oscilloscope and the modulating waveform to the horizontal system.

a) Simulate this using a digital computer and a graphics monitor. Let $x(t) = \cos \pi t$,

$$y(t) = (1 + m \cos \pi t)(\cos 100\pi t)/(1 + m)$$

be the horizontal and vertical functions, respectively. Compute and plot for $0 \leq t < 1$, $m = 0.5$, using 500 sample points. You should observe a trapezoidal-type pattern.

b) Repeat part (a) for $m = 1.0, 1.5$.

c) Define a parameter A as the maximum vertical width of the patterns observed, and define B as the minimum width. Derive a formula for $m(A, B)$.

d) Investigate what happens to the patterns for $m > 1$ if no phase reversal of carrier is permitted (as, e.g., in a nonlinear-type transmitter design).

5.2.6 A given AM (DSB-LC) transmitter develops an unmodulated power output of 1 kW† across a 50-ohm resistive load. When a sinusoidal test tone with a peak amplitude of 5.0 V is applied to the input of the modulator, it is found that the spectral line for each sideband in the magnitude spectrum for the output is 40% of the carrier line. Determine the following quantities in the output signal:

a) The modulation index.

b) The peak amplitude of the lower sideband.

c) The ratio of total sideband power to carrier power.

d) The total power output.

e) The total average power in the output if the peak amplitude of the modulation sinusoid is reduced to 4.0 V.

5.2.7 A given AM (DSB-LC) transmitter is tested using a resistive ("dummy") load. With no modulation, the average output power is 5 kW. With a 1-kHz sinusoidal input $f(t)$ with 1-V peak amplitude, the output power is 7 kW.

a) Assume that typical programming material has a peak-to-rms ratio of 8. If the peaks are set to 100% modulation, what is the resulting average transmitter power for programming?

b) Calculate the equivalent sinusoidal modulation index for part (a).

c) To increase effective power output for the given 5 kW rating (see footnote for Problem 5.2.6), it is decided that the broadcast material can be "colored." This results in some loss in fidelity for program material, but this may be tolerated in some cases to gain more effective radiated power. Although more elaborate methods could be chosen, we assume here that $f(t)$ is applied to a nonlinear amplifier with 120% gain on positive excursions relative to the negative excursions to achieve the "color." Using the equivalent sinusoidal modulation index in part (b), estimate how much added effective power output can be gained by use of this method.

√ **5.2.8** Use a digital computer and graphics monitor to compute and display the magnitude of the FFT using $N = 128$ over an interval $T = 1$ to simulate a spectrum analyzer. Display the magnitudes over $(0, N/2)$. This will result in a display with the integers as frequency (in Hz) for the horizontal axis, with a resolution of 1 Hz. Because only relative

† FCC power ratings for AM broadcast transmitters are for an average carrier (i.e., unmodulated) power.

magnitudes are important here, normalize the vertical display to the maximum (magnitude) value.

a) Compute and display the magnitude spectrum for

$$x(t) = [1 + m \cos 8\pi t] \cos 64\pi t$$

for several values of m, $0 \le m \le 1$. Record the sideband magnitudes (lines 28, 36 in your display) and carrier magnitudes (line 32) for each value of m.

b) Compute and display the magnitude spectrum for

$$y(t) = [1 + m \cos 8\pi t][\cos 64\pi t + \beta \sin 8\pi t]$$

for $\beta = 0.2$ and several values of m. (β represents the incidental frequency modulation present.) Observe that now the sideband magnitudes are unequal. Demonstrate that the correct AM modulation index m can be obtained by adding the two sideband magnitudes and dividing by the carrier magnitude, provided that $\beta \le m$. [*Note:* The incidental frequency modulation also generates some higher harmonic content in the spectrum.]

5.2.9 A sinusoidal carrier waveform with a peak amplitude of 80 V is added to a low-frequency sinusoid. This sum is applied to the input of an amplifier (class C) whose output-input gain characteristic is

$$e_o(t) = \begin{cases} 50[e_i(t) - 10] & \text{for } e_i(t) > 10 \text{ V}, \\ 0 & \text{for } e_i(t) < 10 \text{ V}. \end{cases}$$

The output of the amplifier is applied to a narrowband filter tuned to the carrier frequency (cf. Fig. 5.21).

a) Determine the maximum modulation index that this modulator can handle without significant envelope distortion.

b) Sketch the resulting modulated waveform if the peak amplitude of the low-frequency sinusoid is 80 V.

c) The situation in part (b) results in serious interference with other users operating in adjacent frequency allocations. Explain why this is a much more serious problem than the use of nonlinearities in the mixers used for shifting frequencies prior to transmission.

5.2.10 Let $f(t) = \cos \omega_m t$ in Eq. (5.28) and add the term $a_3 e^3(t)$ to Eq. (5.27); then revise Eq. (5.29) and derive an expression for the modulation index m.

5.2.11 A sinusoidally modulated DSB-LC waveform $\phi(t)$ [cf. Eq. (5.18)] is applied to a square-law device, such that the output voltage $e_o(t)$ is $e_o(t) = [f(t)]^2$. Show that the ratio of the second harmonic to the first harmonic in $e_o(t)$ is equal to $m/4$.

5.2.12 For each of the following conditions find the largest value of RC that will allow an envelope detector to follow the envelope of the sinusoidally modulated waveform of Eq. (5.18).

a) Assume that (ω_m/ω_c) is very small, and therefore any effects of ω_c can be neglected.

b) Assume that (ω_m/ω_c) is small.

[*Hint:* Approximate each point at which the capacitor charges by an impulse function.]

√ **5.2.13** Use a digital computer and graphics monitor to compute and display the magnitude of the FFT using $N = 128$ over an interval $T = 1$ to simulate a spectrum analyzer. (See

Problem 5.2.8.) (For higher resolution, one can double the number of points N and the interval T, etc., and scale the results—but at the expense of added computation time.) Because only relative magnitudes are important here, normalize the vertical display to the maximum value of the frequency components.

a) Use this spectrum analyzer to display the magnitude spectrum of the input to the BPF in Fig. 5.21(b); let the input to the (diode) nonlinearity be

$$x(t) = 2 \cos 32\pi t + a \cos 4\pi t$$

and let the output-input characteristic of the diode-resistor combination be

$$y(t) = x(t)u[x(t)].$$

Begin by letting $a = 1$.

b) Identify the DSB-LC terms in the display, and calculate the modulation index m by taking the ratio of the sum of the two sidebands to the carrier.

c) Vary the value of $a \leq 1$ to determine a straight-line equation for $m(a)$.

d) Monitor the third-harmonic components away from carrier in the DSB-LC spectrum and increase $a > 1$. Note the point beyond which the third-harmonic content begins to increase rapidly with further increases in a. (There will be some residual level as a result of the computational methods.) Determine the largest modulation index m for which the third-harmonic components are $\leq 0.1\%$ of the carrier component.

√ **5.2.14** Suppose that a given diode can be approximated by a one-sided square-law behavior for low voltages and a linear behavior for larger voltages. Let the voltage

$$e(t) = A(1 + m \cos \omega_m t) \cos \omega_c t$$

be applied to the diode, and let $m = 0.40$. The current-voltage characteristic of the diode is (in mA):

$$i(t) = \begin{cases} 0 & e(t) < 0 \\ 5e^2(t) & 0 < e(t) \leq 1 \text{ V} \\ 10[e(t) - 0.5] & 1 < e(t) \end{cases}$$

a) Using a digital computer, compute and plot the ratio of the magnitudes of the second harmonic to the first harmonic of the modulating waveform present in $i(t)$ for $0 < A \leq 2$, in 0.1 step increments. [It may be helpful to refer to the result of Problem 5.2.11.]

b) Determine the largest value of A for which the above ratio is equal to $m/4$. Explain your answer.

c) Determine the smallest value of A for which the above ratio is $\leq 1\%$. Explain your answer.

5.3.1 Sketch a typical spectrum of a frequency-division multiplexing system for each of the following combinations, and list some advantages and disadvantages of each system.

a) DSB-LC for both subcarrier and the main carrier modulation.

b) DSB-SC for subcarrier modulation, DSB-LC for main carrier modulation.

c) DSB-SC for both subcarrier and the main carrier modulation.

5.3.2 A television receiver using the superheterodyne principle has a video IF carrier frequency of 45.75 MHz (and a sound carrier at 41.25 MHz). With the aid of the frequency listings in Appendix C, make a table of those television channels, by channel numbers, that could be image frequencies for desired Channels 2–15. Also, list the beat (i.e., the difference) frequencies between the video carrier frequencies for your entries.

5.3.3 Commercial television transmissions use two carrier frequencies, one for video and one for audio. A listing of channel assignments in the United States is given in Appendix C. Make a sketch of the spectral densities corresponding to those illustrated in Fig. 5.26 for a superheterodyne television receiver tuned to Channel 6, and its image. The receiver LO operates above station carrier and the video IF is 45.75 MHz. Show both the video carrier and the audio carrier; draw the video carrier arrows taller than the audio carriers to distinguish them (stations transmit more power for the video carrier than for the audio). Label all important points on the frequency axes.

5.4.1 The signal $f(t) = \cos 2000\pi t$ is used to modulate a 5-kHz carrier. Sketch the time waveforms and line spectra for the modulation methods listed below. In each case, normalize the waveform so that the peak envelope power (PEP) is 1 W across 1 Ω. The PEP of a modulated waveform is the average power that would be developed if it were replaced by a sinusoid with an equal peak amplitude.

a) DSB-SC
b) DSB-LC, $m = 1$
c) SSB-SC$_-$
d) SSB-SC$_+$

5.4.2 A SSB-LC$_+$ transmitter is modulated with the input $f(t) = \cos 2000\pi t$. In complex-valued notation, the output waveform can be written as

$$\phi_{\text{SSB-LC}_+}(t) = \mathcal{R}e\{A[1 + mf(t) + jm\hat{f}(t)]e^{j\omega_c t}\}.$$

a) Evaluate $\hat{f}(t)$.
b) What is the maximum value of m if envelope detection is used?
c) Determine the numerical values of A, m if the unmodulated power across 50 ohms is 500 W, and the average power is found to increase 50% when the modulation is applied.
d) Find the peak envelope power (PEP; see Problem 5.4.1) under both the unmodulated and the modulated condition in part (c).

5.4.3 A potential problem in the design of superheterodyne-type television receivers is a possibility that if the LO generates second harmonics then these harmonics can produce a difference frequency to bring picture content of an unwanted channel into the IF bandwidth. For a video IF carrier frequency of 45.75 MHz in the receiver, determine the interfering channel (use frequency listings in Appendix C) for this type of interference when the receiver is tuned to

a) Channel 5.
b) Channel 6.

5.4.4 Assume that the first IF amplifier in the dual-conversion receiver shown in Fig. 5.27 passes all frequencies between $(\omega_1 - W)$ and ω_1. Determine the receiver oscillator frequencies, ω_0 and ω_d, required to properly demodulate the following transmissions.

a) A SSB$_+$ transmission occupying the frequency range from ω_c to $(\omega_c + W)$.
b) A SSB$_-$ transmission occupying the frequency range from $(\omega_c - W)$ to ω_c.
c) What happens if the receiver is set for receiving upper sideband and the actual received signal is lower sideband?

5.4.5 A certain (simplified) speech scrambler to aid in communications privacy is shown in Fig. P–5.4.5. The speech bandwidth is divided into three bands, $(0–f_1)$, $(f_1–f_2)$, $(f_2–f_3)$, by three bandpass filters BPF_1, BPF_2, BPF_3, respectively. Polarities of the inverting/noninverting unit-gain amplifiers are preset.

a) Analyze system operation by sketching spectra at (1) the outputs of the first bank of bandpass filters, (2) the inputs of the second bank of bandpass filters, and (3) the output. Assume a triangular (low-pass) input spectral density, decreasing linearly toward zero at $f = f_3$. Let all amplifiers be set to the inverting position.

b) Investigate what effect amplifier polarity may have on system operation.

c) Can the same system be used for unscrambling the signal? Explain.

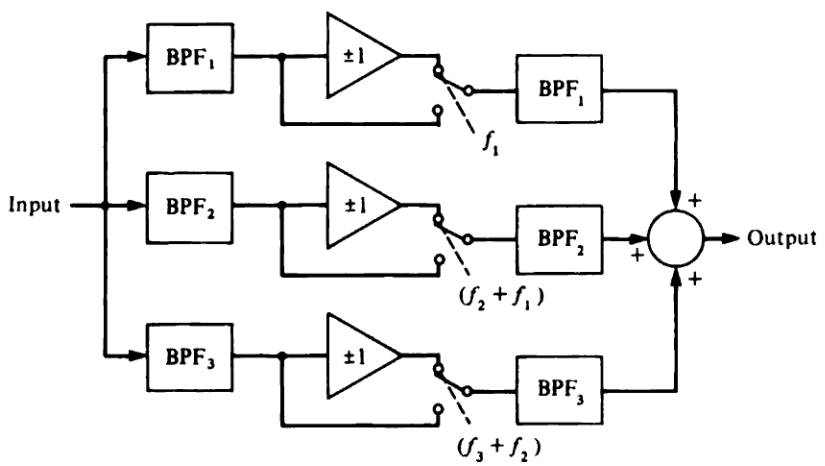


Figure P–5.4.5.

★ **5.4.6** In the SSB-SC modulation system shown in Fig. 5.31, the signal $f(t)$ is shifted in phase ideally by exactly 90° at all frequencies. This is difficult to accomplish in practice. Show that if it is shifted by $(90 + \alpha)°$ then the ratio of the undesired output sideband magnitude to the desired sideband magnitude is given by $\tan(\alpha/2)$.

★ **5.4.7** a) Find the Hilbert transform of $f(t) = a/(a^2 + t^2)$.

b) Determine the magnitude of $z(t) = f(t) + j\hat{f}(t)$.

★ **5.4.8** a) Find the Hilbert transform of $f(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t}$.

b) Determine the magnitude of $z(t) = f(t) + j\hat{f}(t)$.

c) What happens in part (b) if the F_n are even and real-valued?

d) Check your results above using $f(t) = \cos \omega_0 t$.

√★ **5.4.9** Use a 32-point FFT to compute and display the Hilbert transform, $\hat{f}(t)$, for each $f(t)$ given below over (0, 1). Normalize your results so that the magnitudes of $f(t)$ and $\hat{f}(t)$ are equal in part (a). Also, compute and display the magnitude of $z(t) = f(t) + j\hat{f}(t)$.

a) $f(t) = \cos 4\pi t$. [*Note:* Correct signs in the exponential are essential for a correct phasing of the results.]

b) $f(t)$ is one cycle of a periodic symmetric triangular waveform with unit peak amplitude.

5.5.1 A vestigial sideband signal is generated from an input $f(t)$ by first generating a DSB-LC signal ($m = 0.8$) with a carrier frequency $f_c = 10$ kHz and then passing this signal through a filter whose magnitude frequency transfer function is

$$|H(\omega)| = \begin{cases} 1 + \cos[(\omega - \omega_0)/8000] & |\omega - \omega_0| < 8000\pi \\ 0 & \text{elsewhere} \end{cases},$$

where $\omega_0 = 24{,}000\pi$. Find an expression for the resulting VSB signal, sketch the spectral density, and calculate the ratio of the peak envelope power (cf. Problem 5.2.3) to average power if

a) $f(t) = \cos 1000\pi t$.
b) $f(t) = \cos 2000\pi t$.
c) $f(t) = \cos 4000\pi t$.

5.5.2 Calculate the envelope of the resulting VSB signal for each case in Problem 5.5.1.

5.5.3 As illustrated in Fig. 5.38, the receiver tuning characteristic plays an important role in the overall VSB system for commercial television systems.

a) Explain what effect(s) you would expect to observe in the video display if the receiver tuning were shifted slightly too high in frequency.
b) Repeat part (a) if the receiver tuning were slightly too low in frequency.

5.6.1 A noise waveform with a power spectral density $S_n(\omega) = 10^{-7}e^{-|\omega|/\alpha}$ W/Hz, where $\alpha = 2\pi \times 10^7$/sec, is passed through an ideal BPF with unity gain, unity resistance levels, and a bandwidth of 200 kHz centered at 10 MHz.

a) Determine the mean-square values of the in-phase and quadrature components of the bandpass time representation given in Eq. (5.59).
b) Repeat part (a) assuming that $S_n(\omega)$ is constant (at the center frequency value) across the bandwidth of the filter.
c) Repeat parts (a) and (b) for a bandwidth of 2 MHz.

5.6.2 Suppose that the galactic noise plots can be approximated, over the range 0.01 GHz $< f <$ 1 GHz, by a straight line in Fig. 4.16 through the points ($f = 0.01$ GHz, $T = 10^6$ K), ($f = 1$ GHz, $T = 20$ K). Determine an expression for the rms noise as a function of frequency over this range and for the conditions of Problem 5.6.1.

5.7.1 A DSB-SC and a SSB-SC transmission are each sent at 1 MHz in the presence of additive noise. The modulating signal in each case is band-limited to 3 kHz. The received signal power in each case is 0.2 mW. The received noise is assumed to be white with a (two-sided) power spectral density of 10^{-3} μW/Hz. The receiver consists of a bandpass filter whose bandwidth matches the bandwidth of each transmission, followed by a synchronous detector.

a) Compare the signal-to-noise ratios at the detector inputs.
b) Compare the signal-to-noise ratios at the detector outputs.
c) Repeat part (a) if the (two-sided) power spectral density were $10^3/|f|$ μW/Hz. Would a "white noise" assumption be valid here?

5.7.2 Show that, strictly on an average power basis, the use of an envelope detector in a DSB-LC system results in a linear $(S/N)_o$ versus $(S/N)_i$ relationship. What makes this relationship nonlinear for commercial AM systems?

5.7.3 Compare the output S/N ratio, on a scale relative to that for DSB-SC, for signals received using DSB-SC, SSB-SC, DSB-LC with $m = 0.1$, DSB-LC with $m = 0.5$, for each of the conditions listed below. Assume single-tone sinusoidal modulation and synchronous detection in all cases.

a) Equal average power.
b) Equal peak signal strength.

5.7.4 The DSB-LC signal

$$\phi(t) = 3 \cos (10{,}000\pi t) + \cos (1000\pi t) \cos (10{,}000\pi t) \text{ V}$$

is present with additive white noise whose (two-sided) power spectral density is 1 μW/Hz. This signal-plus-noise is passed through an ideal low-pass filter with a bandwidth of 10 kHz. Assume all resistance levels are 1 ohm.

a) Compute the average S/N ratio at the output of the low-pass filter.
b) A synchronous detector is used to demodulate the above signal. Compute the average S/N ratio if the output of the detector is filtered to $0 < f < 1$ kHz.
c) An ideal square-law detector (i.e., the signal-plus-noise is multiplied by itself) is used to demodulate the above signal. Compute the average S/N ratio if the output of the detector is filtered to $0 < f < 1$ kHz. [*Hint:* Use frequency convolution, and consider signal-signal, signal-noise, and noise-noise products.]

5.8.1 When an electromagnetic wave arrives via two differing paths, the sum may add or subtract, depending upon the phase difference between them. Designate the relative attenuation of one wave with respect to the second by α, and the difference in phase by θ.

a) Develop an expression for the magnitude of the resultant.
b) Plot the maximum and the minimum of the above expression, in dB, as a function of α, $0 \leq \alpha < 1$.

5.8.2 Assume that the distance from a transmitter to a receiver is 200 km, and a strong reflector exists at midpath and at a height of 2 km (e.g., an airplane). Calculate how far one needs to move the receiving antenna (along a line joining the transmitter and receiver) to change the phase of the received signal by 180°. Express your answer in terms of the carrier wavelength λ.

a) Assume a flat earth and direct line-of-sight propagation.
b) Repeat part (a) using a 4/3 spherical earth model.

★ **5.8.3** A certain SSB-SC transmitter operates at 10 MHz with an effective output power of 100 W. Bandwidth of transmission is 3 kHz. Assume that transmission is primarily by sky wave, and that the path loss formula given in Section 4.7.7 is accurate for purposes of night operation (i.e., strong reflections off the ionosphere). (Neglect possible transmitting and receiving antenna gain factors.)

a) Estimate the maximum range, in miles, for night-time reception under the following conditions:
   (i) Receiver noise temperature is 1000 K.
   (ii) Antenna noise temperature is given by Fig. 4.16.
   (iii) A minimum (output) S/N of 10 dB is required.
b) What is the minimum difference in path length to cause a selective fading null within the passband?

5.9.1 Amplitude modulation is to be considered for each of the following objectives. State which type of amplitude modulation you would recommend and the reason(s) for your choice. Also give your recommendation for the type of detector to be used.

a) A voice intercom using the 60-Hz power lines to conduct the modulated signal.
b) Remote control (from the cockpit) of the landing flaps using the 400 Hz ac power supply in an aircraft.
c) Weather broadcasts to marine users.
d) Biweekly voice contact between two hospitals in the U.S. and Peru, South America.
e) Transmission of weather photofax information (bandwidth of 3200 Hz, centered at 2400 Hz) over a standard voice-grade telephone circuit (usable bandwidth of 300–3000 Hz).
f) Injection of a probing signal into a heated plasma, and measurement of the amplitude and phase of the response.

DESIGN PROBLEMS

Design Problem 5.1 Voice Intercom System Design

Design a communication system using amplitude modulation for a voice intercom in a large building complex. This intercom is to make use of the 60 Hz power lines for the transmission medium. Assume that there are no power transformers within the building complex itself, but that power for the entire building complex is supplied by a single outside transformer. This transformer is not shared by other users in the area. The frequency of transmission chosen for the intercom must be high enough that the power transformer outside the building will effectively ground the transmission, thus avoiding reception by other residents in the area. However, the frequency should not be so high that the transmission will become very weak in some portions of the building complex. Transmitted power levels are constrained by these considerations also, as well as by the fact that the impedance of a power distribution line is very low. Assume that the power spectral density of the noise environment behaves as $1/|f|$ with frequency (f), and assume a receiver noise temperature of 1000 K. Design a prototype system stating what type of modulation and detection you choose, and your choice of carrier frequency and bandwidth. Also estimate how much power you believe you need to transmit. State your assumptions.

Design Problem 5.2 Satellite Data Link Design

Design a communication system using amplitude modulation for a 1 kHz data channel to a geostationary satellite. Assume that the satellite is in the equatorial plane at an altitude of 22,300 miles, and at the longitude of the ground station. (You will supply your latitude.) Design constraints are: the carrier frequency must be in the range 1–6 GHz (actually, there are specific allowed frequencies, but neglect that fact here in your preliminary design study), maximum receiver noise temperature is 150 K, and maximum effective antenna areas are 1 m^2 for the satellite and 100 m^2 for the ground station. Determine the best modulation and detector type, bandwidth and carrier frequency, and minimum estimated transmitter power required to give an output S/N ratio of 6 dB. State your assumptions.

CHAPTER

6 Angle Modulation

A continuous-wave (CW) sinusoidal signal can be varied by changing its amplitude and its phase angle. Recalling Eq. (5.2), we write

$$\phi(t) = a(t) \cos [\omega_c t + \gamma(t)].$$

In Chapter 5 we kept $\gamma(t)$ constant and varied $a(t)$ proportional to $f(t)$. This introduced the concept of amplitude modulation. Now we shall investigate the case in which $a(t) = A$ (a constant) and the phase angle $\gamma(t)$ is varied in proportion to $f(t)$. This introduces the concept of angle modulation.

6.1 FM AND PM

The angle of a sinusoidal signal is described in terms of a frequency and/or a phase angle. Before proceeding here, however, we must decide precisely what we mean by the frequency of a sinusoid. If a sinusoid has a constant angular rate ω_0, then we say that the frequency of the sinusoid is ω_0 radians per second. But what happens if the angular rate is not constant? It is helpful at this point to return to a phasor representation.

The phasor representation of a constant-amplitude sinusoid is shown in Fig. 6.1. This phasor has a magnitude A and a phase angle $\theta(t)$. If $\theta(t)$ increases linearly with time [that is, $\theta(t) = \omega_0 t$], we say that the phasor has an angular rate, or "frequency," of ω_0 radians per second. If the angular rate is not constant, we can still write a relation between the instantaneous angular rate $\omega_i(t)$ and $\theta(t)$:

$$\theta(t) = \int_0^t \omega_i(\tau)\, d\tau + \theta_0. \tag{6.1}$$

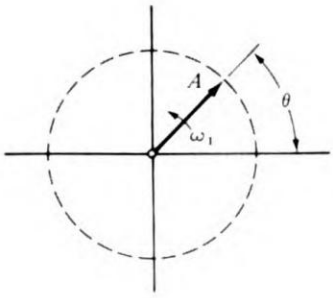


Figure 6.1 A general phasor representation.

Taking the derivative of both sides of Eq. (6.1), we have

$$\omega_i(t) = \frac{d\theta}{dt}. \tag{6.2}$$

Therefore we conclude that the instantaneous frequency of a sinusoidal signal is given by the time derivative of its phase. Note that this definition agrees with our usual concept of frequency when the phase is linear with time.

EXAMPLE 6.1.1

Determine the instantaneous frequency of the signal $\phi(t) = A \cos(10\pi t + \pi t^2)$.

Solution

$$\theta(t) = 10\pi t + \pi t^2$$

$$\omega_i(t) = \frac{d\theta}{dt} = 10\pi + 2\pi t = 2\pi(5 + t)$$

The frequency of $\phi(t)$ is 5 Hz at $t = 0$ and increases linearly at a rate of 1 Hz per second. Thus a quadratic phase shift gives a linear frequency dependence.

DRILL PROBLEM 6.1.1

Determine the instantaneous frequency of the following signal at $t = 0$: $\phi(t) = 5 \cos(10t + \sin 5t)$.

ANSWER: 15 rad/sec.

The concept of instantaneous frequency now permits us to describe two obvious possibilities for angle modulation (there are many more). If the phase angle $\theta(t)$ is varied linearly with the input signal $f(t)$, we can write

$$\theta(t) = \omega_c t + k_p f(t) + \theta_0 \tag{6.3}$$

where ω_c, k_p, θ_0 are constants. Because the phase is linearly related to $f(t)$, this type of angle modulation is called *phase modulation* (PM). The instantaneous frequency of this phase-modulated signal is

$$\omega_i = \frac{d\theta}{dt} = \omega_c + k_p \frac{df}{dt}. \tag{6.4}$$

Another possibility is to make the instantaneous *frequency* proportional to the input signal,

$$\omega_i = \omega_c + k_f f(t), \tag{6.5}$$

where ω_c, k_f are constants. Because the frequency is linearly related to $f(t)$, this type of angle modulation is called *frequency modulation* (FM). The phase angle of this frequency-modulated signal is

$$\theta(t) = \int_0^t \omega_i(\tau)\,d\tau = \omega_c t + \int_0^t k_f f(\tau)\,d\tau + \theta_0 . \tag{6.6}$$

A comparison of Eqs. (6.3)–(6.6) shows that PM and FM are closely related. The phase angle of the PM carrier signal is proportional to the modulating signal. But in FM the instantaneous frequency of the carrier signal is proportional to the modulating signal, so that the phase angle of the carrier is proportional to the integral of the modulating signal. Therefore, if the modulating signal $f(t)$ is first integrated and then used to phase modulate a carrier, the result is a signal that is frequency-modulated. Figure 6.2 is an illustration of FM and PM waveforms for given $f(t)$.

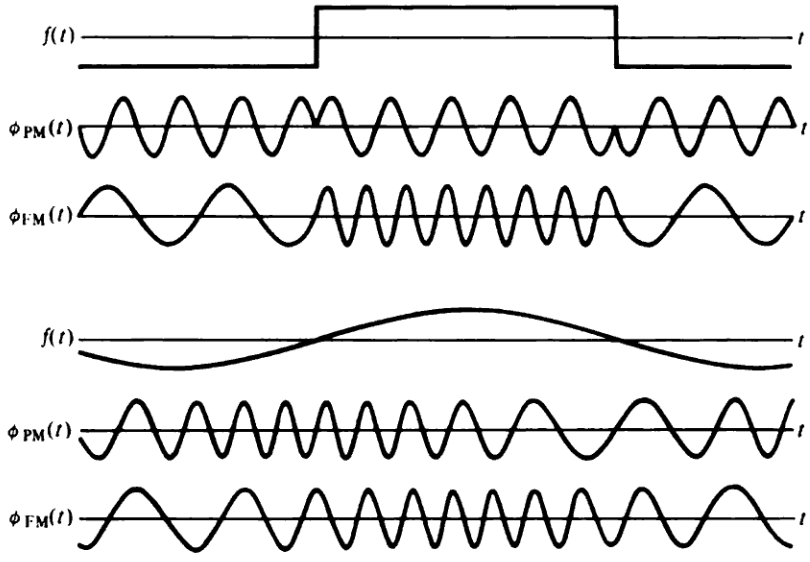


Figure 6.2 Examples of frequency and phase modulation.

Because frequency and phase modulation are so closely related, any variation in phase will necessarily result in a variation in frequency and vice versa. The essential difference between FM and PM is the nature of the dependency on the modulating signal. Although we shall discuss FM in more detail, our discussion is also valid for PM with only minor differences and these are pointed out in a later section.

In the case of AM signals, there was always a one-to-one correspondence between the modulated signal and the modulating signal. When this condition holds, the modulation is said to be *linear*.† For PM and FM this is not always true, however, as can be seen from the following reasoning.

† More formally, if $f(t)$ is the modulating signal and $\phi(t)$ is the modulated signal, the modulation is linear if $d\phi/df$ is independent of $f(t)$.

A general PM (or FM, with the appropriate modifications) signal can be represented by (note our return to complex notation)†

$$\phi_{PM}(t) = Ae^{j\theta(t)} = Ae^{j(\omega_c t+\theta_0)}e^{jk_p f(t)}. \tag{6.7}$$

Using a series expansion for the exponential modulation factor in Eq. (6.7), we have

$$\phi_{PM}(t) = Ae^{j(\omega_c t+\theta_0)}\left[1 + jk_p f(t) - \frac{1}{2!}k_p^2 f^2(t) - j\frac{1}{3!}k_p^3 f^3(t) + \cdots\right]. \tag{6.8}$$

From this result we conclude that, unless $|k_p f(t)| \ll 1$, angle modulation—in this case PM—is not linear. Therefore we can expect that, in general, the sidebands arising in angle modulation will not obey the principle of superposition. An analysis of spectra, etc., will have to be carried out choosing a particular waveform. When confronted by this choice, we shall use the sinusoidal waveform unless otherwise specified.

6.2 NARROWBAND FM

The linear condition in Eq. (6.8) maintains a linear modulation for FM and this appears to be a good place to begin. To lay the groundwork for the nonlinear modulation case we shall use a sinusoidal modulating signal. To be specific, let

$$f(t) = a \cos \omega_m t. \tag{6.9}$$

Because we are dealing with FM [cf. Eq. (6.5)],

$$\begin{aligned}\omega_i &= \omega_c + k_f f(t)\\ &= \omega_c + ak_f \cos \omega_m t,\end{aligned} \tag{6.10}$$

where k_f is the frequency modulation constant; typical units are in radians per second per volt. Defining a new constant called the *peak frequency deviation*,

$$\Delta\omega = ak_f, \tag{6.11}$$

we can rewrite Eq. (6.10) as

$$\omega_i = \omega_c + \Delta\omega \cos \omega_m t. \tag{6.12}$$

The phase of this FM signal is [cf. Eq. (6.6)] (let $\theta_0 = 0$ for convenience)

$$\theta(t) = \omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t = \omega_c t + \beta \sin \omega_m t, \tag{6.13}$$

† Although these are written in terms of PM signals, the conclusions are applicable to FM as well by substituting

$$k_f \int_0^t f(\tau)\, d\tau \quad \text{for} \quad k_p f(t).$$

where

$$\beta = \Delta\omega/\omega_m \tag{6.14}$$

is a dimensionless ratio of the peak frequency deviation to the modulating frequency.

The resulting FM signal is

$$\begin{aligned}\phi_{FM}(t) &= \mathcal{R}e\{Ae^{j\theta(t)}\} \\ &= \mathcal{R}e\{Ae^{j\omega_c t}e^{j\beta \sin \omega_m t}\}.\end{aligned} \tag{6.15}$$

Note that Eq. (6.15) can be rewritten as

$$\phi_{FM}(t) = A \cos(\omega_c t + \beta \sin \omega_m t). \tag{6.16}$$

Alternatively, an identity for the real part of a product (cf. Appendix A) may be used to rewrite Eq. (6.16) as

$$\phi_{FM}(t) = A \cos \omega_c t \cos(\beta \sin \omega_m t) - A \sin \omega_c t \sin(\beta \sin \omega_m t). \tag{6.17}$$

It is fairly obvious that either we will have to approximate this result, or seek some alternative approach. For small values of β we can write

$$\cos(\beta \sin \omega_m t) \approx 1, \tag{6.18}$$

$$\sin(\beta \sin \omega_m t) \approx \beta \sin \omega_m t. \tag{6.19}$$

The condition where β is small enough for these approximations is the condition for *narrowband FM* (NBFM). Usually a value of $\beta < 0.2$ is taken to be sufficient to satisfy this condition. Using Eqs. (6.18), (6.19) in Eq. (6.17), we obtain an approximate solution for small β:

$$\phi_{NBFM}(t) = A \cos \omega_c t - \beta A \sin \omega_m t \sin \omega_c t. \tag{6.20}$$

It is instructive to compare Eq. (6.20) with an equivalent expression for an AM signal:

$$\phi_{AM}(t) = A \cos \omega_c t + mA \cos \omega_m t \cos \omega_c t. \tag{6.21}$$

As suggested by a comparison of Eqs. (6.20), (6.21), the parameter β is called the *modulation index* of the FM signal.

Although the narrowband FM signal and the AM signal have similarities, they are distinctively different methods of modulation. The similarities and differences are portrayed by considering their phasor representations. Expanding Eq. (6.20) [or Eq. (6.15)] in phasor form, we have

$$\begin{aligned}\phi_{NBFM}(t) &= \mathcal{R}e\{Ae^{j\omega_c t}(1 + j\beta \sin \omega_m t)\} \\ &= \mathcal{R}e\{Ae^{j\omega_c t}(1 + \tfrac{1}{2}\beta e^{j\omega_m t} - \tfrac{1}{2}\beta e^{-j\omega_m t})\}.\end{aligned} \tag{6.22}$$

Similarly, Eq. (6.21) can be written in phasor form as

$$\begin{aligned}\phi_{AM}(t) &= \mathcal{R}e\{Ae^{j\omega_c t}(1 + m \cos \omega_m t)\} \\ &= \mathcal{R}e\{Ae^{j\omega_c t}(1 + \tfrac{1}{2}me^{j\omega_m t} + \tfrac{1}{2}me^{-j\omega_m t})\}.\end{aligned} \tag{6.23}$$

Taking the term $Ae^{j\omega_c t}$ as the reference (i.e., suppressing the continuous ω_c rotation), we show the phasor representation of each of these waveforms in Fig. 6.3. The resultant waveform can be found by rotating the entire phasor diagram at an angular rate of ω_c rad/sec and then taking the projection of the resultant on the real axis.

From Fig. 6.3, the differences between Eqs. (6.22) and (6.23) become quite evident. In the AM waveform, the modulation is added in phase with the carrier whereas in NBFM the modulation is added in quadrature with the carrier. The NBFM case gives rise to phase variations with very little amplitude change, whereas the AM case gives amplitude variations with no phase deviation.

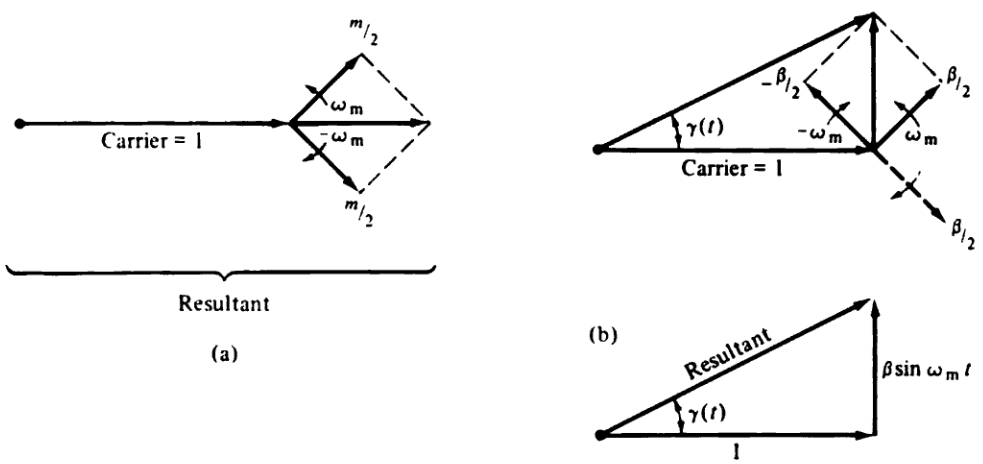


Figure 6.3 Phasor representation of (a) AM and (b) NBFM.

It is instructive to determine the limits on β from the results of the phasor diagrams in Fig. 6.3(b); the phase angle from the carrier is

$$\gamma(t) = \tan^{-1}(\beta \sin \omega_m t). \tag{6.24}$$

The instantaneous frequency deviation from the carrier frequency should be equal to $\Delta\omega \cos \omega_m t = \beta\omega_m \cos \omega_m t$ and is found by taking the derivative of this phase angle, or,

$$\frac{d\gamma}{dt} = \frac{\beta\omega_m \cos \omega_m t}{1 + \beta^2 \sin^2 \omega_m t} \approx \beta\omega_m \cos \omega_m t, \quad \text{if} \quad \beta^2 \sin^2 \omega_m t \ll 1. \tag{6.25}$$

The amplitude of the resultant phasor should be a constant (A); checking from the phasor diagram, we find

$$A\sqrt{1 + \beta^2 \sin^2 \omega_m t} \approx A \quad \text{if} \quad \beta^2 \sin^2 \omega_m t \ll 1. \tag{6.26}$$

Because $\sin^2 \omega_m t \leq 1$, these approximations are valid if $\beta^2 \ll 1$. Choosing $\beta^2 < 0.1$, we find that $\beta < 1/\sqrt{10} = 0.316$ is a reasonable bound for the narrowband approximation. Values as high as 0.50 can be used in practice if the resulting amplitude modulation is removed by amplitude-limiting the angle-modulated waveform.

The addition of the modulation in quadrature with the carrier in NBFM, in contrast to that in phase in AM, suggests a method of generation for the NBFM or NBPM case using phase shifters and balanced modulators as shown in Fig. 6.4. This method is commonly used in the generation of NBFM and NBPM signals. Note that even though we have been discussing the FM case, the PM case follows in the same manner.

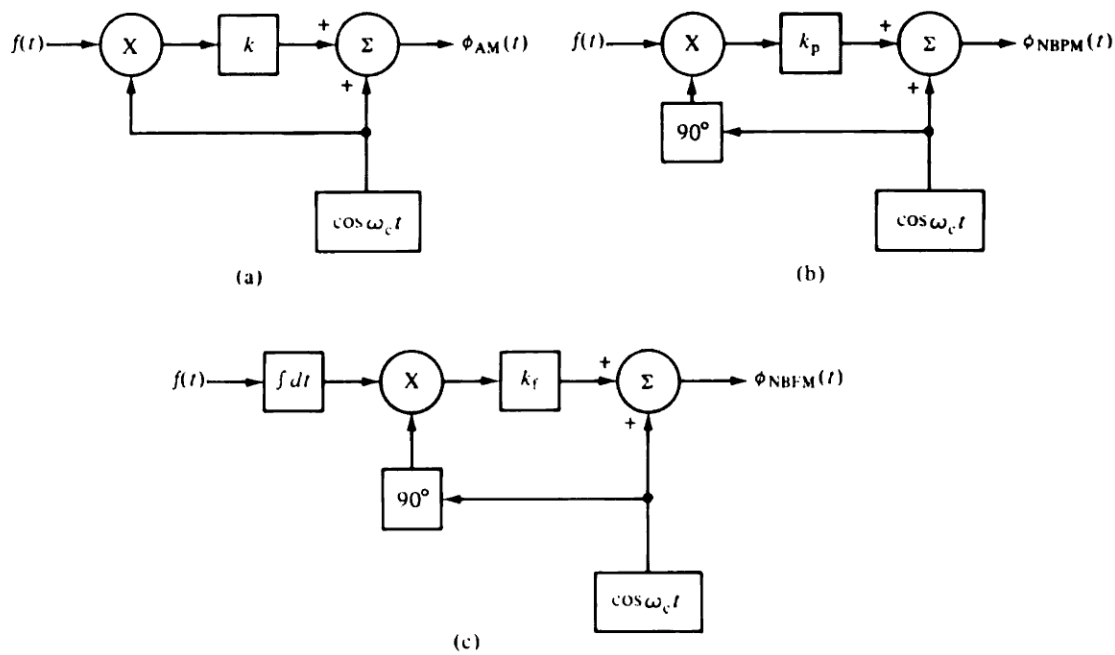


Figure 6.4 Generation of signals using balanced modulators: (a) AM; (b) NBPM; (c) NBFM.

Summarizing, narrowband FM (and PM), like AM, is an example of linear modulation. A major difference is that the modulation is added in phase with the carrier in AM, whereas it is added in quadrature with the carrier in NBFM. Both systems require a bandwidth of $W = 2\omega_m$ to transmit a signal of ω_m rad/sec in spectral width. The modulation index in FM is $\beta = \Delta\omega/\omega_m$ and a useful criterion for NBFM is $\beta < 0.2$.

DRILL PROBLEM 6.2.1

Calculate the maximum (peak)-percentage amplitude, phase, and frequency error incurred in using the phasor approximations to narrowband FM for the sinusoidal case when (a) $\beta = 0.20$; (b) $\beta = 0.50$.

ANSWER: (a) 2.0%, 1.3%, 3.8%; (b) 11.8%, 7.3%, 20.0%.

EXAMPLE 6.2.1

A comparison of the two phasor representations in Fig. 6.3 motivates us to investigate the simultaneous use of both amplitude modulation and narrowband frequency modulation for the possible elimination of one sideband. Investigate the possible use of this technique, for sinusoidal modulation, to generate SSB-LC signals.

Solution For simultaneous amplitude and frequency modulation we write

$$\phi(t) = \mathcal{R}e\{A(1 + m \cos \omega_m t) \exp [j(\omega_c t + \beta \sin \omega_m t)]\},$$

where m, β are the amplitude- and frequency-modulation indexes, respectively. For the NBFM condition, $\beta \ll 1$ so that

$$\exp (j\beta \sin \omega_m t) \approx 1 + j\beta \sin \omega_m t.$$

Under these conditions, we have

$$\phi(t) \approx \mathcal{R}e\{A(1 + m \cos \omega_m t)(1 + j\beta \sin \omega_m t) \exp (j\omega_c t)\},$$

which can be rewritten as

$$\phi(t) \approx \mathcal{R}e\left\{A\left(1 + \frac{m}{2}e^{j\omega_m t} + \frac{m}{2}e^{-j\omega_m t}\right)\left(1 + \frac{\beta}{2}e^{j\omega_m t} - \frac{\beta}{2}e^{-j\omega_m t}\right)e^{j\omega_c t}\right\}.$$

If $\beta \ll 1$ and $m \ll 1$, then the second-order term $(m\beta)$ in the approximation is very small. Neglecting these second-order effects, we get

$$\phi(t) \approx \mathcal{R}e\left\{A\left[1 + \left(\frac{m + \beta}{2}\right)e^{j\omega_m t} + \left(\frac{m - \beta}{2}\right)e^{-j\omega_m t}\right]e^{j\omega_c t}\right\}.$$

Note that, in contrast to AM, the two sidebands may have unequal magnitudes if some FM is also present. Now if we set $\beta = m$, we obtain the (approximate) SSB-LC signal

$$\phi(t) \approx \mathcal{R}e\{A[1 + me^{j\omega_m t}]e^{j\omega_c t}\},$$

or,

$$\phi(t) \approx A \cos \omega_c t + mA \cos (\omega_c + \omega_m)t.$$

DRILL PROBLEM 6.2.2

Sometimes when AM is the desired modulation, a combination of AM and FM can actually occur as a result of an imperfect modulator. Combined AM and NBFM is characterized on a spectrum-analyzer display by two sidebands of unequal amplitude. This arises because the AM sidebands are of the same phase but the NBFM sidebands

are of opposite phase. Because the intended modulation is AM, the incidental FM introduced is assumed to be the smaller of the two effects.

As an example, suppose that a spectrum-analyzer measurement of the output of a modulator using sinusoidal modulation indicates a carrier line of unit magnitude, an upper sideband line of magnitude 0.45 and a lower sideband line of magnitude 0.35. Calculate the percentage AM and the percentage FM present using the result of Example 6.2.1.

ANSWER: 80%; 10%.

6.3 WIDEBAND FM

Up to this point we have relied heavily on the use of the Fourier transform of a general signal $f(t)$ to give us the spectral density $F(\omega)$. However, if the value of β is not small, the Fourier transform of a general angle-modulated waveform cannot be evaluated. For specific cases the integration can be performed numerically or in terms of tabulated values. Therefore we shall first try to establish some bounds on the spectral density before we are forced to restrict the analysis to a few given modulating signal waveforms.

A measure of the peak amplitude-to-frequency conversion is the peak frequency deviation, $\Delta\omega$. This represents the maximum amount that ω_i deviates from the "average" value of ω_c. This is demonstrated for two differing cases in Fig. 6.5.

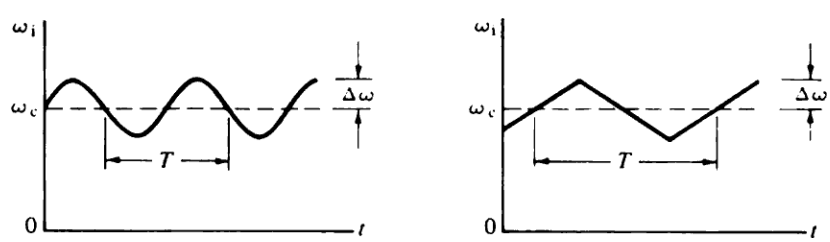


Figure 6.5 Definition of maximum (peak) frequency deviation.

There are two identifiable mechanisms in the description of the spectrum of an FM waveform. The first is attributable to the rate of change of the modulating signal; i.e., the frequency content of the modulating signal. The second effect, peculiar to FM, is the proportionality between the amplitude of the modulating signal and the instantaneous frequency of the FM signal. The instantaneous frequency follows the amplitude of the modulating signal, but this does not imply necessarily that the spectral density follows the same pattern. The concept of instantaneous frequency and the frequency used in the Fourier transform are not identical.

In the NBFM approximation, it is seen that the second effect was neglected in favor of the first since $\Delta\omega \ll \omega_m$. In fact, we now see that for the sinusoidal case the

modulation index $\beta = \Delta\omega/\omega_m$ gives us a relative measure as to the importance of these two effects in FM.

The idea of a modulation index can be extended to more general waveforms. For a general pulse waveform we can define a peak frequency deviation $\Delta\omega$ and a time duration T; if the waveform is periodic, then T is the period. The product $\beta_1 = (\Delta\omega/2\pi)T$ is a dimensionless number, called a *dispersion index*, which takes the place of the modulation index for more general modulation waveforms. It is easy to see that $\beta_1 \to \beta$ for sinusoidal modulation. For very low dispersion indexes the spectral content of a modulating signal largely controls the magnitude of the FM spectral density. For very high dispersion indexes, the amplitude-to-frequency conversion largely controls the magnitude spectral density. Phase effects are not as predictable because they depend on the relative phasing between signals. What happens for intermediate values must be examined on the basis of each given type of signal.

Returning to the sinusoidal case of FM, let $f(t) = a \cos \omega_m t$ and $\omega_i = \omega_c + \Delta\omega \cos \omega_m t$. The spectral content of the modulating signal is at ω_m rad/sec. The peak amplitude-to-frequency conversion is $\Delta\omega$ rad/sec. Then for very low values of $\beta = \Delta\omega/\omega_m$ (that is, $\Delta\omega \ll \omega_m$) the spectrum will be band-limited to $2\omega_m$. On the other hand, for very high values of $\beta = \Delta\omega/\omega_m$ (that is, $\Delta\omega \gg \omega_m$) the amplitude-to-frequency conversion will predominate and we would expect the bandwidth to be on the order of $2\Delta\omega$. We therefore have some rather intuitive bounds on the bandwidth at both extremes.

6.3.1 ★ General Approximations

Another general intuitive comment can be made here before restricting ourselves to specific waveforms. If we let $\beta_1 \to \infty$ ($\beta \to \infty$ for the sinusoidal case), we would expect the amplitude-to-frequency conversion to completely predominate. From the concept of a spectral density we would then expect the spectral magnitudes to be in proportion to the fractional time spent at each frequency.† For example, let $f(t) = a \cos \omega_m t$ so that the frequency deviation about the carrier, $\omega_i' = \omega_i - \omega_c$, is

$$\omega_i' = \Delta\omega \cos \omega_m t\,, \tag{6.27}$$

or

$$t = \frac{1}{\omega_m} \cos^{-1}\left(\frac{\omega_i'}{\Delta\omega}\right) \qquad \text{for } |\omega_i'| \le \Delta\omega\,. \tag{6.28}$$

The fractional amount of time per unit of frequency is‡:

$$\frac{1}{T}\left|\frac{dt}{d\omega_i'}\right| = \frac{1/(2\pi)}{\Delta\omega\sqrt{1 - (\omega_i'/\Delta\omega)^2}} \qquad \text{for } |\omega_i'| \le \Delta\omega\,. \tag{6.29}$$

† This is sometimes referred to as Woodward's theorem.

‡ The reader with some knowledge of probability will recognize this as the probability density function of the modulating waveform for uniform phase.

Therefore as $\beta_1 \to \infty$ (in this case $\beta \to \infty$) the magnitude weighting of the spectral density of the FM waveform will approach the shape shown in Fig. 6.6 over band limits $2\Delta\omega$ in width. Note that this is based on a signal T units long; for the periodic case the spectral density will be composed of impulses with weights determined from this curve. Effects of phase may cause the individual components to vary somewhat from this approximation. This result gives a relative distribution of power; the correct scaling factors can be found by setting the integral of this result equal to the average power in the modulated waveform.

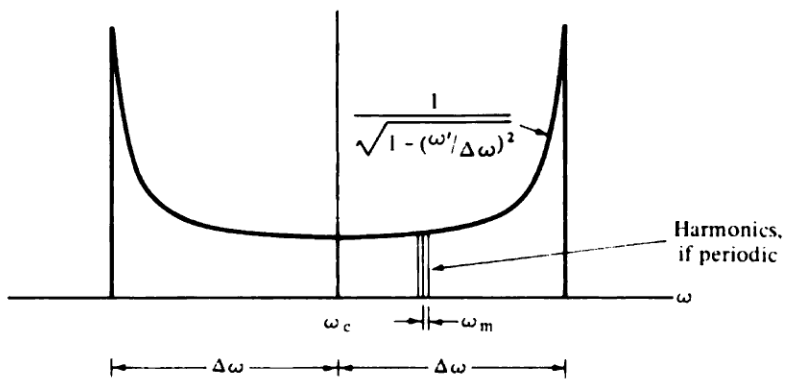


Figure 6.6 Approximation to the magnitude FM spectral density as $\beta \to \infty$, sinusoidal case.

EXAMPLE 6.3.1

A sinusoidal signal at a frequency of ω_c rad/sec is frequency-modulated by the sawtooth waveform shown in Fig. 6.7(a). The peak frequency deviation on each side of the carrier is $\Delta\omega$ rad/sec, as shown in Fig. 6.7(b). Describe the approximate magnitude spectral density as the dispersion index of the system becomes very large.†

Solution As $\beta_1 \to \infty$, the bandwidth approaches $2\Delta\omega$ and the magnitude spectrum is approximated by

$$\omega_i' = \frac{\Delta\omega}{T/2}t, \qquad -T/2 < t < T/2,$$

$$t = \frac{T}{2\Delta\omega}\omega_i',$$

$$\frac{1}{T}\left|\frac{dt}{d\omega_i'}\right| = \frac{1}{2\Delta\omega}, \qquad -\Delta\omega < \omega_i' < \Delta\omega.$$

† This is a simplified version of the type of modulation that a bat uses (at ultrasonic frequencies) for navigation and target location. It is also used for radar purposes.

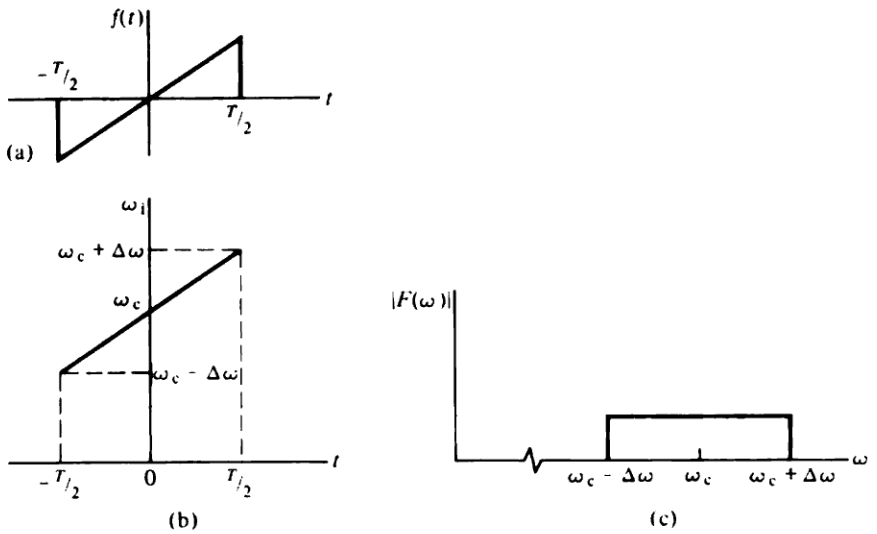


Figure 6.7 Example of FM spectral density as $\beta_1 \to \infty$.

This is shown in Fig. 6.7(c). If the modulating signal were repeated periodically, a series of impulses would be present spaced by ω_0 units. A numerical example of the computation of a magnitude spectrum for $\beta_1 = 50$ is shown in Fig. 6.8.

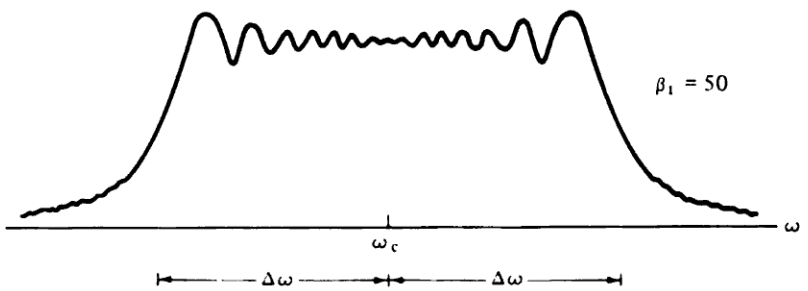


Figure 6.8 Computed magnitude spectrum for the FM discussed in Example 6.3.1.

DRILL PROBLEM 6.3.1

A sinusoidal signal at a frequency of ω_c rad/sec is frequency-modulated by a symmetrical square wave. The peak frequency deviation is $\Delta\omega$. Describe the approximate magnitude spectral density as the dispersion index becomes very large.

ANSWER: $\frac{1}{2}\delta(\omega - \omega_c + \Delta\omega) + \frac{1}{2}\delta(\omega - \omega_c - \Delta\omega)$.

6.3.2 Sinusoidal Case

Having obtained some intuitive insights into the mechanisms of FM, we now seek to extend our knowledge by using the Fourier transform. As pointed out earlier,

however, it is not possible to evaluate the Fourier transform of a general FM waveform and so we restrict the analysis here to pure sinusoids. Although pure sinusoids make for rather uninteresting communications, the results of the analysis hopefully will permit us to draw some more general conclusions. Because FM is a nonlinear modulation we cannot appeal to superposition here.

As before, we choose $f(t) = a \cos \omega_m t$; for FM,[†]

$$\begin{aligned}\omega_i(t) &= \omega_c + ak_f \cos \omega_m t \\ &= \omega_c + \Delta\omega \cos \omega_m t\,,\end{aligned}$$

and

$$\begin{aligned}\theta(t) &= \int_0^t \omega_i(\tau)\, d\tau \\ &= \omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t \\ &= \omega_c t + \beta \sin \omega_m t\,.\end{aligned}$$

Using complex notation,

$$\begin{aligned}\phi_{FM}(t) &= \mathcal{R}e\{Ae^{j\theta(t)}\} \\ &= \mathcal{R}e\{Ae^{j\omega_c t} e^{j\beta \sin \omega_m t}\}\,. \qquad (6.30)\end{aligned}$$

The second exponential in Eq. (6.30) is a periodic function of time with a fundamental frequency of ω_m rad/sec. It can be expanded in a Fourier series,

$$e^{j\beta \sin \omega_m t} = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_m t}\,, \qquad (6.31)$$

where

$$F_n = \frac{1}{T} \int_{-T/2}^{T/2} e^{j\beta \sin \omega_m t} e^{-jn\omega_m t}\, dt\,. \qquad (6.32)$$

Making a change of variable $\xi = \omega_m t = (2\pi/T)t$, we get

$$F_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(\beta \sin \xi - n\xi)}\, d\xi\,. \qquad (6.33)$$

This integral can be evaluated numerically in terms of the parameters n and β, and because it occurs in many physical problems it has been tabulated extensively. It is a function of n and β, denoted by $J_n(\beta)$, and is called the Bessel function of the first kind (signified by the "J") of order n and argument β.[‡] Note that in our case n is an integer (negative and positive) and β is a continuous variable (positive values only). Some of these functions are plotted in Fig. 6.9. Though we do not wish to get in-

[†] A constant term is introduced if the lower limit in the integral is not zero; this does not change the analysis and will be omitted for convenience.

[‡] A table of Bessel functions is given in Appendix G.

volved in discussing the detailed characteristics of Bessel functions, we do have use for the following properties:

1. $J_n(\beta)$ are real valued,
2. $J_n(\beta) = J_{-n}(\beta)$, for n even,
3. $J_n(\beta) = -J_{-n}(\beta)$, for n odd, (6.34)
4. $\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1$.

Using these results we can rewrite Eq. (6.31) as

$$e^{j\beta \sin \omega_m t} = \sum_{n=-\infty}^{\infty} J_n(\beta) e^{jn\omega_m t}, \tag{6.35}$$

and Eq. (6.30) becomes

$$\phi_{FM}(t) = \mathcal{R}e\left\{A e^{j\omega_c t} \sum_{n=-\infty}^{\infty} J_n(\beta) e^{jn\omega_m t}\right\}, \tag{6.36}$$

$$\phi_{FM}(t) = A \sum_{n=-\infty}^{\infty} J_n(\beta) \cos(\omega_c + n\omega_m)t. \tag{6.37}$$

This can be expanded, if desired, using properties (2) and (3) of Eq. (6.34).

From these results, it is evident that an FM waveform with sinusoidal modulation, in contrast to AM, has an infinite number of sidebands. However, the magnitudes of the spectral components of the higher-order sidebands become negligible and, for all practical purposes, the power is contained within a finite bandwidth. Plots of the sideband magnitudes for several different values of β are shown in Fig. 6.10. Note that β can be varied by varying $\Delta\omega$ or by varying ω_m, as demonstrated in Fig. 6.10.

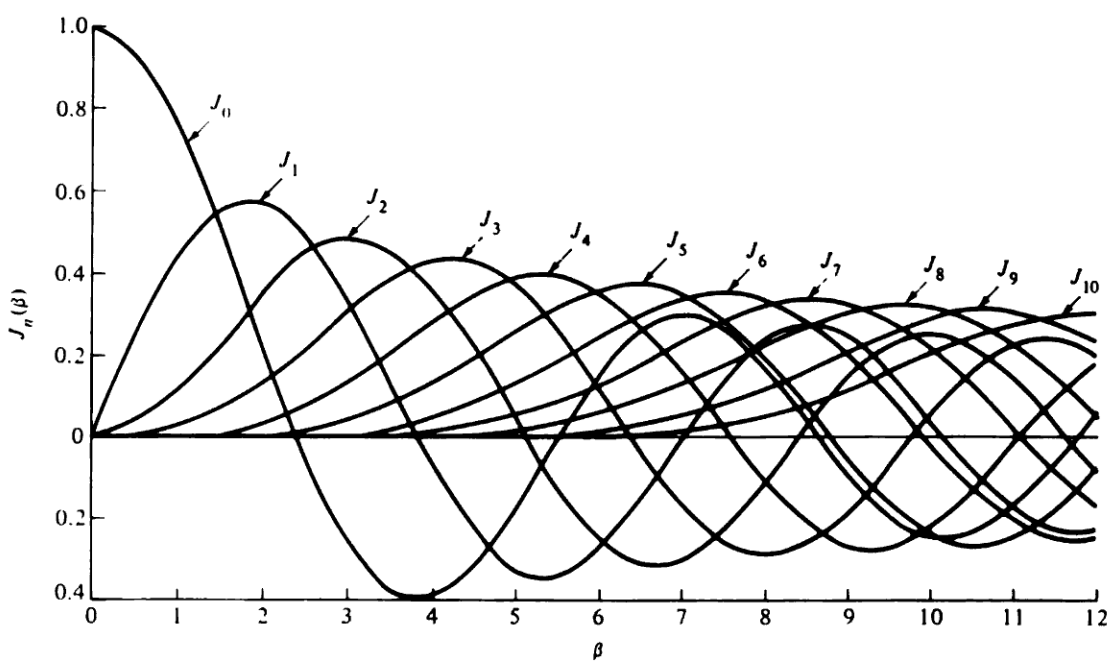


Figure 6.9 Plot of Bessel function of the first kind, $J_n(\beta)$.

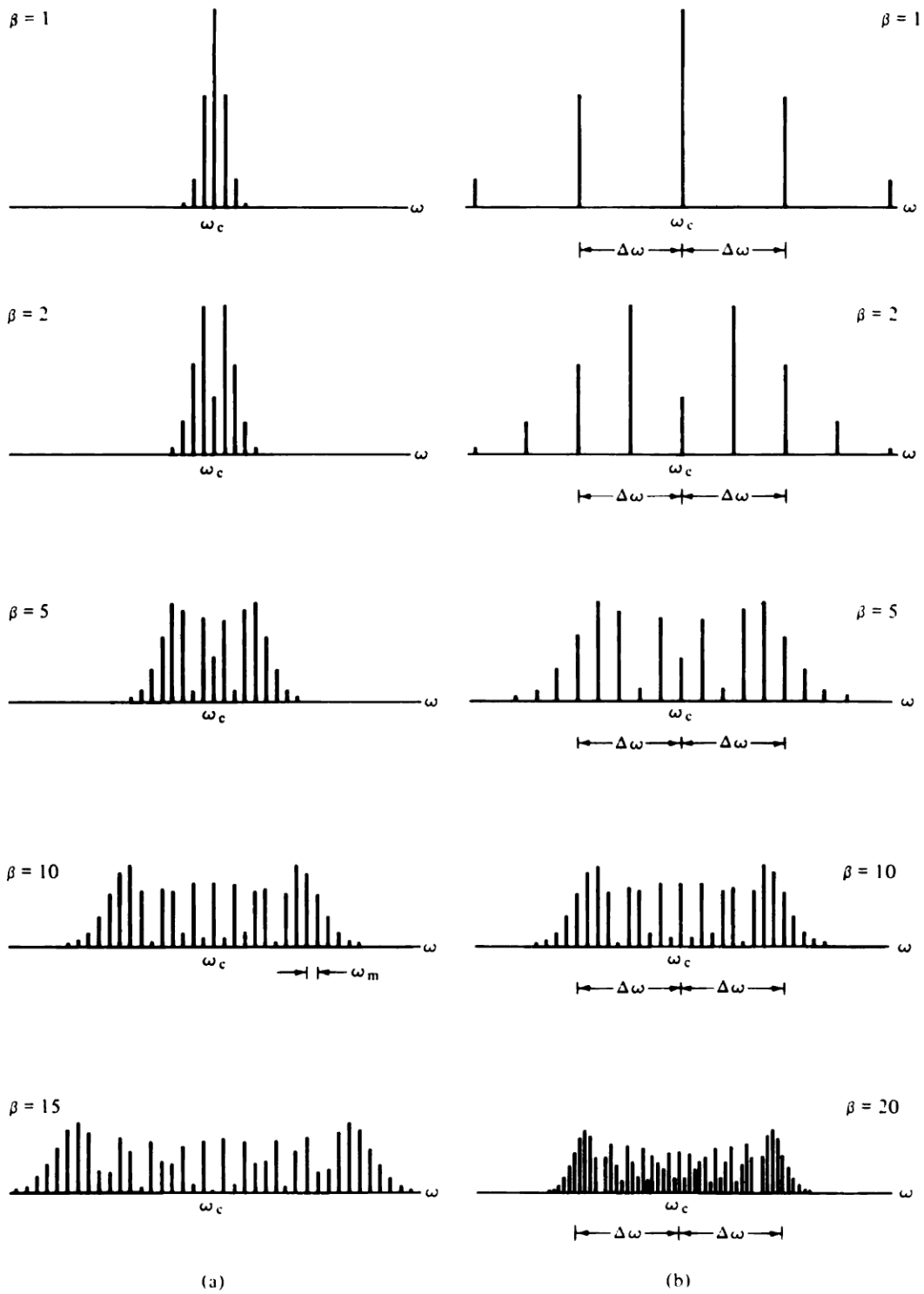


Figure 6.10 Magnitude line spectra for FM waveforms with sinusoidal modulation: (a) for constant ω_m; (b) for constant $\Delta\omega$.

How many sidebands are important to the FM transmission of a signal? This will depend on the intended application and the fidelity requirements. A rule commonly adopted is that a sideband is *significant* if its magnitude is equal to or exceeds 1% of the unmodulated carrier, i.e., if

$$|J_n(\beta)| \geq 0.01 . \tag{6.38}$$

The actual number of significant sidebands for different values of β can be found from a plot or a table of Bessel functions. It can be seen from the plots of Fig. 6.10 that the $J_n(\beta)$ diminish rapidly for $n > \beta$, particularly as β becomes large. A graph of the ratio n/β for $|J_n(\beta)| \geq 0.01$ is shown in Fig. 6.11 and it is seen that the ratio approaches one as β becomes very large. The bandwidth for very large β can then be approximated by taking the last significant sideband at $n = \beta$ so that

$$W = 2n\omega_m \approx 2\beta\omega_m = 2\frac{\Delta\omega}{\omega_m}\omega_m ,$$

or

$$W \approx 2\Delta\omega \quad \text{for large } \beta. \tag{6.39}$$

For very small values of β, the only Bessel functions of significant magnitude (see Fig. 6.9) are $J_0(\beta)$ and $J_1(\beta)$. Therefore the bandwidth for the narrowband case (verifying an earlier result) is

$$W \approx 2\omega_m \qquad \text{for small } \beta. \tag{6.40}$$

We now have bounds on the limiting cases. It would be convenient to have a more general rule to take care of the intermediate cases and, if possible, also approach the limiting cases in a continuous manner. One such rule was proposed by J. R. Carson†:

$$W \approx 2(\Delta\omega + \omega_m) , \tag{6.41a}$$

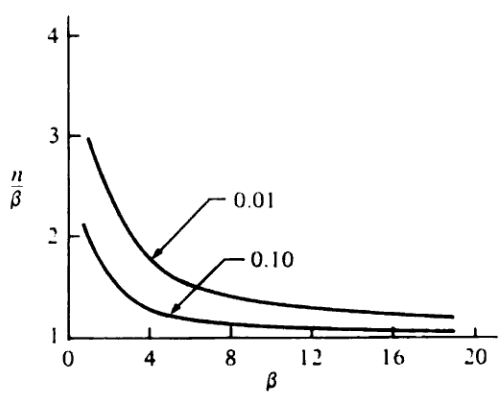


Figure 6.11 Number of FM sidebands for which $|J_n(\beta)| \geq 0.01$ and $|J_n(\beta)| \geq 0.10$.

† J. R. Carson was one of the first to investigate FM in the 1920s. See, for example, J. R. Carson, "Notes on the Theory of Modulation," reprinted in *Proceedings of the IEEE*, vol. 51 (1951): 893–896.

which can also be written as

$$W \approx 2\omega_m(1 + \beta). \tag{6.41b}$$

Carson's rule approaches the correct limits for both very large and very small β; it is widely used in practice because it gives a very convenient approximation that is reasonably accurate. It always gives less bandwidth than our definition of significant sidebands, with a maximum bandwidth error in the neighborhood of $\beta = 1$. The average power in the sidebands neglected, however, is small and less than 1% of the total average power in the FM waveform. In fact, the approximation is good enough that we now release the restriction that the modulating signal be purely sinusoidal and make the wide generalization that Carson's rule holds for general modulating signals that are band-limited and have finite power. An intuitive justification for this is that the two terms in Carson's rule display the effects of the two mechanisms in the generation of FM, and that these effects on the bandwidth are additive. It works surprisingly well.

The analytical method used here in expanding in a Fourier series is a very powerful one, and can be used for more general periodic modulating signals. This is explored in problems at the end of this chapter. Because we must integrate the modulating signal in FM to obtain the phase, we require that the modulating signal for FM have zero mean value.

EXAMPLE 6.3.2

A 10 MHz carrier is frequency-modulated by a sinusoidal signal such that the peak frequency deviation is 50 kHz. Determine the approximate bandwidth of the FM signal if the frequency of the modulating sinusoid is (a) 500 kHz; (b) 500 Hz; (c) 10 kHz.

Solution

a) $\beta = \dfrac{\Delta f}{f_m} = \dfrac{50}{500} = 0.10$

This is a narrowband FM signal; $B \approx 2f_m = 1$ MHz.

b) $\beta = 100$; this is the wideband case and

$$B \approx 2\Delta f = 100 \text{ kHz (Carson's rule gives 101 kHz).}$$

c) $\beta = 5$; use of Carson's rule gives $B \approx 2(\Delta f + f_m) = 120$ kHz. A more accurate method is to use Fig. 6.9 or Appendix G to find the number n of significant sidebands:

$$B = 2nf_m = 2(8)(10 \text{ kHz}) = 160 \text{ kHz}.$$

A magnitude line spectrum for $\beta = 5$ is shown in Fig. 6.10; in this case the spacing between lines would be 10 kHz.

DRILL PROBLEM 6.3.2

Repeat Example 6.3.2 if the peak frequency deviation were decreased to 20 kHz.

ANSWER: (a) 1 MHz; (b) 41 kHz; (c) 80 kHz (60 kHz if you use Carson's rule).

DRILL PROBLEM 6.3.3

A given FM signal is

$$\phi_{FM}(t) = 10 \cos [10^6 \pi t + 8 \sin (10^3 \pi t)].$$

Determine the following: (a) the carrier frequency, f_c; (b) the modulation index, β; (c) the peak frequency deviation, Δf.

ANSWER: (a) 500 kHz; (b) 8; (c) 4 kHz.

6.3.3 Commercial FM Transmissions

As noted earlier, narrowband FM is linear and therefore much of the analysis for AM applies. Advantages in using narrowband FM over AM include the possibility of a response to zero Hz (important in telemetry and recording) and the rejection of large noise pulses (as a result of clipping, or limiting, the amplitude of the waveform) which may tend to saturate the receiver. Narrowband FM is used primarily in telemetry and mobile communications.

Provided that we are content with only the bandwidth, we can apply our knowledge of purely sinusoidal FM to more general waveforms also in the wideband case. For wideband FM we noted that the bandwidth depended mainly on the peak frequency deviation, Δf. This in turn depends, for a given modulator constant, on the amplitude of the modulating signal. Therefore some limit must be placed on the modulating signal to avoid excessive bandwidths even though the bandwidth of the modulating signal may be well-defined.

For commercial FM broadcasting, the Federal Communications Commission (FCC) in the United States assigns carrier frequencies spaced at 200 kHz intervals in the range 88–108 MHz and fixes the peak frequency deviation at 75 kHz.[†] The 200 kHz between station assignments, in comparison with 10 kHz for AM broadcasting, allows the transmission of high-fidelity program material with room to spare, and wideband FM is used to fill the band. Suppose we take the modulating frequency f_m to be 15 kHz (typically the maximum audio frequency in FM transmissions). Use of Carson's rule then yields a bandwidth of $B \approx 2(\Delta f + f_m) = 180$ kHz, well within the required

[†] See Appendix C.

bandwidth. Our sinusoidal analysis indicates that $\beta = 5$ and the bandwidth occupied by significant sidebands is

$$2(8)(15 \text{ kHz}) = 240 \text{ kHz}$$

(see Fig. 6.9, 6.10, or 6.11). The discrepancy, of course, lies in the definition of bandwidth. However, we chose an extreme case as far as typical audio transmission is concerned, because we assumed that the 15 kHz tone was set at the maximum amplitude to produce a peak frequency deviation of 75 kHz. Typical program material does not contain as much at the higher frequencies. For lower audio frequencies the value of β increases and the bandwidth occupied by the significant sidebands approaches the wideband limit of $2\Delta f = 150$ kHz. (For audio signals with full maximum amplitude and frequencies below about 5 kHz all significant sidebands are within the 200 kHz bandwidth.) Note that if the amplitude weighting is uniform, it is the highest modulating frequency that governs the final bandwidth.

The transmission of one audio channel leaves room for additional program material within the bandwidth allocated. Stereo multiplexing and other auxiliary transmissions often occupy the higher frequency portions of the modulating spectrum. These were discussed in Chapter 5. To keep the bandwidth restricted, the maximum amplitude of these transmissions is reduced. The spectrum of a typical commercial transmission before the FM transmitter is shown in Fig. 5.12.

In the FM station, the left (L) and right (R) audio signals are derived from microphones, records, tapes, etc., and a preemphasis is applied to each channel (this is discussed later in this chapter). For stereo broadcasts, a pilot subcarrier at 19 kHz is permitted 10% of the total peak-frequency deviation (of 75 kHz). When there is a pause in program material (and no auxiliary transmissions), the modulation index is $\beta = (10\%)(75 \text{ kHz})/(19 \text{ kHz}) = 0.395$. This is approximately in the narrowband condition. Thus when there is a pause in program material, a stereo FM broadcast can be identified on a spectrum analyzer by a large carrier line plus two first-order sidebands each spaced 19 kHz from the carrier.

The Subsidiary Communications Authorization (SCA) system permits a commercial FM station to add another broadcasting channel in addition to the monophonic and stereo channels. The SCA transmissions carry no commercial messages and are intended for private subscribers who pay a fee for background music in stores, physicians' offices, etc. In contrast, the other FM transmissions are for general public use and are supported by commercial advertisements. The SCA channel uses narrowband FM. The subcarrier center frequency is usually set at 67 kHz, although this choice is not set by the FCC. A total peak-frequency deviation not exceeding 75 kHz is still required. For monophonic transmission only, this entire 75 kHz is available. If SCA is used with mono, the FCC limits the SCA portion to 30% of the maximum peak-frequency deviation, leaving 70% for the mono channel.

In stereo broadcasting with no SCA, 10% of the maximum peak deviation is used for the 19 kHz pilot subcarrier, leaving 90% to be divided between the (L + R) and (L − R) stereo channels. The average amplitudes of the L and R channels are normally kept equal. The maximum (L + R) amplitude is set to provide 90% modula-

tion when the (L − R) amplitude is zero. Now if (L − R) is maximum, then (L + R) will go to zero, and if either L or R goes to zero, (L + R) and (L − R) will each take a maximum of 45% of the total peak frequency deviation. Thus there is a seesaw effect between the (L + R) and (L − R) channels such that the total does not exceed 90% of the peak frequency deviation capability allowed.

When used with stereo multiplexing, the SCA channel is limited to 10% of the maximum peak frequency deviation. This leaves 80% for the stereo channels (i.e., 10% for SCA plus 10% for the 19 kHz pilot subcarrier). The system performance is the same as in the preceding paragraph except that now the stereo channels are allowed 80% of the maximum peak frequency deviation instead of 90%. Also, with the relatively low peak frequency deviation allowed, the SCA transmission does not have a very good signal-to-noise ratio and is used for only local coverage. Station muting (e.g., dropping the subcarrier to actuate an audio silencer circuit in the receiver) is often employed to silence the noise between records or tapes in the SCA transmissions.

Public FM broadcasting is an example of the use of DSB-SC and NBFM methods to frequency-multiplex several channels before using wideband FM for the final transmission. Frequency modulation is also used for the audio in commercial television transmissions. The peak frequency deviation for this use is fixed at 25 kHz by the FCC. Assuming a maximum audio frequency of 15 kHz, use of Carson's rule gives a bandwidth of 80 kHz for the sound channel of a television receiver.

The relatively large bandwidth required for commercial FM, as compared with AM, is the penalty for obtaining substantial improvement in noise and interference rejection. This noise rejection increases with increasing Δf and, therefore, with increasing bandwidth. These topics will be discussed in a later section.

6.4 AVERAGE POWER IN ANGLE-MODULATED WAVEFORMS

For sinusoidal modulation, we can write [cf. Eq. (6.16)]

$$\phi_{FM}(t) = A \cos(\omega_c t + \beta \sin \omega_m t).$$

The mean-square value of this expression is

$$\overline{\phi_{FM}^2(t)} = A^2/2, \tag{6.42}$$

showing that the total average power in an FM waveform is a constant regardless of the modulation index. This is in contrast to AM where the total average power was proportional to the modulation index. This conclusion can be extended to any arbitrary band-limited modulating waveform.

Equation (6.42) can be verified by writing $\phi_{FM}(t)$ in a series expansion [cf. Eq. (6.37)],

$$\phi_{FM}(t) = A \sum_{n=-\infty}^{\infty} J_n(\beta) \cos(\omega_c + n\omega_m)t.$$

As a result of the orthogonality of the cosine terms, the mean-square value of the sum is equal to the sum of the mean-square values and we get

$$\overline{\phi_{\mathrm{FM}}^2(t)} = \tfrac{1}{2}A^2 \sum_{n=-\infty}^{\infty} J_n^2(\beta) . \tag{6.43}$$

But from property (4) in Eq. (6.34),

$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 ,$$

so that

$$\overline{\phi_{\mathrm{FM}}^2(t)} = A^2/2 .$$

The mean-square value of the unmodulated carrier is $A^2/2$. As the modulation index β is increased from zero and the sidebands are nonzero, the carrier component decreases. According to Eqs. (6.42) and (6.43), this takes place in such a manner as to always keep the total mean-square value constant. The mean-square value of each sideband is $\frac{1}{2}A^2J_n^2(\beta)$ (recall also that sidebands occur in pairs). The mean-square value, of course, is identical to the average power if the resistance is 1 ohm and is related to the average power by a constant (i.e., the resistance) for all other cases, so the conversion to units of power is straightforward.

It is possible to make any particular sideband, including the carrier, as small as desired by a proper choice of the modulation index β. From a table or graph of Bessel functions (e.g., Fig. 6.9), we see that the carrier term, $J_0(\beta)$, can be made zero for $\beta = 2.405, 5.52, \ldots$, and in these cases all of the average power is in the sidebands. These points are easy to read with a spectrum analyzer and serve as very convenient calibration points for β and Δf.

EXAMPLE 6.4.1

A given FM transmitter is modulated with a single sinusoid. The output for no modulation is 100 watts into a 50-ohm resistive load. The peak frequency deviation of the transmitter is carefully increased from zero until the first sideband amplitude in the output is zero. Under these conditions determine (a) the average power at the carrier frequency; (b) the average power in all the remaining sidebands; and (c) the average power in the second-order sidebands.

Solution a) Using Fig. 6.9 and Appendix G, we see that $J_1(\beta) = 0$ first occurs at $\beta \approx 3.8$ and that $J_0(3.8) \approx -0.40$. The average carrier power is then

$$P_c = \frac{J_0^2(3.8)}{J_0^2(0)}(100 \text{ W}) = 16 \text{ W}.$$

b) The average power in the sum of the remaining sidebands is

$$P_s = P_t - P_c = 100 \text{ W} - 16 \text{ W} = 84 \text{ W}.$$

c) $J_2(3.8) \approx 0.41$. The average power in the second-order sidebands is

$$P_c = 2\frac{J_2^2(3.8)}{J_0^2(0)}(100 \text{ W}) = 34 \text{ W}.$$

DRILL PROBLEM 6.4.1

Determine the peak amplitude of (a) the total waveform and (b) the upper second-order sideband in Example 6.4.1.

ANSWER: (a) 100 V; (b) 41 V.

DRILL PROBLEM 6.4.2

Show that the rms value of Eq. (6.16) can be written as

$$\sqrt{\overline{\phi^2(t)}} = \frac{A}{\sqrt{2}}\sqrt{J_0^2(\beta) + 2\sum_{n=1}^{\infty} J_n^2(\beta)}.$$

6.5 PHASE MODULATION

There is no basic difference between the mechanisms involved in the generation of phase modulation (PM) and frequency modulation (FM). In fact, the only difference is that the phase in the modulated waveform is proportional to the input signal amplitude in PM and to the integral of the input signal in FM. This introduces only a slight modification and we shall point that out here.

For an FM signal with the sinusoidal modulation $f(t) = a \cos \omega_m t$, the instantaneous frequency is

$$\begin{aligned}\omega_i(t) &= \omega_c + ak_f \cos \omega_m t \\ &= \omega_c + \Delta\omega \cos \omega_m t,\end{aligned}$$

where $\Delta\omega$ is the peak frequency deviation (in radians per second) and k_f is the frequency-modulator constant (in radians per second per volt). The modulation index, $\beta = \Delta\omega/\omega_m$, is a dimensionless number and serves as a guide to the behavior of the carrier and sidebands.

For PM with the same modulating signal we have

$$\begin{aligned}\theta(t) &= \omega_c t + ak_p \cos \omega_m t + \theta_0 \\ &= \omega_c t + \Delta\theta \cos \omega_m t + \theta_0 ,\end{aligned}$$

where $\Delta\theta$ is the peak phase deviation (in radians) and k_p is the phase-modulator constant (in radians per volt). The instantaneous frequency is

$$\begin{aligned}\omega_i(t) &= \frac{d\theta}{dt} \\ &= \omega_c - ak_p\omega_m \sin \omega_m t \\ &= \omega_c - \Delta\omega \sin \omega_m t .\end{aligned}$$

Thus we see that the peak frequency deviation in PM is proportional not only to the amplitude of the modulating waveform but also to its frequency; that is,

$$\Delta\omega = \begin{cases} ak_f & \text{for FM} \\ ak_p\omega_m = (\Delta\theta)\omega_m & \text{for PM} \end{cases} . \tag{6.44}$$

This makes PM less desirable to transmit when $\Delta\omega$ is fixed (as in commercial FM). There are some advantages in the demodulation of PM, however, which make its use desirable. (These will become more evident later in this chapter.) The role of the modulation index β remains the same as in FM. Formally, then, we can compute $\Delta\omega = ak_p\omega_m = \Delta\theta\omega_m$ and then proceed as if the modulation were FM as far as bandwidth, sidebands, etc. are concerned. Note that the numerical value of β is the peak phase deviation, $\Delta\theta$, in the PM case.

EXAMPLE 6.5.1

A carrier is phase modulated by a sinusoidal signal of 5 kHz and unit amplitude and the peak phase deviation is one radian. Calculate the bandwidth of the PM signal (a) using Carson's rule; and (b) using the definition of significant sidebands.

Solution a) $\Delta f = (\Delta\theta)f_m = 5$ kHz and Carson's rule gives

$$B \approx 2(\Delta f + f_m) = 20 \text{ kHz} .$$

b) $\beta = \Delta\theta = 1$; using a table of Bessel functions,

$$B = 2nf_m = 2(3)(5 \text{ kHz}) = 30 \text{ kHz} .$$

DRILL PROBLEM 6.5.1

Here we consider a PM system in which the phase may take on only two possible values—known as *phase-shift-keying* (PSK). Assume that a phase modulator is modu-

lated by a periodic symmetric square wave of unit amplitude. Determine the required value of the peak phase deviation, $\Delta\theta$ ($-\pi/2 \leq \Delta\theta < \pi/2$), such that the spectral component at the carrier frequency is zero in the output.

ANSWER: $\pm 90°$.

6.6 GENERATION OF WIDEBAND FM SIGNALS

One method of generating wideband FM signals is to produce a narrowband FM signal first and then to use frequency multiplication to increase the modulation index to the desired range of values. This is known as the indirect method of generating wideband FM signals. A second method—known as the direct method—is to vary the carrier frequency directly with the modulating signal. We shall now examine these two methods.

6.6.1 Indirect FM

Section 6.2 showed that the generation of narrowband PM is relatively easy and that narrowband FM can then be generated by first integrating the modulating signal. However, the modulation index obtainable by use of this method is restricted to very low values ($\beta < 0.2$ in theory; $\beta < 0.5$ in practice). To generate wideband FM, a method of increasing the modulation index must be used in this approach. The method used is that of the frequency multiplier.

A *frequency multiplier* is a nonlinear device designed to multiply the frequencies of the input signal by a given factor. For example, the input-output characteristic of an ideal square-law device is

$$e_o(t) = ae_i^2(t). \tag{6.45}$$

If the input signal is the FM signal,

$$e_i(t) = A \cos(\omega_c t + \beta \sin \omega_m t),$$

the output is

$$\begin{aligned} e_o(t) &= aA^2\cos^2(\omega_c t + \beta \sin \omega_m t) \\ &= (1/2)aA^2[1 + \cos(2\omega_c t + 2\beta \sin \omega_m t)]. \end{aligned} \tag{6.46}$$

The first term in this result is simply a constant level and is easily removed with a filter. We conclude that both the carrier frequency and the modulation index have been doubled in this process. In a similar manner, use of an nth law device followed by a filter yields a carrier and a modulation index that have been increased by a factor

of n. Equivalently, the peak frequency deviation $\Delta\omega$ has been increased by n in the multiplication (ω_m has remained unaltered).

In practice, very abrupt nonlinearities can be generated using special diodes (e.g., the varactor and the step-recovery diode) which yield many harmonic terms. It is possible to multiply by an order of magnitude or more in one step using these techniques. Limitations include the fact that losses incurred in the harmonic generation require additional amplification and small phase instabilities in the multiplication process accumulate and appear as noise in the output. With good design techniques, multiplication factors on the order of 10^3 are achievable with only a few degrees of phase noise.

Use of frequency multiplication increases the carrier of the FM waveform as well as the modulation index. This may result in very high carrier frequencies in order to achieve a given modulation index. To avoid this, frequency converters are often used to control the value of the carrier frequency. The frequency converter is essentially the same as discussed in connection with AM and translates the spectrum of a signal by a given amount but does not alter its spectral content. Block diagrams of the frequency multiplier and the frequency converter are shown in Fig. 6.12. Note carefully the differences between these two operations. In the frequency multiplier all spectral components of the input signal are multiplied by themselves (so that all cross-products are present), whereas in the frequency converter all spectral components of the input signal are multiplied by a sinusoid of a fixed frequency. The former operation spreads the spectral content (this can be verified using the frequency convolution property discussed in Chapter 3) and the latter translates the spectral content in frequency.

The method of obtaining a wideband FM waveform from a narrowband one using frequency multiplication is called the Armstrong indirect FM transmitter.† A block diagram of a typical Armstrong-type transmitter is shown in Fig. 6.13. As a result of the multiplication and heterodyne operations, it is difficult in this system to maintain the correct magnitude of carrier to sidebands, and thus it could not be used for an information signal with dc content. This drawback can be remedied with use of the phase-locked loop, a subject introduced in a subsequent section.

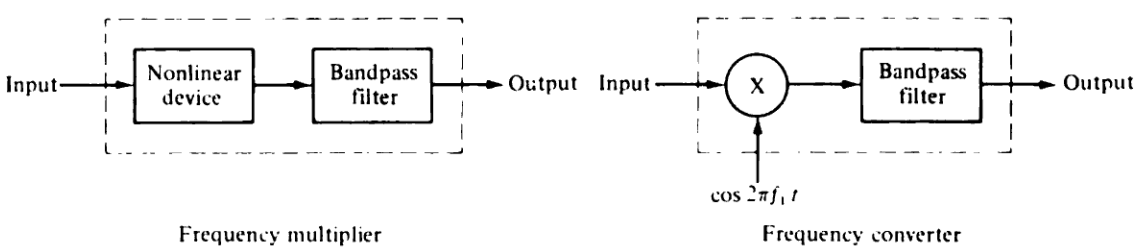


Figure 6.12 Block diagrams of frequency multiplication and frequency conversion.

† E. H. Armstrong was one of the first engineers to recognize the possible merits of FM broadcasting in the 1930s. See, for example, E. H. Armstrong, "A Method of Reducing Disturbance in Radio Signaling by a System of Frequency Modulation," *Proceedings of the IRE*, vol. 24 (1936); 689–740.

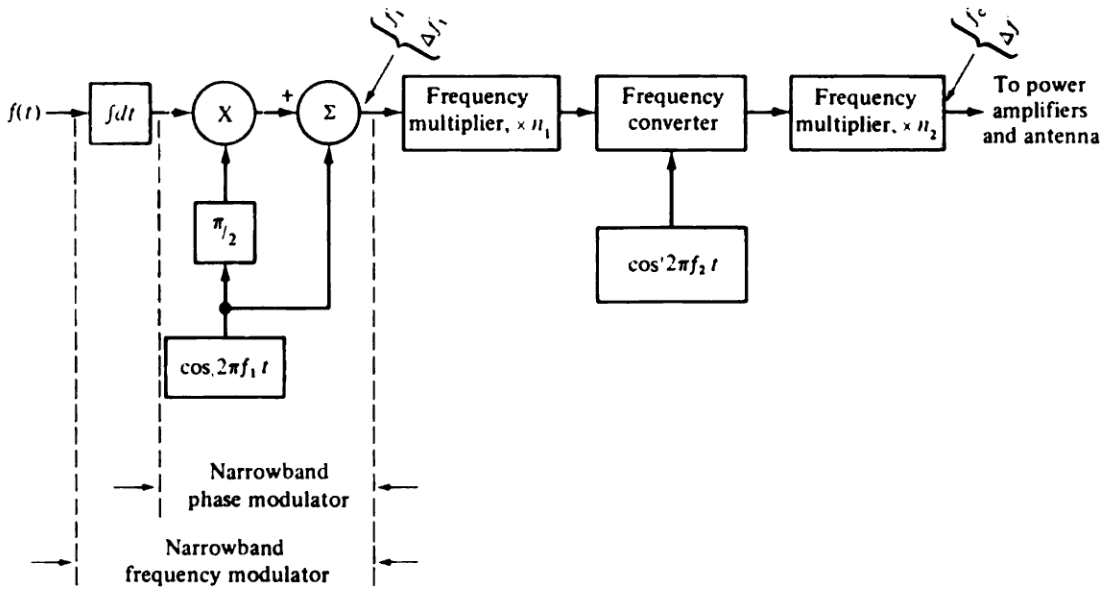


Figure 6.13 Block diagram of an indirect (Armstrong) FM transmitter.

EXAMPLE 6.6.1

A given angle-modulated signal has a peak frequency deviation of 20 Hz for an input sinusoid of unit amplitude and a frequency of 50 Hz. Determine the required frequency multiplication factor, n, to produce a peak frequency deviation of 20 kHz when the input sinusoid has unit amplitude and a frequency of 100 Hz, and the angle-modulation used is (a) FM; (b) PM.

Solution a) $\Delta f_2 = 20$ kHz; $\Delta f_1 = 20$ Hz; $n = \Delta f_2/\Delta f_1 = 1000$

b) $\Delta f_2 = 20$ kHz; $\Delta f_1 = (100/50)(20\text{ Hz}) = 40$ Hz; $n = \Delta f_2/\Delta f_1 = 500$

DRILL PROBLEM 6.6.1

Compute the carrier frequency f_c and the peak frequency deviation Δf of the output of the FM transmitter shown in Fig. 6.13 if $f_1 = 200$ kHz; $f_2 = 10.8$ MHz; $\Delta f_1 = 25$ Hz; $n_1 = 64$; $n_2 = 48$.

ANSWER: 96.0 MHz or 1132.8 MHz; 76.8 kHz.

6.6.2 Direct FM

In the direct method of generating FM the modulating signal directly controls the carrier frequency. An attempt is usually made to generate as wide a frequency deviation

as possible and thus these systems often require less frequency multiplication, if any, than those using the indirect method. On the other hand, because the frequency is controlled by the modulating voltage, the long-term frequency stability is not as good as the crystal-stabilized oscillators generally used in the indirect method. Thus the direct FM methods usually employ some auxiliary methods for frequency stabilization.

A common method used for generating FM directly is to vary the inductance or capacitance of a tuned electronic oscillator. If L and C are the inductance and capacitance, respectively, of a simple tuned circuit, the frequency of oscillation is

$$\omega_0 = \frac{1}{\sqrt{LC}}.$$

If L or C is varied, the output frequency will also vary. For very small variations (that is, $\Delta\omega \ll \omega_c$, where ω_c is the carrier frequency), the square-root relationship can be approximated by a linear term and the conversion can be made quite linear.

There are various ways to make the capacitance or inductance of a tuned circuit dependent on the input signal. One common method at medium and high frequencies is to use a reverse-biased semiconductor diode as a voltage-variable capacitance. Although any semiconductor diode exhibits some capacitance change with change in reverse bias, the type of diode frequently used for this application is the varactor diode. The percentage frequency deviation that can be attained in this manner is quite small. To increase the percentage frequency deviation, the frequency modulation is performed at a high frequency and then heterodyned down to a lower frequency.

Other methods that are used successfully at high frequencies include the reflex klystron and the reactance-tube modulator. The latter consists of a pentode that is operated in such a manner as to produce a capacitance which is proportional to the grid voltage over a wide range. At lower frequencies the control of RC oscillators with FET's and similar devices has been used. Any oscillator whose frequency is controlled by the modulating-signal voltage is called a *voltage-controlled oscillator*, or VCO.

EXAMPLE 6.6.2

A reverse-biased semiconductor diode can be used as a voltage-variable capacitance for frequency modulation. Assume that the capacitance of a given PN junction is given in terms of its reverse-bias voltage V by $C = C_0/\sqrt{1 + 2V}$. Such a diode is to be used as the capacitance in a parallel LC circuit tuned to a center frequency of 10 MHz when the reverse-bias voltage is 4 V.

a) Determine the modulation constant k_f (i.e., the frequency-voltage slope near center frequency).

b) Determine the peak frequency deviation permissible for a maximum error of 1% from a linear frequency-voltage characteristic.

Solution a) We can write the frequency f as

$$f = \frac{1}{2\pi\sqrt{LC}} = \frac{(1 + 2V)^{1/4}}{2\pi\sqrt{LC_0}}.$$

Letting $f = f_0$ when $V = V_0$ (i.e., at the operating point), we get

$$f = f_0 \frac{(1 + 2V)^{1/4}}{(1 + 2V_0)^{1/4}}.$$

Now let v be an incremental voltage about the operating point so that $V = V_0 + v$; also let $K = (1 + 2V_0)$ so that

$$f = f_0[1 + (2v/K)]^{1/4}.$$

Assuming that $v \ll K$, we can use the binomial expansion to obtain

$$f \approx f_0\left[1 + \frac{1}{4}\left(\frac{2v}{K}\right) - \frac{3}{32}\left(\frac{2v}{K}\right)^2 + \cdots\right].$$

The modulation constant k_f is the slope of the linear frequency-voltage characteristic and is given by

$$k_f = \frac{1}{4}\left(\frac{2v}{K}\right)\frac{f_0}{v} = \frac{f_0}{2K} = \frac{f_0}{2(1 + 2V_0)}.$$

For the given operating point, $f_0 = 10$ MHz and $V_0 = 4$ volts so that

$$k_f = 0.56 \text{ MHz/V}.$$

b) Most of the error will arise from the second-order term in the series expansion so that we require

$$\frac{\frac{3}{32}\left(\frac{2v}{K}\right)^2}{\frac{1}{4}\left(\frac{2v}{K}\right)} \leq 0.01$$

which gives

$$v \leq \frac{4K}{300}.$$

The peak frequency deviation is then

$$\Delta f = k_f v_{max} = \left(\frac{f_0}{2K}\right)\left(\frac{4K}{300}\right) = \frac{f_0}{150}$$

$$= 66.7 \text{ kHz}.$$

In some cases it is not necessary that the output voltage be sinusoidal, or the output may be wave-shaped by nonlinear shaping circuits or by filtering. In these cases it is possible — and quite attractive — to generate wideband FM digitally. Perhaps the simplest way is to control the oscillation frequency of a relaxation oscillator or a multivibrator with the modulating-signal voltage. A more accurate and stable method is to generate the zero crossings of a PM waveform using digital techniques. Basically this involves sampling the input waveform and then using a precision ramp generator that resets at each sample point, followed by a voltage-variable threshold. The point in time at which the threshold is exceeded is used to generate a short pulse to signify a zero crossing. Applying this sequence of pulses to a bandpass filter results in a wideband PM signal. This method is stable and accurate and is capable of generating signals of very wide bandwidth. A simplified block diagram is shown in Fig. 6.14. These operations are discussed in more detail in Chapter 7.

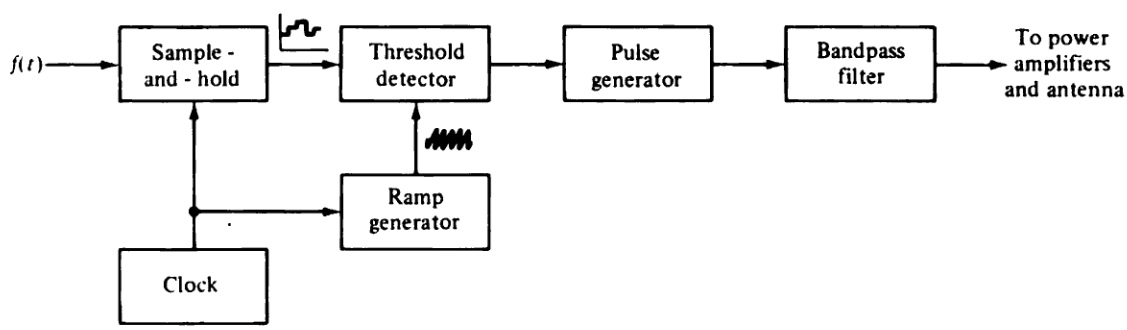


Figure 6.14 Digital generation of wideband PM.

6.6.3 FM Multiplexing

It is a common practice in data transmission to combine several channels of modulated signals using frequency multiplexing methods and then modulate a high-frequency carrier with the composite multiplexed signal. To do this, the individual data signals each modulate an assigned subcarrier. These subcarriers are arranged so that the channels occupy adjacent frequency bands with some frequency space between them, known as *guard bands*. The modulated subcarriers are used to angle-modulate a high-frequency carrier, as shown in Fig. 6.15.

If FM is used for the subcarrier modulation and for the main carrier modulation, the composite modulation is referred to as FM-FM; if AM is used, then it is referred to as AM-FM. The amplitude-modulation methods used for the subcarrier modulation are DSB-SC or SSB-SC. Large-carrier methods are avoided because too much of the peak frequency deviation would be used merely to send the AM carrier. Usually a pilot subcarrier is also sent for demodulation. Note that the stereo multiplexing used in commercial FM is an example of an AM-FM system.[†]

[†] Applications of both amplitude modulation and angle modulation to proposed stereo AM transmissions are discussed in Appendix H.

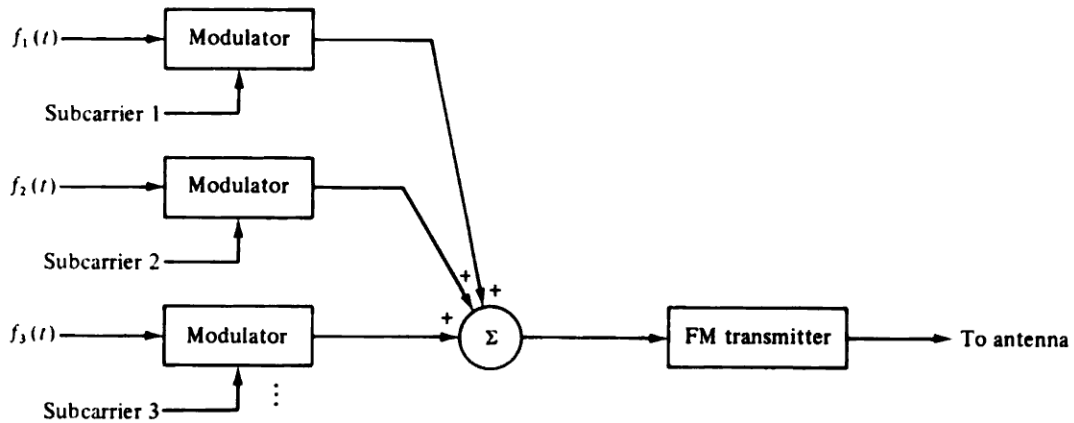


Figure 6.15 A simplified composite modulation system.

Angle modulation (both FM and PM) is widely used in multichannel data transmission and telemetry systems. Standards exist for the assignment of subcarriers and guard bands for the latter.† To allow for realizable filter designs to separate adjacent channels, it is common to allow some frequency separation between channels.

6.7 DEMODULATION OF FM SIGNALS

There are a number of ways to recover the modulating signal from the FM waveform and we shall discuss only some of them. The overall characteristic must be the same, however—to provide an output signal whose amplitude is linearly proportional to the instantaneous frequency of the input waveform.

6.7.1 Direct Method

One method is to use some system that has a linear frequency-to-voltage transfer characteristic. Such a system is called a frequency *discriminator*. In our search for a simple discriminator, we need something with a linear amplitude versus frequency characteristic. The simplest conceptually is that of the ideal differentiator, for we recall that its transfer function is given by $H(\omega) = j\omega$. (Certainly the magnitude characteristic is very linear!)

An expression for the general FM waveform is

$$\phi_{\text{FM}}(t) = A \cos\left[\omega_c t + k_f \int_0^t f(\tau)\, d\tau\right].$$

† See, for example, E. L. Gruenberg (ed.), *Handbook of Telemetry and Remote Control*, New York: McGraw-Hill, 1967.

Assuming that A is a constant (a limiter can be inserted prior to the differentiator to ensure this), we have

$$\frac{d\phi}{dt} = -A[\omega_c + k_f f(t)] \sin\left[\omega_c t + k_f \int_0^t f(\tau)\, d\tau\right]. \tag{6.47}$$

If $k_f f(t) \ll \omega_c$, we see that Eq. (6.47) is in the form of an AM signal whose envelope is

$$A\omega_c\left[1 + \frac{k_f}{\omega_c} f(t)\right] \tag{6.48}$$

and whose frequency is

$$\omega_c + k_f f(t). \tag{6.49}$$

The differentiator has therefore changed FM into AM with only the slight difference that the new carrier frequency has some frequency variation. The resulting AM signal can be detected by an envelope detector and as long as $k_f f(t) \ll \omega_c$, the slight variation in the carrier frequency would not be detectable by the envelope detector.

EXAMPLE 6.7.1

The angle-modulated signal waveform $\phi(t) = A \cos(\omega_c t + \beta \sin \omega_m t)$ is applied to the input of an RC high-pass filter, shown in Fig. 6.16. If $\omega \ll (RC)^{-1}$ in the frequency band occupied by the signal $\phi(t)$, show that the output voltage is amplitude modulated and can be detected by an envelope detector. Also, find an expression for the modulation index of the AM signal in the filter output.

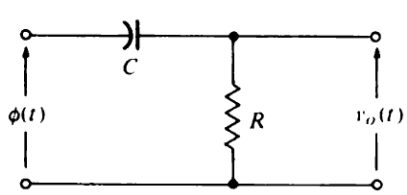


Figure 6.16 The RC high-pass filter.

Solution The frequency transfer function of the high-pass filter is

$$H(\omega) = \frac{R}{R + (j\omega C)^{-1}} = \frac{j\omega RC}{1 + j\omega RC}.$$

For frequencies of interest within the condition $\omega \ll (RC)^{-1}$ this transfer function can be approximated by

$$H(\omega) \approx j\omega RC \quad \text{for } \omega \ll (RC)^{-1},$$

from which the impulse response is

$$h(t) \approx RC\frac{d}{dt}.$$

Using the angle-modulated signal waveform $\phi(t)$ as the input, the output voltage waveform, within these approximations, is

$$v_o(t) \approx RC\frac{d}{dt}[A \cos(\omega_c t + \beta \sin \omega_m t)]$$

$$v_o(t) \approx -ARC\omega_c\left[1 + \frac{\beta\omega_m}{\omega_c}\cos \omega_m t\right]\sin(\omega_c t + \beta \sin \omega_m t).$$

We observe that $v_o(t)$ is an angle-modulated waveform that is itself amplitude-modulated. The fact that the carrier waveform is angle-modulated is not critical to the use of an envelope detector for demodulation of the AM, and the equivalent AM modulation index is given by

$$m = \frac{\beta\omega_m}{\omega_c} = \frac{\Delta\omega}{\omega_c}.$$

The action of the ideal differentiator can be approximated by any device whose magnitude transfer function is reasonably linear within the range of frequencies of interest. Figure 6.17(a) illustrates this type of discriminator. In Fig. 6.17(b) an *RL* circuit approximation to a differentiator is used followed by an envelope detector. A bandpass version of this circuit is shown in Fig. 6.17(c). These discriminators are known as *slope detectors*. Although the slope detector is economical, it has a very limited linear range and its use is restricted to input signals with small frequency variations.

A more linear response can be obtained by taking the difference between two bandpass magnitude responses. The triple-tuned balanced discriminator shown in Fig. 6.17(d) uses this principle and has better sensitivity and linearity. It also provides a zero output response at the carrier frequency, which is an advantage in recording equipment design.

Another possibility is to use a time-delay/phase-shift approximation to the differentiator. It avoids the multiple tuning problems while retaining high sensitivity and good linearity. The most straightforward approach is the time-delay (phase-shift) demodulator shown in block diagram form in Fig. 6.18(a). The time-delay demodulator finds applications in demodulating FM microwave signals.

At frequencies below the microwave region, one can take advantage of the nearly linear phase-shift characteristic of a tuned circuit near resonance. One circuit is the Foster–Seeley discriminator circuit shown in Fig. 6.18(b).[†] In the Foster–Seeley discriminator both tuned circuits are tuned to the carrier frequency and the output voltage varies with the frequency deviation as a result of the phase shifts in the secondary circuit. Circuit operation can best be traced using phasor diagrams; we shall not pursue

[†] D. F. Foster and S. H. Seeley, *Proceedings of the IRE*, vol. 25 (1937): 289.

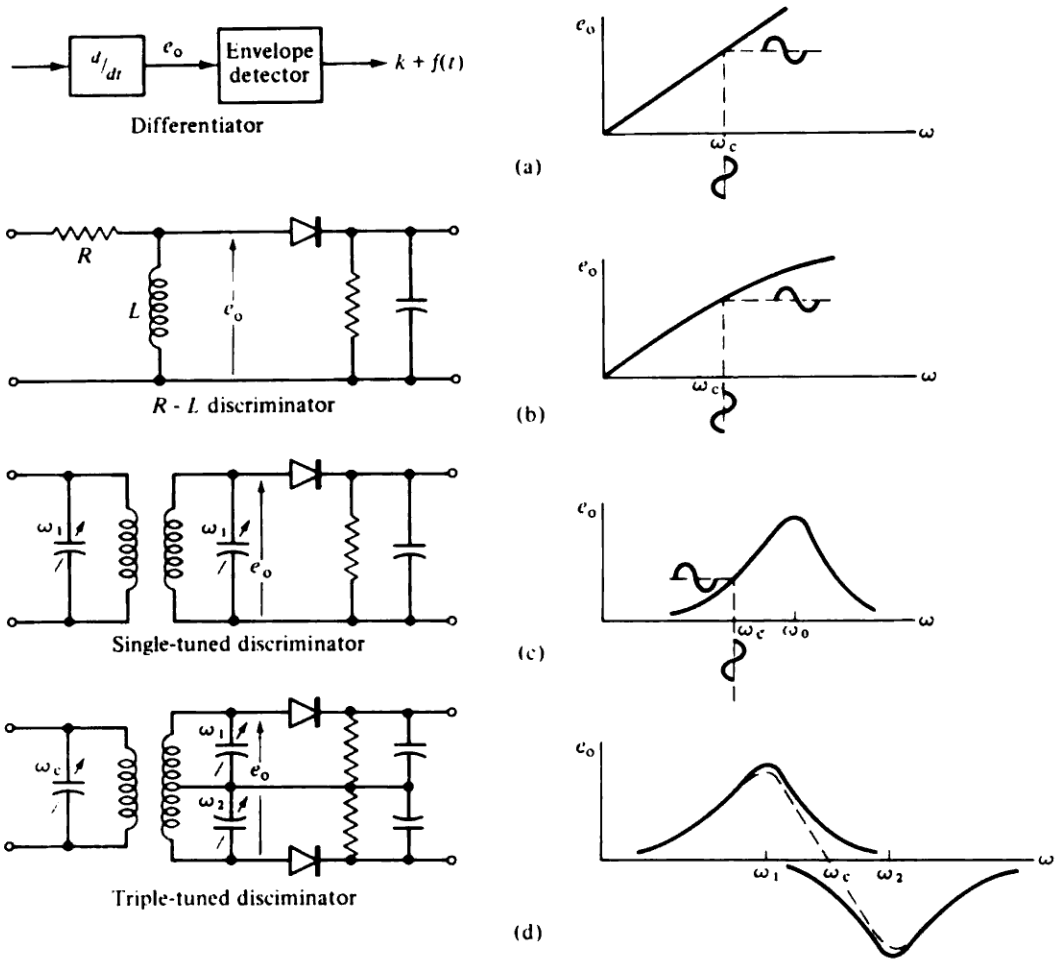


Figure 6.17 FM demodulation using discriminators.

this here.[†] The operation of the Foster–Seeley discriminator is very linear and produces a balanced output. It is amplitude-sensitive and requires prior amplitude limiting.

A more commonly used circuit in commercial receivers that is similar and yet requires no prior limiting is the ratio detector shown in Fig. 6.18(c). In contrast to the Foster–Seeley circuit, the diodes in the ratio detector are connected in series with the tuned secondary circuit and the output is taken across a bridge circuit. A very large capacitor placed across the opposite diagonal of the bridge keeps the total voltage relatively constant, suppressing the effects of amplitude variations. The ratio detector is used in many entertainment-type FM receivers.

Because the information about the modulating signal is contained in the zero crossings of the FM waveform, another approach is to severely clip (limit) the ampli-

[†] See, for example, J. J. DeFrance, *Communications Electronics Circuits*, Second ed., San Francisco: Rinehart Press, 1972, p. 337; or A. B. Cook and A. A. Liff, *Frequency Modulation Receivers*, Englewood Cliffs, N.J.: Prentice-Hall, 1968, Chapter 10.

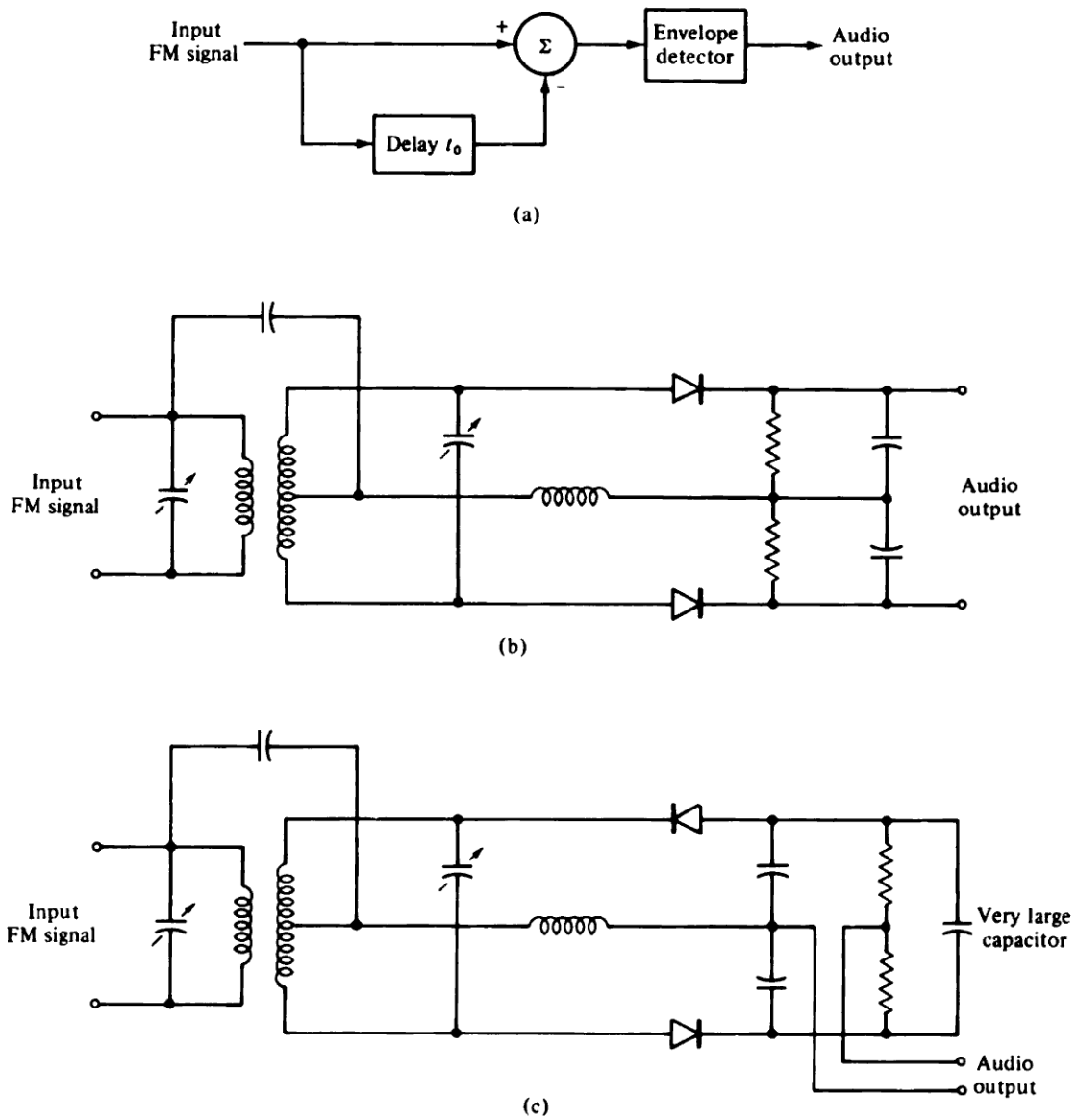


Figure 6.18 Some FM demodulators used in practice: (a) the time-delay demodulator, (b) the Foster–Seeley discriminator, (c) the ratio detector.

tude of the FM waveform and then detect the variation in the zero crossings of the resulting square wave. The spacing of the zero crossings can be measured by counting the number of zero crossings in a given time interval or by measuring the time interval for a given number of zero crossings.

A block diagram of a typical receiver for the commercial FM broadcast band covering 88–108 MHz is shown in Fig. 6.19. The receiver is similar to the AM superheterodyne receiver with the exception of the addition of a discriminator and possibly a limiter. The common choice of an intermediate frequency is 10.7 MHz.

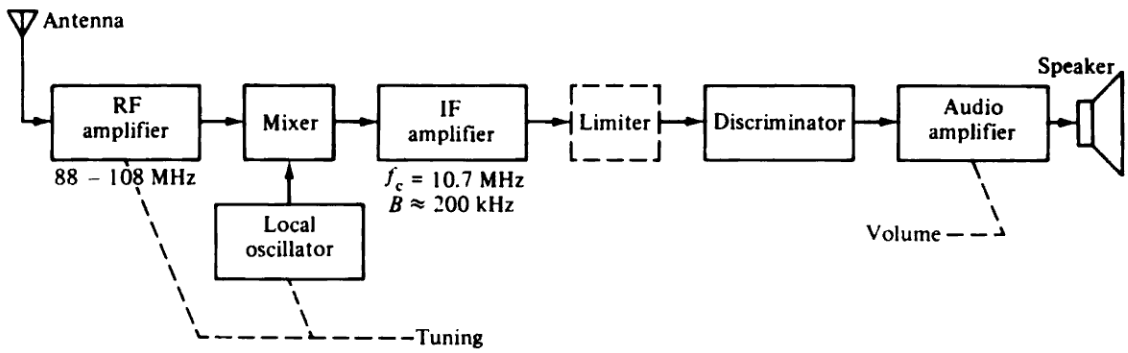


Figure 6.19 A typical FM broadcast receiver.

★ EXAMPLE 6.7.2

Suppose one attempts to convert a standard AM broadcast receiver to receive FM transmissions by only changing the frequency of the local oscillator. By tuning slightly to one side of the (nonideal) IF response, it is reasoned that the IF response and the envelope detector will then act as a slope detector. Choosing a simple parallel *RLC* network to model the IF response, estimate the performance of this receiver.

Solution For a parallel *RLC* network (see Fig. 6.20), we have

$$|H(\omega)| = \left|\frac{V(\omega)}{I(\omega)}\right| = \frac{1/C}{\sqrt{4\alpha^2 + [(\omega_0^2 - \omega^2)/\omega]^2}}$$

where

$$\omega_0 = 1/\sqrt{LC} \quad \text{and} \quad \alpha = 1/(2RC).$$

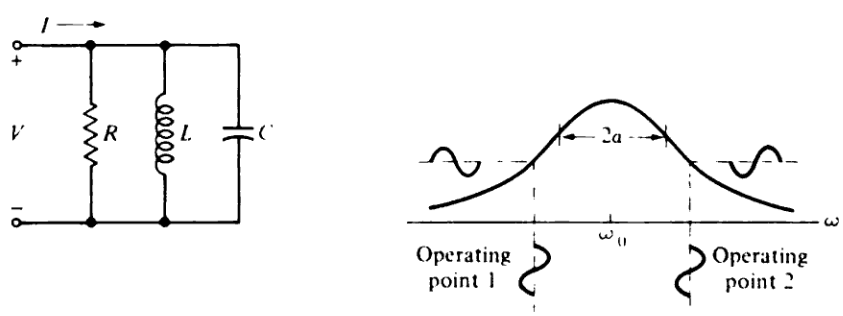


Figure 6.20 FM slope demodulation using a tuned circuit.

The −3-dB bandwidth is 2α; if $2\alpha \ll \omega_0$ (high-Q case), we can simplify things a little by translating to an equivalent low-pass case (that is, $\omega_0 = 0$) so that

$$|H_1(\omega)| = \frac{1/C}{\sqrt{4\alpha^2 + \omega^2}} = \frac{R}{\sqrt{1 + (\omega/2\alpha)^2}}.$$

A reasonable choice of operating point for maximum linear operating range is at an inflection point (see Fig. 6.20); i.e., the point where

$$\frac{d^2}{d\omega^2}|H_1(\omega)| = 0 .$$

Solutions to this are at $\omega = \pm\sqrt{2}\,\alpha$.

Next we expand the tuning characteristic about the operating point. Taking the upper choice of operating points, let $\Delta\omega = \omega - \sqrt{2}\,\alpha$ and

$$|H_1(\omega)| = \frac{R}{\sqrt{1 + [(\Delta\omega + \sqrt{2}\,\alpha)/2\alpha]^2}}$$

If we assume that $\Delta\omega \ll \alpha$, this can be approximated using the binomial expansion so that

$$|H_1(\omega)| \approx \sqrt{2/3}\,R[1 - 1/(3\sqrt{2}\,\alpha)\Delta\omega], \qquad \Delta\omega \ll \alpha .$$

Thus, within these assumptions, we have verified that $|H(\omega)|$ does have a linear slope component. Continuing our example, the bandwidth of this circuit for AM is on the order of 10 kHz so that $\alpha/2\pi = 5$ kHz and $\sqrt{2}\,\alpha/2\pi = 7$ kHz. Therefore $\Delta\omega$ should not exceed 0.7 kHz for our linearity assumption to hold and we conclude that, although this scheme will work, it is limited to very narrowband inputs. This approach is actually used in inexpensive converters to allow one to use an AM receiver to listen to police calls transmitted via narrowband FM.

6.7.2 Indirect Method—the Phase-Locked Loop

All of the preceding methods have some similarities and the choice of which one to use is governed by the required linearity, zero balance, ease of alignment, amplitude sensitivity, reliability, and economic factors. A different approach is to place a frequency modulator in the return branch of a feedback system. A feedback system with sufficient loop gain performs in its forward branch the inverse operation of that which is performed in its return branch and thus demodulates the signal. Among the detectors in this group are the FM demodulator with feedback (FMFB) and the phase-locked loop (PLL). Block diagrams of the FMFB and PLL demodulation systems are shown in Fig. 6.21. We shall center our discussion on the PLL. Its ease of alignment and efficient operation in the presence of noise makes the PLL increasingly more popular. The reader is referred to more advanced texts for details on the operation of the FMFB system.

A block diagram of the PLL is shown in Fig. 6.21(b). Both inputs to the phase comparator are assumed to be periodic and to have the same fundamental frequency. The phase comparator detects the timing difference between the two signals and produces an output voltage which is proportional to this difference. For sinusoidal

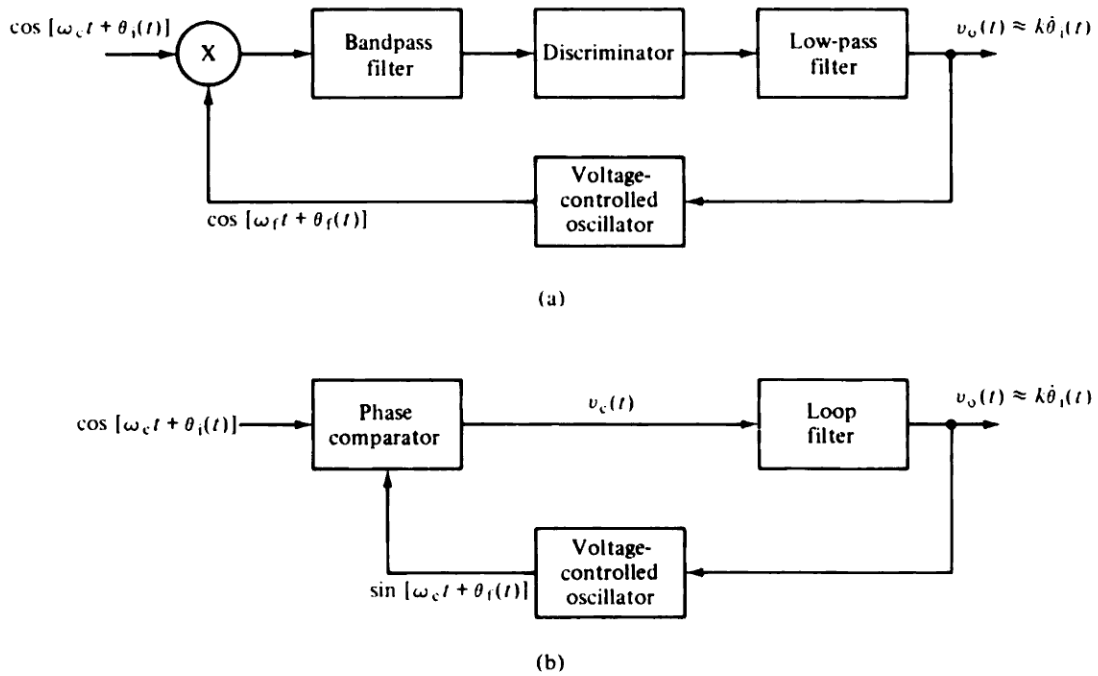


Figure 6.21 Block diagram of (a) the FM demodulator with feedback (FMFB) and (b) the phase-locked loop (PLL).

inputs, this timing difference can also be expressed as a phase difference. One way to construct a phase comparator is to use a multiplier and low-pass filter, as shown in Fig. 5.13. The purpose of the low-pass filter in the phase comparator of Fig. 5.13 is to attenuate the second-harmonic frequency components that arise in the multiplication. We have added a loop filter in Fig. 6.21(b) whose function will be explained shortly.

In general, the output of the phase comparator is proportional to the average value of the product of the two inputs as a function of their relative time displacement. If the two inputs are $x(t)$, $y(t)$, this can be expressed as

$$\frac{1}{T}\int_0^T x(t)y(t+\tau)\,dt = R_{xy}(\tau) \tag{6.50}$$

where T is the period of the input waveform. Therefore the phase-comparator output is proportional to the crosscorrelation between the input waveforms. If both inputs are sinusoids, that is,

$$x(t) = A_1 \cos \omega_c t\,, \tag{6.51}$$

$$y(t) = A_2 \cos \omega_c t\,, \tag{6.52}$$

then Eq. (6.50) becomes (cf. Example 4.4.2)

$$R_{xy}(\tau) = \tfrac{1}{2}A_1 A_2 \cos \omega_c \tau\,. \tag{6.53}$$

Because the inputs are sinusoidal, this can also be expressed in terms of the phase $\theta = \omega_c \tau$:

$$R_{xy}(\theta) = \tfrac{1}{2} A_1 A_2 \cos \theta . \tag{6.54}$$

This is shown in Fig. 6.22(a).

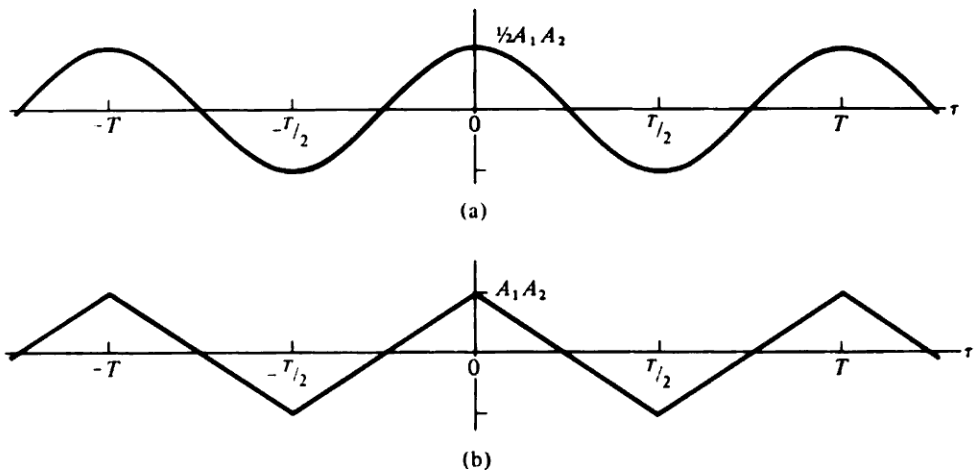


Figure 6.22 Phase comparator characteristic for (a) two sinusoidal inputs and (b) two square-wave inputs.

We could also use square waves for both input signal waveforms. For example, the input could be a sinusoid that is amplitude-limited and the VCO could be designed for a square-wave output. In this case, the phase-comparator characteristic is a triangular characteristic (cf. Example 4.4.1), as shown in Fig. 6.22(b) (zero average value with peak values A_1 and A_2 assumed). Other cases (e.g., a sinusoid and a square wave) can be worked out using Eq. (6.50).

From the phase-comparator characteristics shown in Fig. 6.22, we observe that the output is zero at odd multiples of $\pm T/4$; these points are called *nulls*. The nulls form the desired operating points for our purposes and the remainder of the PLL is designed to keep the timing difference between input waveforms as near an operating point as possible. Note that the phase-comparator output goes through a null with a positive slope for $\tau = -T/4 \pm T$ ($\theta = -\pi/2 \pm 2\pi$ for sinusoids) and with a negative slope for $\tau = T/4 \pm T$ ($\theta = \pi/2 \pm 2\pi$ for sinusoids).

Now suppose that the loop is operating at the null point $\tau = T/4 \pm T$ and that the input is delayed by an increment $\Delta\tau$. This results in a small negative voltage out of the phase comparator (cf. Fig. 6.22). If the VCO frequency increases with input voltage and the gain in the feedback loop is noninverting, the loop will produce an incremental delay of opposite sign to the input which will tend to keep the loop near the desired operating point. On the other hand, if the feedback gain is inverting, the loop will produce an incremental delay which will add to $\Delta\tau$ and will drive the loop away from the operating point. Thus this null is a stable operating point for a noninverting

gain and an unstable operating point for an inverting gain. In contrast, the slope is positive at the null $\tau = -T/4 \pm T$, and thus this point is a stable operating point for an inverting loop gain and an unstable operating point for a noninverting loop gain. Note that in either case the relationships are such that the timing difference between the two waveforms is $\pm T/4$ (that is, $\pm\pi/2$ for sinusoids).

Both phase-comparator characteristics shown in Fig. 6.22 have linear slopes near the nulls. The square-wave characteristic has some advantage in that the slope is a constant over one-half a period whereas the slope decreases away from the null points for the sinusoidal case. This latter effect results in a nonlinear loop-gain factor.

Although the phase-comparator characteristic depends on the timing difference between inputs, it is convenient to use a phase designation such that $\theta_i(t)$ is the input phase and $\theta_f(t)$ is the feedback phase. The timing difference is $[\theta_i(t) - \theta_f(t)]$ and the phase-comparator output voltage $v_c(t)$ about an operating point is then

$$v_c(t) = k_c[\theta_i(t) - \theta_f(t)]. \tag{6.55}$$

The proportionality constant k_c in Eq. (6.55) is the phase-comparator gain factor and has the dimension of volt per radian. From Fig. 6.22 we observe that k_c is proportional to the product of the amplitudes of the two input waveforms.

For the sinusoidal case, the two inputs must be 90° out of phase at a null. Using a multiplier and low-pass filter, the phase-comparator output is then

$$\begin{aligned} v_c(t) &\propto A_1 \cos[\omega_c t + \theta_i(t)] A_2 \sin[\omega_c t + \theta_f(t)]_{LP} \\ &\propto \tfrac{1}{2} A_1 A_2 \sin[\theta_i(t) - \theta_f(t)], \end{aligned}$$

or

$$v_c(t) = k_c \sin[\theta_i(t) - \theta_f(t)]. \tag{6.56}$$

Equation (6.55) follows from Eq. (6.56) if $[\theta_i(t) - \theta_f(t)]$ is small. Of course, Eq. (6.55) must be restricted to an operating range $(-T/4, T/4)$ even for the square-wave case. The use of Eq. (6.55) in the analysis of the PLL is referred to as a "linearized" PLL.

The loop filter controls the dynamic response of the PLL. It has been separated from the low-pass filter of the phase comparator to emphasize the different roles of the two filters. The purpose of the low-pass filter in the phase comparator is to suppress second-harmonic terms while the loop filter controls the dynamic response of the loop. Designating the impulse response of the loop filter by $h(t)$, the output voltage $v_o(t)$ is

$$v_o(t) = v_c(t) \circledast h(t). \tag{6.57}$$

The voltage-controlled oscillator (VCO) generates a constant-amplitude periodic waveform (e.g., a sinusoid, square-wave, etc.) whose fundamental frequency is proportional to the input voltage. Letting ω_f be the instantaneous frequency with respect to the fundamental frequency, we have

$$\omega_f(t) = k_f v_o(t). \tag{6.58}$$

where the voltage $v_o(t)$ is the output of the loop filter (cf. Fig. 6.21b). The constant k_f has the dimensions of radians per second per volt [cf. Eq. (6.5)]. The phase of the VCO is then, within an arbitrary constant,

$$\theta_f(t) = k_f \int_0^t v_o(\xi)\, d\xi\,. \tag{6.59}$$

Equations (6.55) [(6.56) for the nonlinear case], (6.57), and (6.59) describe the dynamic response of the PLL. The controlling parameters are the loop-gain factors k_c, k_f and the parameters of the loop filter.

Taking the derivative of Eq. (6.59), we get

$$v_o(t) = \frac{1}{k_f}\frac{d\theta_f}{dt}\,. \tag{6.60}$$

If the loop gain is high, $[\theta_i(t) - \theta_f(t)]$ is small, $\theta_f(t) \approx \theta_i(t)$ and Eq. (6.60) states that the output voltage is proportional to the instantaneous frequency of the input referred to the carrier. The PLL therefore accomplishes the desired demodulation of the input FM signal.

We have assumed that the input and the VCO are at the same frequency. Initially, this may not be true and the loop must go through a mode of acquisition to gain lock. If the frequency difference between the input and the VCO is less than the loop bandwidth (i.e., the closed-loop bandwidth of the PLL; this will be discussed in Section 6.7.3) the PLL will lock quickly. This is the usual case for FM detection.

If the range of expected frequency differences is much larger than the loop bandwidth, the VCO frequency can be swept at a suitable rate to search for the signal. In the absence of noise, this has similarities to the scan rate of a scanning spectrum analyzer for which (cf. Problem 3.13.3) the usual restriction is that

$$\dot{f} < B^2, \tag{6.61}$$

where $\dot{f}$ is the frequency scan rate and B is the bandwidth. Equation (6.61) is often used as an upper bound for the sweep rate of the PLL, but it must be lowered considerably in the presence of noise.

The frequency sweep can be obtained by applying a ramp voltage to the input of the VCO. The ramp is reset when a preset voltage level is attained and the sweep is then reinitiated. If the frequency sweep is slow enough, the loop will lock when the input and VCO frequencies coincide. Although the sweep voltage can often be left on and overcome by the locked loop, it is advantageous to have a lock indicator which automatically disengages the sweep when lock is attained.

A lock indicator widely used is the quadrature phase comparator shown in Fig. 6.23. In the unlocked condition, the outputs of the phase comparators have average values near zero. In the locked condition, the main phase comparator has an output proportional to $\sin(\theta_i - \theta_f)$ (for sinusoidal inputs) while the quadrature output is proportional to $\cos(\theta_i - \theta_f)$. Since $(\theta_i - \theta_f)$ is small in the locked condition, $\cos(\theta_i - \theta_f) \approx 1$. The locked condition can be sensed by a threshold circuit which then disconnects the frequency sweep and connects the PLL for normal operation. Without the sweep

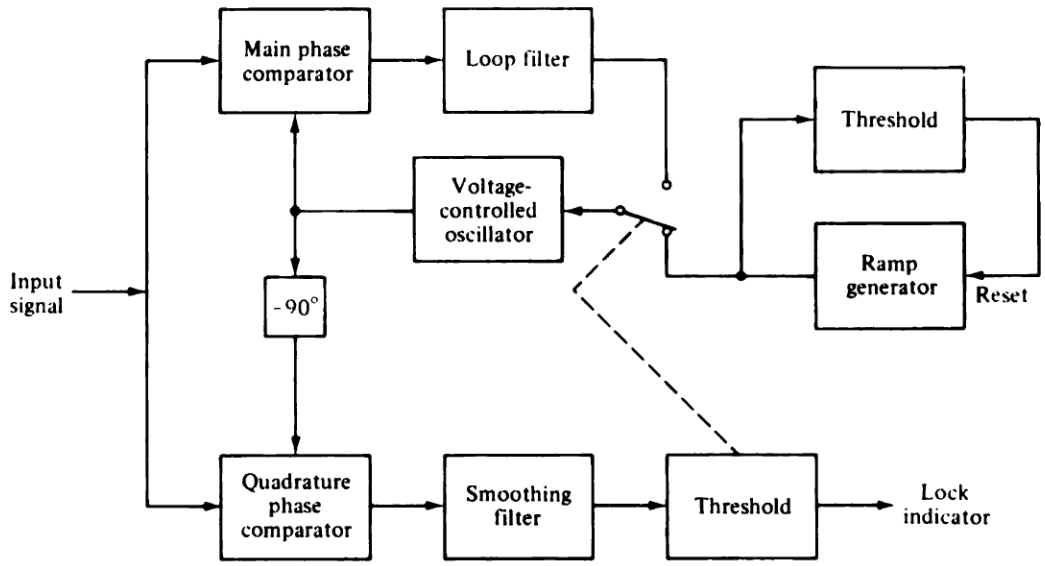


Figure 6.23 PLL sweep acquisition.

acquisition circuitry, the lock indicator may be used to signal the lock condition (e.g., the stereo indicator in stereo FM detection).

EXAMPLE 6.7.3

(a) Derive an integro-differential equation for $\theta_f(t)$ in terms of $\theta_i(t)$ for sinusoidal PLL inputs. (b) Derive a differential equation for the case of no loop filter and small phase errors.

Solution a) Differentiating Eq. (6.59) and substituting Eqs. (6.56) and (6.57), we get

$$\frac{d\theta_f}{dt} = k_c k_f \sin\left[\theta_i(t) - \theta_f(t)\right] \circledast h(t).$$

b) For $h(t) = 1$ and $\sin\left[\theta_i(t) - \theta_f(t)\right] \approx \left[\theta_i(t) - \theta_f(t)\right]$, we get the first-order linear differential equation,

$$\frac{d\theta_f}{dt} + k_c k_f \theta_f(t) = k_c k_f \theta_i(t).$$

Thus the linearized PLL for no loop filter is a first-order system with a time constant of $(k_c k_f)^{-1}$.

DRILL PROBLEM 6.7.1

(a) Rewrite the linear differential equation in Example 6.7.3(b) in terms of frequency.
(b) Find ω_f for $t > 0$ if $\omega_i = u(t)$.

ANSWER: (a) $d\omega_f/dt + k_c k_f \omega_f(t) = k_c k_f \omega_i(t)$; (b) $1 - \exp(-k_c k_f t)$.

6.7.3 ★ The Linearized Phase-Locked Loop†

To proceed farther into an analysis of the PLL, we linearize the phase-comparator characteristic and apply the Laplace transform to Eqs. (6.55), (6.57), and (6.59) to obtain, respectively,

$$V_c(s) = k_c[\theta_i(s) - \theta_f(s)], \tag{6.62}$$

$$V_o(s) = V_c(s)H(s), \tag{6.63}$$

$$\theta_f(s) = \frac{k_f}{s} V_o(s). \tag{6.64}$$

Combining Eqs. (6.62)–(6.64), we obtain the linearized PLL model shown in Fig. 6.24. Thus we can write

$$\theta_f(s) = k_c[\theta_i(s) - \theta_f(s)]H(s)k_f/s, \tag{6.65}$$

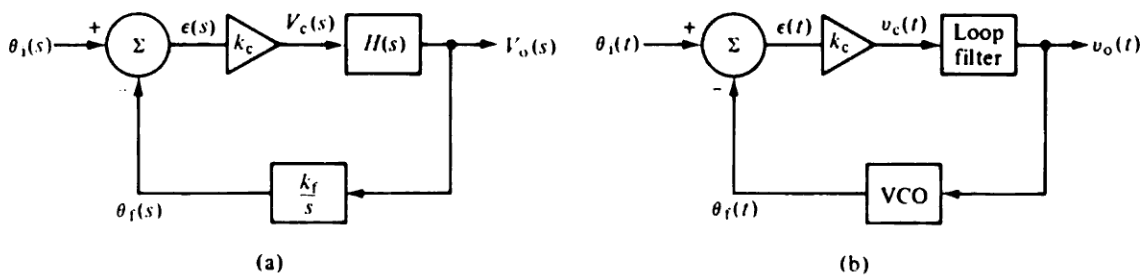


Figure 6.24 The linearized PLL model in (a) frequency and (b) time.

or

$$\frac{\theta_f(s)}{\theta_i(s)} = \frac{k_c k_f H(s)}{s + k_c k_f H(s)}. \tag{6.66}$$

Equation (6.66) is the basic result for the linearized PLL.

† This section requires some knowledge of the Laplace transform.

Two quantities of interest to us are the loop phase error $\epsilon(t)$ and the input-output transfer function for phase and frequency demodulation. The former can be found from Eq. (6.66) as follows (cf. Fig. 6.24a):

$$\frac{\epsilon(s)}{\theta_i(s)} = \frac{\theta_i(s) - \theta_f(s)}{\theta_i(s)} = 1 - \frac{\theta_f(s)}{\theta_i(s)}$$

$$= \frac{s}{s + k_c k_f H(s)}. \tag{6.67}$$

The latter can be found from (cf. Fig. 6.24a)

$$V_o(s) = k_c \epsilon(s) H(s),$$

or, using Eq. (6.67),

$$\frac{V_o(s)}{\theta_i(s)} = \frac{k_c s H(s)}{s + k_c k_f H(s)}. \tag{6.68}$$

We have thus expressed the desired output voltage $V_o(s)$ and the loop phase error $\epsilon(s)$ in terms of the input phase $\theta_i(s)$ and the loop filter $H(s)$. To proceed farther requires some choice of the loop-filter transfer function $H(s)$. We examine two cases.

The First-Order PLL

The simplest case is to let

$$H(s) = 1 \tag{6.69}$$

so that Eq. (6.68) becomes

$$\frac{V_o(s)}{\theta_i(s)} = \frac{k_c s}{s + k_c k_f}. \tag{6.70}$$

Because we are interested in the frequency-demodulation properties of the PLL, we apply a unit impulse in *frequency*. Then we have

$$\theta_i(s) = 1/s, \tag{6.71}$$

and Eq. (6.70) gives

$$V_o(s) = \frac{k_c}{s + k_c k_f}. \tag{6.72}$$

Because the PLL behaves like a first-order system, it is called a *first-order loop*. The −3-dB frequency is

$$\omega_{-3\,\mathrm{dB}} = k_c k_f, \tag{6.73}$$

and therefore, to avoid attenuation of high-frequency content, we require that the modulation bandwidth W satisfy the restriction

$$W < k_c k_f . \tag{6.74}$$

For a step change, ω, in the input frequency, we have

$$\theta_i(s) = \omega/s^2, \tag{6.75}$$

and Eq. (6.70) gives

$$V_o(s) = \frac{k_c \omega}{s(s + k_c k_f)}, \tag{6.76}$$

or

$$v_o(t) = \frac{\omega}{k_f}[1 - e^{-k_c k_f t}]u(t) . \tag{6.77}$$

The output voltage therefore attempts to follow the frequency step and, after the transients go to zero, it produces the desired voltage [cf. Eq. (6.60)]

$$v_o(\infty) = \frac{\omega}{k_f} . \tag{6.78}$$

The loop phase error for this step in frequency is found from Eqs. (6.67), (6.69), and (6.75):

$$\epsilon(s) = \frac{\omega}{s(s + k_c k_f)}, \tag{6.79}$$

or,

$$\epsilon(t) = \frac{\omega}{k_c k_f}[1 - e^{-k_c k_f t}]u(t) . \tag{6.80}$$

The steady-state phase error is then

$$\epsilon(\infty) = \frac{\omega}{k_c k_f} . \tag{6.81}$$

This assumes, of course, that $|\epsilon(t)| < \pi/2$ for the square-wave inputs or that $|\epsilon(t)| \ll \pi/2$ for the sinusoidal inputs to remain within the linear range.

The product $k_c k_f$ is the loop gain. A high loop gain is desirable for small phase errors [see Eq. (6.81)]. However, this product is also the loop bandwidth [cf. Eq. (6.73)] so that a high loop gain implies a large bandwidth. As a result of these restrictions, the first-order PLL is seldom used.

The Second-Order PLL

The loop-filter transfer function often used (or approximated) in the second-order PLL is[†]

$$H(s) = \frac{s\tau_2 + 1}{s\tau_1}, \tag{6.82}$$

where τ_1, τ_2 are adjustable parameters. Using Eq. (6.82) in Eq. (6.68), we get

$$\frac{V_o(s)}{\theta_i(s)} = \frac{k_c s(s\tau_2 + 1)}{\tau_1(s^2 + 2\zeta\omega_n s + \omega_n^2)} \tag{6.83}$$

where the *natural frequency* ω_n and the *damping factor* ζ of the loop are parameters normally used in the theory of second-order systems and are given here by

$$\omega_n = \sqrt{k_c k_f/\tau_1}, \tag{6.84}$$

$$\zeta = \omega_n \tau_2/2. \tag{6.85}$$

For a unit impulse in frequency, we use Eq. (6.71) in Eq. (6.83):

$$V_o(s) = \frac{k_c(s\tau_2 + 1)}{\tau_1(s^2 + 2\zeta\omega_n s + \omega_n^2)}, \tag{6.86}$$

which exhibits the second-order behavior of the PLL. Usually the PLL is operated with slightly less damping than "critical damping" ($\zeta = 1$) to obtain a fast response. Therefore we are primarily interested in the range $\frac{1}{2} < \zeta < 1$($\zeta = 1/\sqrt{2} = 0.707$ is a popular choice). Within this range, the -3-dB bandwidth of Eq. (6.86) is approximately

$$\omega_{-3\,\mathrm{dB}} \approx (1 + \sqrt{2}\,\zeta)\omega_n \quad \text{for} \quad \tfrac{1}{2} < \zeta < 1. \tag{6.87}$$

A criterion often used for the modulating bandwidth W is the slightly more pessimistic choice,

$$W < 2\omega_n\zeta. \tag{6.88}$$

For a step change in frequency, Eqs. (6.75) and (6.83) give

$$V_o(s) = \frac{k_c\omega}{\tau_1}\frac{s\tau_2 + 1}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)}, \tag{6.89}$$

whose time solution, for $\zeta^2 < 1$, is

$$v_o(t) = \frac{\omega}{k_f}\left[1 - \frac{1}{\sqrt{1-\zeta^2}}e^{-\zeta\omega_n t}\cos(\omega_n\sqrt{1-\zeta^2}\,t + \phi)\right]u(t), \tag{6.90}$$

[†] This is known as a *lead-lag* filter in the theory of feedback systems.

where $\phi = \tan^{-1}(\zeta/\sqrt{1-\zeta^2})$. After all transients have disappeared, Eq. (6.90) gives†

$$v_o(\infty) = \omega/k_f. \tag{6.91}$$

The loop phase error is found using Eqs. (6.67) and (6.82):

$$\frac{\epsilon(s)}{\theta_i(s)} = \frac{s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{6.92}$$

As in the case of the first-order loop, this exhibits a high-pass characteristic and shows that the loop cannot keep up with rapid phase changes. Applying a frequency step, we use Eq. (6.75) in Eq. (6.92) to give

$$\epsilon(s) = \frac{\omega}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \tag{6.93}$$

whose time solution, for $\zeta^2 < 1$, is

$$\epsilon(t) = \frac{\omega}{\omega_n\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \sin(\omega_n\sqrt{1-\zeta^2}\,t)u(t). \tag{6.94}$$

After all transients have disappeared, Eq. (6.94) gives

$$\epsilon(\infty) = 0. \tag{6.95}$$

In contrast to the first-order loop, the second-order PLL tracks out the static phase error resulting from a step change in frequency and returns the phase-comparator characteristic to the null point. This, together with the added versatility in setting the loop dynamic response, is why most PLL systems are operated as second-order systems.

The peak phase error is found by setting $d\epsilon/dt = 0$ and solving to give

$$\epsilon_p = \frac{\omega}{\zeta\omega_n\chi} \exp\left[-\frac{\tan^{-1}\chi}{\chi}\right] \sin(\tan^{-1}\chi), \tag{6.96}$$

where

$$\chi = \sqrt{1-\zeta^2}/\zeta.$$

For a given value of ζ, Eq. (6.96) sets the maximum phase error for a step-frequency change.

Now let us investigate the sinusoidal case. Assume that the frequency modulation is sinusoidal with frequency ω_m and a peak frequency deviation $\Delta\omega$. The peak phase deviation is $\Delta\omega/\omega_m = \beta$ and from Eq. (6.92) the magnitude of the phase error is

$$|\epsilon(\omega)| = \frac{\beta\omega^2}{\sqrt{(\omega_n^2 - \omega^2)^2 + 4\zeta^2\omega_n^2\omega^2}} \qquad \text{(sinusoidal case)}. \tag{6.97}$$

† The final-value theorem of the Laplace transform may be used with Eqs. (6.89) and (6.84), and yields this result for general ζ.

We let $\omega = \omega_m$ and keep $\Delta\omega$ fixed. Setting $d\epsilon/d\omega_m = 0$ and solving, we find that the maximum error occurs for $\omega_m = \omega_n$. Using this value in Eq. (6.97), we get

$$\epsilon_p = \left(\frac{\Delta\omega}{\omega_n}\right)\frac{1}{2\zeta} \qquad \text{(sinusoidal case)}. \tag{6.98}$$

For no distortion in demodulating FM, this peak phase must be kept within the linear range of the phase-comparator characteristic.

The second-order PLL has the two parameters ω_n, ζ; these can be controlled by the choices of τ_2 and $k_c k_f/\tau_1$. The second-order PLL provides good performance and is widely used. Only in very special applications does the required complexity warrant a higher-order system.

EXAMPLE 6.7.4

Determine the design parameters of a minimum-bandwidth second-order PLL to demodulate commercial FM. Base your design on a sinusoidal FM signal with $\Delta f = 75$ kHz, $0 \le f_m \le 75$ kHz. Assume a triangular phase-comparator characteristic, $\zeta = 0.707$ and $\tau_1 = \tau_2$.

Solution The PLL must both demodulate the signal with uniform gain and stay within the linear phase-comparator characteristic. From Eq. (6.88), we have

$$f_n > f_m/(2\zeta) = 53 \text{ kHz}.$$

From Eq. (6.98), the peak phase error for $\zeta = 0.707$ is

$$\epsilon_p = \frac{1}{\sqrt{2}}\left(\frac{\Delta\omega}{\omega_n}\right) < \frac{\pi}{2},$$

or

$$f_n > \sqrt{2}\,\Delta f/\pi = 34 \text{ kHz}.$$

Therefore we choose $f_n = 53$ kHz; from Eq. (6.85),

$$\tau_2 = 2\zeta/\omega_n = 4.24\ \mu\text{sec}.$$

Using Eq. (6.84) and noting that $\tau_1 = \tau_2$, we find

$$k_c k_f = \frac{2}{\omega_n}\zeta\omega_n^2 = 4.71 \times 10^5/\text{sec}.$$

DRILL PROBLEM 6.7.2

A frequency step of 25 kHz is applied to the PLL of Example 6.7.4. If $k_c = 1$ V/rad, determine (a) the peak phase error, and (b) the steady-state voltage output.

ANSWER: (a) 12.3°; (b) 0.333 V.

6.8 SIGNAL-TO-NOISE RATIOS IN FM RECEPTION

Our purpose here is to examine the effects of band-limited white noise on the frequency demodulation process. The block diagram of an idealized FM receiver is shown in Fig. 6.25. We define the signal-to-noise ratio to be the ratio of the mean signal power without noise to the mean noise power in the presence of an unmodulated carrier.†

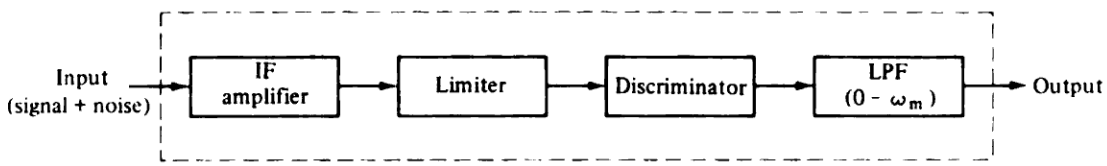


Figure 6.25 The idealized FM receiver.

Assuming that the limiter is ideal and removes all amplitude variations, we can write the FM signal at the discriminator input as

$$s_i(t) = A \cos \theta(t),$$

$$s_i(t) = A \cos \left[\omega_c t + k_f \int_0^t f(\tau)\, d\tau\right]. \tag{6.99}$$

The discriminator output is proportional to the difference between the instantaneous frequency of $s_i(t)$ and the carrier frequency. Letting the constant of proportionality be 1 for convenience,

$$s_o(t) = \left(\frac{d\theta}{dt} - \omega_c\right) = k_f f(t). \tag{6.100}$$

The mean-square value of the output signal is

$$S_o = \overline{s_o^2(t)} = k_f^2 \overline{f^2(t)}. \tag{6.101}$$

Next we turn to calculating the mean output noise power in the presence of an unmodulated carrier. Using the bandpass representation for the band-limited noise [cf. Eq. (5.59)], we can write (see Fig. 6.26),

$$\begin{aligned} A \cos \omega_c t + n_i(t) &= A \cos \omega_c t + n_c(t) \cos \omega_c t - n_s(t) \sin \omega_c t \\ &= r(t) \cos[\omega_c t - \gamma(t)]. \end{aligned} \tag{6.102}$$

Therefore the addition of noise introduces both amplitude noise [in $r(t)$] and phase noise [in $\gamma(t)$]. In the AM case we were interested in the effects of $r(t)$ but in the FM case we can assume amplitude limiting and are interested only in $\gamma(t)$.

† In other words, we assume that the output noise power can be calculated independently of the modulating signal power. A justification for this choice for wideband FM in the presence of white noise is given in B. P. Lathi, *Communication Systems*, New York: John Wiley & Sons, 1968, p. 363.

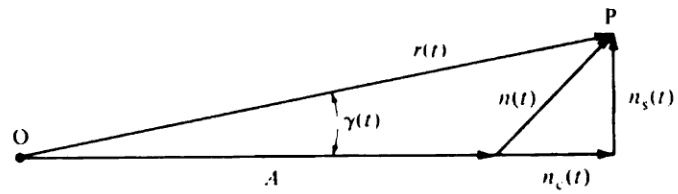


Figure 6.26 Phasor diagram of the sum of a sine wave and narrowband noise, large S/N case.

The principal value of the phase noise angle is

$$\gamma(t) = \tan^{-1}\left[\frac{n_s(t)}{A + n_c(t)}\right]. \tag{6.103}$$

Assuming that the noise is small, i.e., that $n_c(t)$, $n_s(t) \ll A$, we have

$$\gamma(t) \approx \tan^{-1}\left[\frac{n_s(t)}{A}\right] \approx \frac{n_s(t)}{A}. \tag{6.104}$$

The discriminator output is proportional to the difference between the instantaneous frequency and the carrier frequency so that

$$n_o(t) = \frac{d\gamma}{dt} = \frac{1}{A}\frac{d}{dt}[n_s(t)]. \tag{6.105}$$

The power spectral density corresponding to Eq. (6.105) is

$$S_{n_o}(\omega) = \frac{1}{A^2} S_{n_s}(\omega)\,|H(\omega)|^2, \tag{6.106}$$

where $\mathcal{F}^{-1}\{H(\omega)\} = h(t)$ is a time differentiator. But the Fourier transform equivalent of time differentiation is multiplication by $(j\omega)$ so that

$$S_{n_o}(\omega) = \frac{1}{A^2}\omega^2 S_{n_s}(\omega). \tag{6.107}$$

Thus those spectral components at the higher frequencies are emphasized—a general consequence of a differentiator.

The bandwidth of the discriminator output is limited by a low-pass filter with a cutoff frequency of ω_m radians per second. Using Eq. (5.65), we can express the power spectral density within this bandwidth by

$$S_{n_s}(\omega) = [S_n(\omega - \omega_c) + S_n(\omega + \omega_c)]_{LP}. \tag{6.108}$$

If the noise at the discriminator input is white,

$$S_n(\omega) = \eta/2.$$

Eq. (6.108) reduces to

$$S_{n_s}(\omega) = \eta.$$

and Eq. (6.107) becomes

$$S_{n_o}(\omega) = \eta\omega^2/A^2. \tag{6.109}$$

This means that if one assumes a white noise spectral density for the discriminator input, the output noise spectral density will be parabolic (see Fig. 6.27).

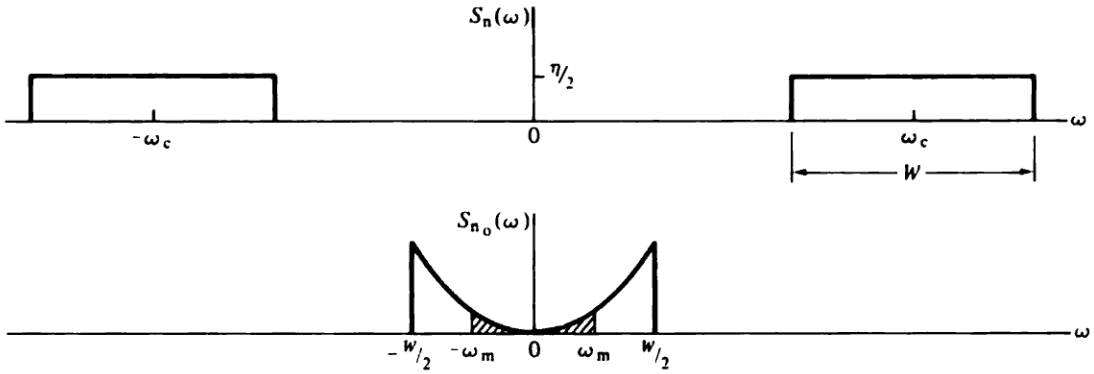


Figure 6.27 Input-output noise spectral densities for the FM demodulator under high carrier-to-noise conditions.

Using Eq. (6.109), we find that the mean-square value of the output noise is

$$N_o = \overline{n_o^2(t)} = \frac{\eta}{\pi A^2}\int_0^{\omega_m} \omega^2\, d\omega\,,$$

$$N_o = \frac{\eta\omega_m^3}{3\pi A^2}. \tag{6.110}$$

The mean carrier power is

$$S_c = A^2/2. \tag{6.111}$$

Comparing Eqs. (6.110) and (6.111), we see that the output noise power is inversely proportional to the mean carrier power in FM. This effect of a decrease in output noise power as the carrier power increases is called *noise quieting*.

Combining Eqs. (6.101) and (6.110), we have

$$\frac{S_o}{N_o} = \frac{3\pi A^2 k_f^2 \overline{f^2(t)}}{\eta\omega_m^3}. \tag{6.112}$$

The peak frequency deviation is proportional to k_f; for the wideband case, the bandwidth increases proportional to k_f. *For wideband FM, we conclude that the output signal-to-noise ratio increases as the square of the bandwidth.* In particular, if $f(t)$ is sinusoidal [that is, $f(t) = a\cos\omega_m t$ and $\Delta\omega = ak_f$], Eq. (6.112) becomes

$$\frac{S_o}{N_o} = \frac{3\pi A^2(\Delta\omega)^2}{2\eta\omega_m^3} = \frac{3\pi A^2\beta^2}{2\eta\omega_m}. \tag{6.113}$$

For a comparison between FM and AM, let the modulation be sinusoidal in each case and let us define the mean-square noise at the IF for the AM case by

$$N_c = \frac{1}{\pi}\int_{\omega_c-\omega_m}^{\omega_c+\omega_m} \frac{\eta}{2}\,d\omega = \frac{\eta\omega_m}{\pi}. \tag{6.114}$$

Using Eqs. (6.111) and (6.114) and noting that $\beta = \Delta\omega/\omega_m$, Eq. (6.113) can be rewritten in terms of the carrier-to-noise ratio

$$\frac{S_o}{N_o} = 3\beta^2\frac{S_c}{N_c}. \tag{6.115}$$

Note that the bandwidth used for the computation of N_c in this result is that for an AM system. Under the most favorable conditions in AM the modulation index is 100% and for these conditions [cf. Eq. (5.90)],

$$\left(\frac{S_o}{N_o}\right)_{AM} = \frac{S_c}{N_c}. \tag{6.116}$$

Using Eq. (6.116) in Eq. (6.115), we get our final result:

$$\left(\frac{S_o}{N_o}\right)_{FM} = 3\beta^2\left(\frac{S_o}{N_o}\right)_{AM}. \tag{6.117}$$

From this, we conclude that the output signal-to-noise ratio can be made much higher in FM than in AM by increasing the modulation index β. It must be remembered, of course, that we assumed that the noise was small in deriving this result. The effects of the noise in the output are reduced and the factor $3\beta^2$ is called the noise quieting of the system [this can be expressed in decibels by taking $10 \log_{10} (3\beta^2)$]. But an increase in β also increases the bandwidth so that FM systems provide an improvement in signal-to-noise at the expense of an increase in bandwidth. For example, when $\beta = 5$ the output FM signal-to-noise ratio is 75 times that of an equivalent AM system but the bandwidth required is approximately 8 times larger (see Fig. 6.11). Thus the use of FM allows one to exchange bandwidth for signal-to-noise ratio.

To realize any signal-to-noise improvements in FM over AM we must have, from Eq. (6.117),

$$\beta > \frac{1}{\sqrt{3}} = 0.577.$$

However, this condition is approximately the transition point between narrowband and wideband FM. We conclude that narrowband FM provides no signal-to-noise improvement over AM. This is, of course, an expected result because NBFM does not use increased bandwidth to exchange for an improvement in output signal-to-noise ratio.

The exchange of bandwidth for signal-to-noise ratio in FM cannot be continued indefinitely. The reason for this is that the noise power increases with the increased receiver bandwidth necessary to demodulate the signal. Assuming a fixed signal power we find the noise power eventually becomes comparable to the signal power and the

results of the above analysis no longer hold. As the noise power becomes large in angle modulation, the phase variations of the noise take over and the performance of the system becomes very poor. This transition depends heavily on the input signal-to-noise ratio, giving rise to a "threshold" effect (Section 6.9).

EXAMPLE 6.8.1

Develop a result similar to Eq. (6.117) for the improvement of PM over AM.

Solution For PM, we write

$$s_i(t) = A \cos [\omega_c t + k_p f(t)] .$$

The phase detector output is proportional to the difference between the instantaneous phase of $s_i(t)$ and the carrier phase,

$$s_o(t) = [\theta(t) - \omega_c t] = k_p f(t) .$$

Similarly, using Eq. (6.104), we get the output noise for high signal-to-noise conditions:

$$n_o(t) = n_s(t)/A ,$$

$$S_{n_o}(\omega) = \frac{1}{A^2} S_{n_s}(\omega) .$$

Assuming that the input noise spectral density is white and equal to $\eta/2$ watts/Hz, we write

$$\overline{n_o^2(t)} = \frac{1}{\pi} \int_0^{\omega_m} \frac{1}{A^2} \eta \, d\omega = \frac{\eta \omega_m}{\pi A^2} .$$

Then the output signal-to-noise ratio is

$$\frac{S_o}{N_o} = \frac{\overline{s_o^2(t)}}{\overline{n_o^2(t)}} = \pi A^2 \frac{k_p^2 \overline{f^2(t)}}{\eta \omega_m} .$$

For the sinusoidal case, $\overline{f^2(t)} = a^2/2$ and $\Delta\theta = a k_p$ so that

$$\frac{S_o}{N_o} = \frac{\pi A^2}{2\eta\omega_m} (\Delta\theta)^2 .$$

Compared to the AM case for 100% modulation, this becomes

$$\left(\frac{S_o}{N_o}\right)_{PM} = (\Delta\theta)^2 \left(\frac{S_o}{N_o}\right)_{AM} .$$

Noting that $\Delta\theta = \beta$ for PM, we conclude that the signal-to-noise improvement in PM is similar to that in FM.

DRILL PROBLEM 6.8.1

The narrowband PM signal

$$\phi_{\text{NBPM}}(t) = \cos \omega_c t - \beta \cos \omega_m t \sin \omega_c t$$

is to be demodulated using a synchronous detector; i.e., multiplication by $\cos(\omega_c t + \theta)$ followed by a low-pass filter. Determine (a) if this will work satisfactorily; and (b) what value of θ is required for best results.

ANSWER: (a) Yes; (b) $\pm\pi/2$.

Before leaving this discussion of S/N ratios, we digress briefly to complete the comparison of CW modulation systems that was started in Chapter 5. The results are summarized in Table 6.1. For convenience, the input S/N ratios have been normalized in this table for the case of band-limited white noise with power spectral density $S_n(\omega) = \eta/2$ W/Hz and a bandwidth f_m equal to the bandwidth of the input modulating signal. This places single-sideband and double-sideband modulation systems on an equal basis for S/N comparisons.

From Table 6.1, it is evident that SSB and VSB systems offer spectral efficiency, requiring only as much (or slightly more) bandwidth as the input modulating signal. In contrast, angle modulation systems offer increased S/N at some expense in required bandwidth. Large-carrier amplitude modulation systems generally use envelope detection to minimize the complexity of the receiver. Angle modulation and

TABLE 6.1 A Comparison of CW Modulation Systems

Modulation	B/f_m	$\frac{S_o}{N_o}\left(\frac{S_i}{\eta f_m}\right)^{-1}$	$\frac{S_o}{N_o}\left(\frac{S_c}{\eta f_m}\right)^{-1}$	Typical Applications
DSB-SC	2	1		Analog instrumentation, multiplexing.
DSB-LC[a]	2	$\frac{\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}$	$\frac{\overline{f^2(t)}}{A^2}$	Broadcast radio, point-to-point voice.
SSB-SC	1	1		Point-to-point voice, multiplexing.
SSB-LC[a]	1	$\frac{\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}$	$\frac{\overline{f^2(t)}}{A^2}$	Point-to-point voice.
VSB-SC	1[+]	1		Facsimile, digital data.
VSB-LC[a]	1[+]	$\frac{\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}$	$\frac{\overline{f^2(t)}}{A^2}$	Television video.
FM[b]	$\approx 2(1 + \beta)$		$\frac{3}{2}\beta^2$	Broadcast radio, mobile radio.
PM[b]	$\approx 2(1 + \beta)$		$\frac{1}{2}(\Delta\theta)^2$	Telemetry, digital data.

[a] Above-threshold envelope detection assumed; for sinusoidal modulation, $f(t) = mA\cos \omega_m t$ so that $\overline{f^2(t)} = m^2A^2/2$.
[b] Above-threshold detection assumed; no deemphasis and sinusoidal modulation.

SSB-SC systems require moderate receiver complexity while other types of suppressed-carrier transmissions require more substantial receiver complexity. Each method has its advantages and disadvantages, and these must be evaluated in each specific situation.

6.9 THRESHOLD EFFECT IN FM

The threshold effect turns out to be quite pronounced in FM. Because this effect occurs when the noise and signal levels are comparable, it is more difficult to analyze than the high signal-to-noise case.

To gain some insight into this, suppose we consider the following simplified problem. Two phasors, A and B, as shown in Fig. 6.28(a), are added to yield a resultant phasor C. For simplicity, we assume that the magnitudes of A and B are fixed and that the angle of the phasor B can take all values in the range $-\pi < \theta_B \leq \pi$ with equal weighting. The angle of phasor A remains fixed. Phasor A can be used to represent the signal and phasor B the noise. We wish to find the mean-square value of the phase angle θ_C of the resultant and this will represent the average phase noise present in the measurement of angle.

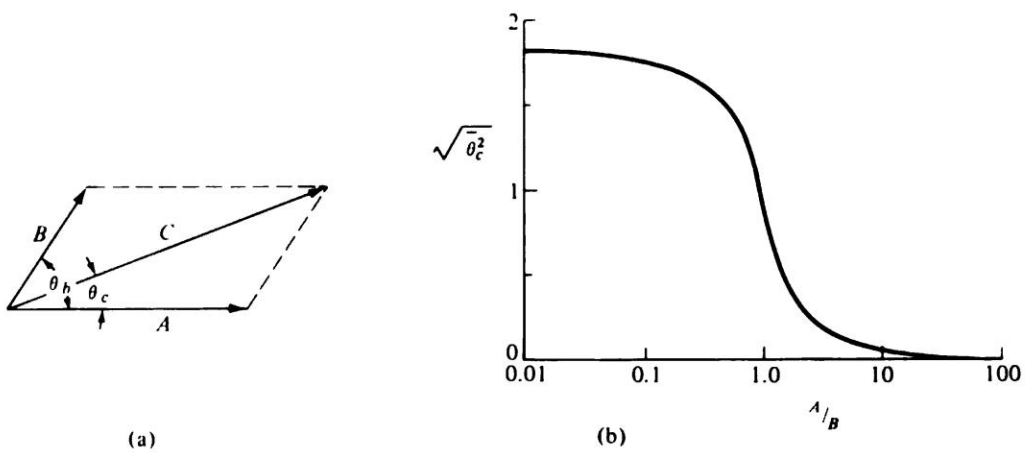


Figure 6.28 Phasor diagram and determination of phase noise.

From Fig. 6.28(a), we can write

$$\theta_C = \tan^{-1}\left(\frac{B \sin \theta_B}{A + B \cos \theta_B}\right), \tag{6.118}$$

so that the mean-square value is

$$\overline{\theta_C^2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\tan^{-1}\left(\frac{B \sin \theta_B}{A + B \cos \theta_B}\right)\right]^2 d\theta_B . \tag{6.119}$$

This can be solved numerically and the results are shown in Fig. 6.28(b). Note that the ratio A/B may be regarded as a voltage signal-to-noise ratio in this simplified model. From these results, we see that the mean-square phase noise increases rapidly when the ratio A/B is smaller than about 3.

In actuality, the problem is more complicated because the amplitude of the noise varies, and the presence of angle modulation in the signal also has some effect. There is also another effect peculiar to FM that makes the threshold effect even more abrupt, and we turn our attention to that next.

Referring back to Fig. 6.26, we note that the tip of the resultant phasor, labeled point P, moves about in the neighborhood of the tip of carrier phasor A due to the random fluctuations in the additive noise $n(t)$. As long as $n(t) \ll A$, point P remains close to the tip of carrier phasor A, and the principal value of the angle $\gamma(t)$ is approximately $n_s(t)/A$, as given by Eq. (6.104). As the noise increases in magnitude, the wandering of point P about the tip of the carrier phasor increases, and the discriminator output noise increases accordingly.

As the noise is increased further, however, a condition is reached where not only may the magnitude of the phasor $n(t)$ exceed the magnitude of the carrier phasor, but the angles may be such that point P sweeps *around* the origin (point O). This latter condition is demonstrated in Fig. 6.29(a). The net effect is that the phase angle $\gamma(t)$

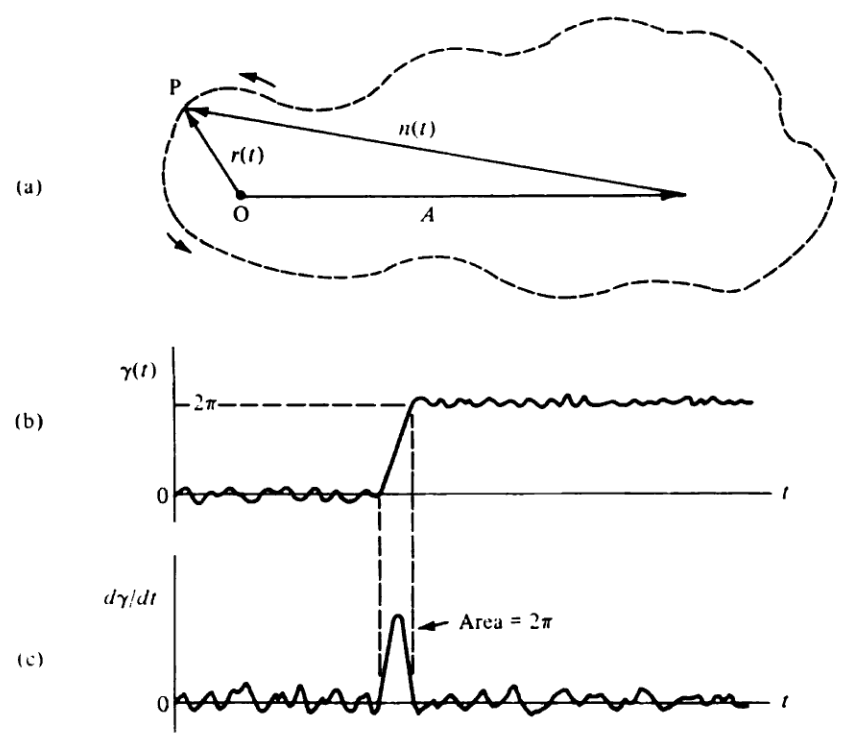


Figure 6.29 (a) Phasor diagram of the sum of a sine wave and narrowband noise, small S/N case. (b) Example of $\gamma(t)$ for a case in which point P encircles the origin (point O). (c) Plot of resulting $d\gamma/dt$, exhibiting spike noise.

may change by $\pm 2\pi$ radians within a very short interval of time, as illustrated in Fig. 6.29(b). Because the output of the FM discriminator is proportional to $d\gamma/dt$, impulselike noise spikes will occur in the output whenever the resultant phasor encircles point O. This is illustrated in Fig. 6.29(c). Heights and polarities of these noise spikes will vary, depending on the particular path taken by P around O, but the area of each is approximately $\pm 2\pi$ radians.

It is found experimentally that the occurrences of occasional noise spikes are heard as individual clicks in the FM discriminator output. Therefore the onset of spike noise in FM is sometimes referred to as click noise. As the signal-to-noise ratio is decreased a little further, these clicks rapidly merge into a crackling sound. As a result, the output signal-to-noise ratio decreases rapidly with only a relatively small decrease in the input S/N, and the results we described in the preceding section now fail to describe the performance adequately. This gives rise to a very pronounced FM threshold effect.

The presence of modulation varies the instantaneous angle of the carrier phasor in Fig. 6.29(a). The resulting behavior near the threshold condition is that, at a given time, there is a greater occurrence of encirclements about O in one direction than in the other. It turns out that the resulting decrease in encirclements in one direction does not quite offset the increase in encirclements in the other direction. Thus there is some dependence on modulation, although this effect is secondary.

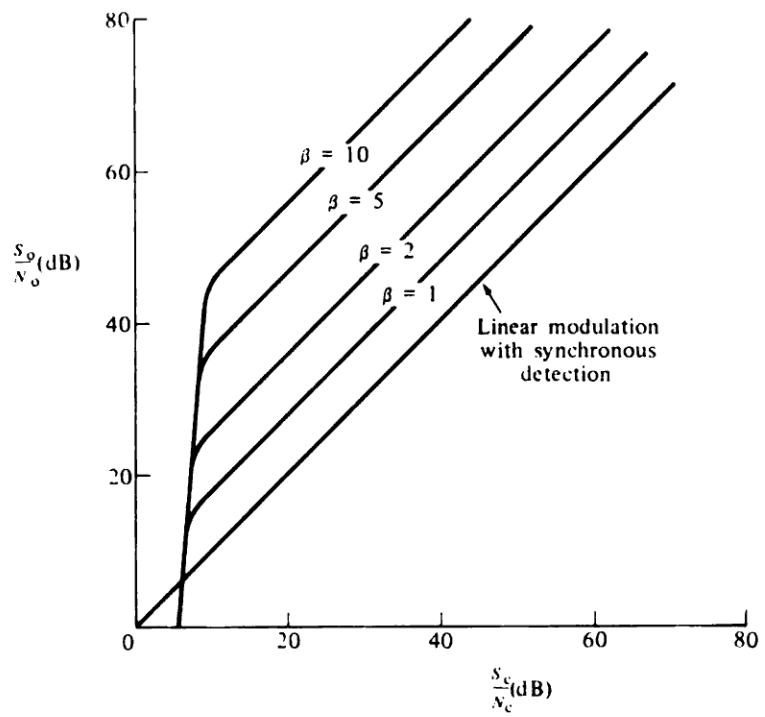


Figure 6.30 Signal-to-noise performance of wideband FM.

Figure 6.30 shows the signal-to-noise performance characteristics of a wideband FM system. The dependence of the threshold on the value of β is small and the input signal-to-noise ratio to achieve above-threshold operation is usually taken as about 10 dB for large β. Note that for small input signal-to-noise conditions the FM system may actually be inferior to an AM system. The amount of noise quieting above threshold increases with β, as noted earlier.

Improvement of the FM threshold has been the subject of considerable research, particularly for space communications where power is a main concern and signal-to-noise ratios may not be very large. One method used successfully to extend the FM threshold to lower values is that of the phase-locked loop (PLL). The occurrence of a noise spike in the PLL causes the phase error to exceed $\pm T/4$ ($\pm 90°$) in the phase-comparator characteristic (cf. Fig. 6.22). However, the dynamic range of the PLL output is limited by the phase-comparator characteristic and therefore the net energy due to the noise spike is decreased. It has been found that the overall threshold improvement is lowered by about 3 dB with the use of the PLL.

Other methods used successfully to extend the FM threshold include the frequency-locked loop and FM with feedback (FMFB). A discussion of the lowered threshold of the PLL and FMFB can be found in Reference 7.

There is no reason, of course, why the second signal in the phasor example above could not have been a second FM or PM signal. In this case we are interested in the signal-to-interference ratio. The rather sharp transition between good and poor interference-to-signal ratios in FM results in a very definite suppression of small signals in the presence of larger ones. This small-signal suppression effect together with a higher carrier frequency results in more closely defined service areas for commercial FM than for AM.

EXAMPLE 6.9.1

A given receiver for commercial FM has a noise temperature of 210 K. (a) Estimate the minimum required signal strength at the test generator terminals to obtain full quieting. The input resistance is 300 Ω; assume that the signal generator is at 290 K and matched to the receiver. (b) What is the theoretical minimum required signal strength for full quieting under these conditions?

Solution a) If we assume an IF bandwidth of 180 kHz, the available noise power is

$$P_n = kTB = (1.38 \times 10^{-23})(500)(180{,}000) = 1.24 \times 10^{-15}\ \text{W}.$$

The assumed threshold is 10 dB so that

$$P_s = 1.24 \times 10^{-14}\ \text{W},$$

$$\overline{e_s^2(t)} = P_s R = (1.24 \times 10^{-14})(300),$$

$$\sqrt{\overline{e_s^2(t)}} = 1.93\ \mu\text{V}.$$

b) For no receiver-contributed noise, $T = 290$ K (from the signal generator) so that

$$P_s = 10\,kT_0B = 7.20 \times 10^{-15}\ \text{W},$$
$$\overline{e_s^2(t)} = (300)(7.20 \times 10^{-15}),$$
$$\sqrt{\overline{e_s^2(t)}} = 1.47\ \mu\text{V}.$$

★ **DRILL PROBLEM 6.9.1**

A telemetry system for VHF weather satellite transmissions uses a (low-pass) bandwidth of 1.6 kHz. This is modulated on a 2.4 kHz carrier using DSB-SC and then frequency-modulated (AM/FM).

a) Use Carson's rule to estimate the required receiver bandwidth if the peak frequency deviation is set at 20 kHz.

b) Estimate the maximum noise temperature of the receiver if the minimum signal level is 1 μV rms across 75 Ω and the expected antenna noise temperature is 1150 K.

ANSWER: (a) 48 kHz; (b) 863 K.

6.10 SIGNAL-TO-NOISE IMPROVEMENT USING DEEMPHASIS

In commercial FM broadcasting it is found that signals arising from speech and music have most of the energy at the lower frequencies. In the output of an FM demodulator, however, the noise-power spectral density rises parabolically with frequency. Therefore the noise power spectral density is largest in the frequency range where the signal spectral density is smallest. This results in a rather inefficient communication system. To remedy this situation, we emphasize the high-frequency components in the input signal at the transmitter, *before the noise is introduced*. At the output of the FM demodulator in the receiver the inverse operation is performed to deemphasize the high-frequency components. The signal spectrum is restored to its original shape but the noise, which was added after the preemphasis, is now reduced.

The use of the preemphasis/deemphasis technique is an example of how a knowledge of the differences between characteristics of signal and noise can be applied to improve the performance of a communication system. This technique finds other applications, particularly in the audio recording area.[†]

[†] Much interest has developed in applying combinations of preemphasis filtering and dynamic-range signal-amplitude compression to suppress the effects of noise in audio recordings. These are discussed in Appendix F.

Some care must be exercised in the selection of the preemphasis to be used. If the high-frequency components are emphasized too much, the bandwidth of the transmitted FM spectrum will increase unless β is lowered, offsetting the advantage. A simple and straightforward approach in choosing a preemphasis characteristic is to choose one that will yield a white noise spectral density after demodulation. Because the noise power spectral density at the FM demodulator output rises parabolically with frequency, we then use a matching preemphasis for the signal; this requires that $|H(\omega)|^2 = \omega^2$, suggesting the possibility $H(\omega) = j\omega$. This transfer function corresponds to a differentiator and we arrive at the surprising result that we now should use PM instead of FM! However, if we had used PM entirely, the bandwidth would be more difficult to control for commercial use. Obviously what we want here is the best combination of both modulation methods.

What is needed, then, is a filter whose transfer function is constant for low frequencies and behaves like a differentiator at the higher frequencies. An example of an *RC* network that approximates this type of response is shown in Fig. 6.31(a). The Bode magnitude plot for this network is also shown and a corresponding deemphasis network for the receiver is shown in Fig. 6.31(b).

The choice of ω_1 and ω_2 in Fig. 6.31 determines the component values to be used. A reasonable choice for $f_1 = \omega_1/2\pi$ is usually taken as that frequency at which the signal spectral density is down by 3 dB; for broadcasting this is taken as 2.1 kHz ($RC = 75$ μsec) and it is assumed that the spectral density falls off very rapidly beyond this frequency.† The choice for $f_2 = \omega_2/2\pi$ is taken well above the highest audio frequency to be transmitted.

How much does the addition of the preemphasis-deemphasis improve the overall signal-to-noise ratio? Assuming that the two networks are chosen properly, there will be no net change in the signal. The noise spectral density at the demodulator output, however, is altered appreciably, as shown in Fig. 6.32.

The improvement in the signal-to-noise ratio can be found by calculating the decrease in the noise power. From Eq. (6.109), the noise power spectral density at the output of the FM demodulator is

$$S_{n_0}(\omega) = \eta\omega^2/A^2.$$

The transfer function of the deemphasis filter can be written as

$$H(\omega) = \frac{1}{1 + j\omega/\omega_1}. \tag{6.120}$$

† The high-frequency response is becoming so good in modern recordings that this assumption is no longer valid. Therefore stations have either had to decrease their modulation or purposely "color" their input spectrum to stay within the specified 75 kHz peak frequency deviation. The FCC has permitted stations to use a 25 μsec time constant (6.4 kHz) and Dolby B compression (see Appendix F) instead of 75 μsec. This choice proves compatible with existing 75 μsec deemphasis and permits full program dynamic range with an increased S/N ratio.

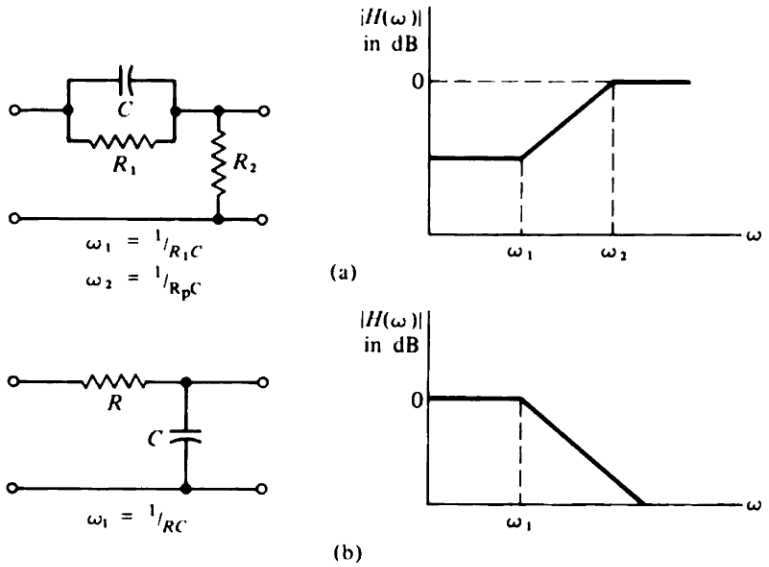


Figure 6.31 Example of a preemphasis/deemphasis network combination.

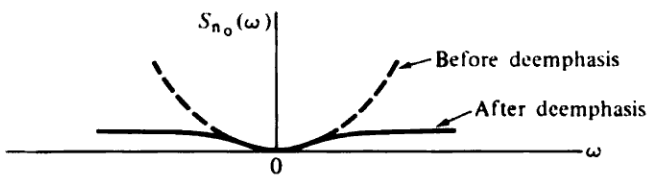


Figure 6.32 Noise spectral density following deemphasis.

The mean-square value of the noise after the deemphasis filter is

$$N_o' = \frac{1}{\pi}\int_0^{\omega_m} S_{n_o}(\omega)\,|H(\omega)|^2\,d\omega\,,$$

$$N_o' = \frac{\eta}{\pi A^2}\int_0^{\omega_m} \frac{\omega^2}{1 + (\omega/\omega_1)^2}\,d\omega\,. \tag{6.121}$$

Without the deemphasis filter, the noise would be

$$N_o = \frac{\eta}{\pi A^2}\int_0^{\omega_m} \omega^2\,d\omega\,. \tag{6.122}$$

Defining a noise improvement factor,

$$\Gamma = N_o/N_o'\,, \tag{6.123}$$

and evaluating Eqs. (6.121) and (6.122), we find that

$$\Gamma = \frac{1}{3}\,\frac{(\omega_m/\omega_1)^3}{(\omega_m/\omega_1) - \tan^{-1}(\omega_m/\omega_1)}\,. \tag{6.124}$$

For example, if $f_m = 15$ kHz, $f_1 = 2.1$ kHz, Γ is approximately 13 dB. A plot of the noise improvement factor Γ is shown in Fig. 6.33.

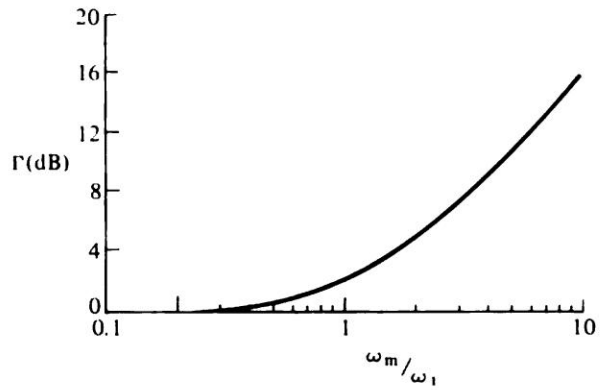


Figure 6.33 Signal-to-noise improvement using deemphasis.

Note that the use of preemphasis/deemphasis has decreased the noise in the FM demodulator output but has left the signal unchanged. Therefore the value of Γ represents a net increase in the effective signal-to-noise ratio without any required increase in the transmitted power. As shown in Fig. 6.33, this increase can be appreciable. With the use of this preemphasis, the transmitter essentially sends out a signal that is frequency-modulated by the lower audio frequencies and phase-modulated by the higher audio frequencies.

EXAMPLE 6.10.1

It was assumed, in the above discussion of preemphasis for commercial FM, that the higher spectral components in the modulating signal are negligibly small. Therefore Eq. (6.124) neglects any increase in the modulation power resulting from the preemphasis of higher spectral components. In modern recordings this assumption may lose some of its validity. For example, suppose that the power spectral density of the signal is $S_f(\omega) = (\text{constant})/[1 + (\omega/\omega_1)^2]$.

a) Determine the required change in modulation level with and without preemphasis for a fixed modulation power.

b) What is the net S/N improvement under these conditions?

Solution a) Without preemphasis, the modulation signal power, within a constant, is

$$P = \frac{1}{\pi}\int_0^{\omega_m} \frac{1}{1 + (\omega/\omega_1)^2}\,d\omega = \frac{\omega_1}{\pi}\tan^{-1}\left(\frac{\omega_m}{\omega_1}\right).$$

With preemphasis, the modulation signal power increases unless an attenuation K is used so that, within a constant, the power is

$$P = \frac{1}{\pi}\int_0^{\omega_m} K\,d\omega = K\omega_m/\pi\,.$$

If this modulation power is to remain fixed, we have

$$K = \frac{\tan^{-1}(\omega_m/\omega_1)}{\omega_m/\omega_1}.$$

For $f_m = 15$ kHz, $f_1 = 2.1$ kHz, this gives $K = 0.200$ (-6.98 dB).

b) From Eq. (6.124), the S/N improvement as a result of preemphasis/deemphasis is 13.28 dB. However, now the modulation level must be decreased by 6.98 dB so that the net S/N improvement is only $(13.28 - 6.98) = 6.30$ dB. This gives some insight into why commercial FM stations are introducing alternative preemphasis/deemphasis methods as the quality of recordings continues to improve.

DRILL PROBLEM 6.10.1

Suppose that the preemphasis/deemphasis networks described in this section are proposed for use in commercial AM. Assuming that the input noise spectral density is white and that an envelope detector is used under high signal-to-noise conditions, calculate the improvement possible for $f_m = 5$ kHz. [*Hint:* Assume that the output noise is white.]

ANSWER: 3.07 dB.

6.11 SUMMARY

The variation of frequency of a continuous sinusoid in proportion to the amplitude of a modulating signal is called frequency modulation (FM). The variation of phase of a continuous sinusoid in proportion to the amplitude of a modulating signal is called phase modulation (PM). Both are examples of angle modulation. The instantaneous frequency is the time derivative of the phase so that PM and FM are closely related.

In the generation of FM there is an amplitude-to-frequency conversion and a frequency-to-frequency conversion. The peak frequency deviation is a measure of the former and the modulating frequency the latter. The ratio is called the modulation index, β.

The value of β determines the spectral characteristics of the FM waveform. For values of β below about 0.2, the spectral density of an FM waveform consists of two sidebands about a large carrier and this condition is called narrowband FM. The spectral details for larger values of β, called wideband FM, depend on each particular modulating signal because the FM generation is nonlinear. However, the overall bandwidth can be approximated by adding twice the peak frequency deviation to twice the modulating signal bandwidth (Carson's rule).

Most of the principles of FM also apply to PM. A difference is that the peak phase deviation is controlled by the modulating signal in PM and therefore the peak frequency deviation varies with the frequency of the modulating signal as well as its amplitude.

The total average power in an angle-modulated signal remains constant regardless of the modulation index.

Generation of angle-modulated signals is accomplished either by first generating a narrowband signal and then increasing the modulation index using frequency multiplication or by generating a wideband signal directly. Demodulation can be accomplished using a frequency-to-voltage converter circuit called a discriminator or by using feedback techniques around a voltage-to-frequency converter.

Wideband angle-modulated systems offer an improvement in output signal-to-noise over AM systems at the expense of increased bandwidths. This improvement can only be realized, however, when the input signal-to-noise ratio is above a threshold level. A preemphasis/deemphasis circuit is used in commercial FM broadcasting to partially suppress high-frequency noise in the demodulator output.

Selected References for Further Reading

1. J. J. DeFrance. *Communications Electronics Circuits*. Second ed. San Francisco: Rinehart Press, 1972.
   Chapter 9 contains a discussion of the circuits and operating characteristics of practical FM systems; easy reading.
2. S. Haykin. *Communication Systems*. Second ed. New York: John Wiley & Sons, 1983.
   Chapter 4 has a treatment of FM at about the same level, and includes the multitone case.
3. H. Stark, F. B. Tuteur, and J. B. Anderson. *Modern Electrical Communications*. Second ed. Englewood Cliffs, N.J.: Prentice Hall, 1988.
   Chapter 7 presents angle modulation with explanations of some of the popular circuits used. (For more detailed circuit descriptions, see Chapters 11 and 12 in Reference 5 of Chapter 5 of the present book.)
4. R. M. Gagliardi. *Introduction to Communications Engineering*. Second ed. New York: John Wiley & Sons, 1988.
   The topics of both receiver frequency acquisition, and phase referencing and timing are covered in Chapter 10.
5. R. E. Best. *Phase-Locked Loops*. New York: McGraw-Hill, 1984.
   An interesting discussion of phase-locked loops, with emphasis on practical applications and the digital PLL.
6. F. M. Gardner. *Phaselock Techniques*. Second ed. Somerset, N.J.: Wiley-Interscience, 1979.
   Contains a good discussion of the linearized second-order phase-locked loop and its performance in the presence of noise; provides useful design criteria.
7. H. Taub and D. L. Schilling. *Principles of Communications Systems*. Second ed. New York: McGraw-Hill, 1986.
   Chapters 4, 9, and 10 contain a well-written discussion of angle modulation at a more advanced level, stressing the noise suppression and threshold effects.
8. P. F. Panter. *Modulation, Noise, and Spectral Analysis*. New York: McGraw-Hill, 1965.
   Written at a more advanced level, this book is probably the most complete reference available on the subject of angle modulation, devoting 10 chapters to the topic.

PROBLEMS

6.1.1 Determine the instantaneous frequency, in Hertz, of each of the following waveforms.

a) $10 \cos (100\pi t + \pi/3)$
b) $10 \cos (200\pi t + 10 \sin \pi t)$
c) $2 \exp [j200\pi t(1 + \sqrt{t})]$
d) $\cos 200\pi t \cos (5 \sin 2\pi t) + \sin 200\pi t \sin (5 \sin 2\pi t)$

6.1.2 Sketch a) the phase versus time, and b) the instantaneous frequency versus time for the signal

$$\phi(t) = A \cos [\omega_c t + p(t)\pi/6],$$

where $p(t)$ is a ± 1 square wave changing states every $2/f_c$ seconds at $(\omega_c t) = \pi/2, 3\pi/2, \ldots$.

6.1.3 a) Find an approximation to the Fourier series expansion of the angle-modulated waveform $\psi(t) = \exp (j\beta \sin \omega_m t)$ for small β by using the MacLaurin series expansion for $\exp (x)$ and retaining only the first two terms in the expansion.
b) Sketch the line spectrum of the approximation to $\psi(t)$, as determined in part (a).
c) Determine the Fourier transform (spectral density) of the approximation to $\psi(t)$, as determined in part (a).

6.2.1 A 1-GHz carrier is frequency-modulated by a 10-kHz sinusoid so that the peak frequency deviation is 100 Hz. Determine

a) the approximate bandwidth of the FM signal;
b) the bandwidth if the modulating signal amplitude were doubled;
c) the bandwidth if the modulating signal frequency were doubled;
d) the bandwidth if both the amplitude and the frequency of the modulating signal were doubled.

6.2.2 Determine the phase angle from carrier, the instantaneous frequency deviation from carrier, and the magnitude (envelope) of the SSB-LC waveform described by

$$\phi(t) = A(1 + me^{j\omega_m t})e^{j\omega_c t}.$$

6.2.3 The upper sideband of an AM waveform (DSB-LC) with sinusoidal modulation and modulation index m is multiplied by a factor α, where: $0 \leq \alpha \leq 1$. Derive a relation for the peak (i.e., maximum) phase deviation from carrier as a function of m and α.

6.2.4 A vestigial sideband (VSB) signal is generated from an input $f(t) = \cos 2000\pi t$ by first generating a DSB-LC signal with a carrier frequency $f_c = 10$ kHz and modulation index m. Then this DSB-LC signal waveform is applied to a VSB filter. The resulting line spectrum at the filter output is a line at 9 kHz with height a, a line at 10 kHz with height 1.0, and a line at 11 kHz with height b.

a) What is the modulation index of the DSB-LC signal, in terms of a and b?
b) Sketch and label a phasor diagram for the VSB signal if $a = 0.40$ and $b = 0.60$. [*Hint:* First express your results in terms of real and imaginary components.]
c) Find an expression for the peak instantaneous frequency from carrier (in terms of a, b, m) present in the VSB signal.

★ **6.3.1** The voltage waveform $f(t)$ shown in Fig. 2.18 is applied to a FM modulator with a modulation sensitivity of $k_f = 10$ kHz/mV about the center frequency f_c.

a) Sketch a diagram of the instantaneous frequency versus time.
b) Sketch an appropriate magnitude spectral density if the period (T) is 1 sec.

√ ★ **6.3.2** Using a FFT algorithm and $N = 256$, compute and plot the magnitude spectral density of a sinusoid over $(0, T)$, $T = 1$, whose instantaneous frequency varies linearly from $(f_c - D)$ to $(f_c + D)$ in the computation interval. [*Hint:* In the computation, let $f_c = N/4$ and $D = N/8$.] Compare your result with Fig. 6.8.

6.3.3 Repeat Problem 6.2.1 with the exception that the peak frequency deviation is 40 kHz. Use both Carson's rule and Bessel functions and compare results.

6.3.4 Using Appendix G, make a comparison between bandwidths using Carson's rule, Bessel functions with a 1% rule for significant sidebands [cf. Eq. (6.38)], and Bessel functions with a 10% rule for significant sidebands. Make the comparison for $\beta = 1, 2, \ldots, 6$.

6.3.5 A 1-kHz sinusoid is used to frequency modulate a 50-kHz carrier signal; the peak frequency deviation from carrier is 200 Hz. This FM signal is applied to a nonlinear system whose input-output transfer characteristic is $e_o(t) = 2e_i^2(t)$. The output is filtered with an ideal bandpass filter (BPF) having a bandwidth of 20 kHz, centered at 100 kHz.

a) Sketch to scale the resulting line spectrum.
b) Using a table of Bessel functions, estimate the modulation index β of the output signal.

6.3.6 A carrier waveform is frequency-modulated by the sum of two sinusoids:

$$\phi(t) = 100 \cos (\omega_c t + \sin \omega_m t + 2 \sin 2\omega_m t),$$

where $f_c = 100$ kHz and $f_m = 1$ kHz.

a) What is the peak frequency deviation from carrier?
b) Estimate the net bandwidth required for transmission of this FM signal.
c) Sketch to scale the resulting magnitude line spectrum (one-sided, above carrier). [*Hint:* Express $\phi(t)$ in complex-valued notation and use Bessel functions for series representations to recognize the required coefficients.]

6.3.7 The output of a frequency modulator is filtered with a memoryless ideal bandpass filter with bandwidth B. On the basis of the concept of instantaneous frequency and sinusoidal modulation, develop an argument why the condition $B \geq 2\Delta f$ is necessary to avoid distortion.

√ **6.3.8** Use a digital computer and graphics monitor to compute and display the magnitude of the FFT using $N = 128$ over an interval $T = 1$ to simulate a spectrum analyzer. Display the magnitudes over $(0, N/2)$. This will result in a display with the integers as frequency (in Hertz) for the horizontal axis, with a resolution of 1 Hz. Because only relative magnitudes are important here, normalize the vertical display to the maximum value.

a) Compute and display the magnitude spectrum for: $x(t) = \cos [64\pi t + 6 \sin (4\pi t)]$. Print a copy of the spectrum display. Also list the computed sideband magnitudes, and compare their relative magnitudes with tabulated values (cf. Appendix G).
b) Compute and display the magnitude spectrum for: $y(t) = \cos (64\pi t + \sin 4\pi t + 2 \sin 8\pi t)$. Print a copy of the spectrum display; explain your results. (It may be helpful to refer to Problem 6.3.6c.)

6.3.9 The analytical method used to find the spectral density of FM with sinusoidal modulation can be used for more general periodic modulating signals with zero mean. Consider the case in which the modulating signal is a periodic square wave of unit amplitude and period T. (For convenience, assume that the square wave has even sym-

metry about the origin. Let the peak frequency deviation from carrier be $\Delta\omega$, and define a modulation index $\beta = \Delta\omega/\omega_0$, where $\omega_0 = 2\pi/T$.

a) Sketch the instantaneous frequency and phase.
b) Derive an expression for the spectral density.
c) Sketch the magnitude spectrum for $\beta = 5$.

6.4.1 The sinusoidal signal $f(t) = a \cos 2\pi f_m t$ is applied to the input of a FM system. The corresponding modulated signal output (in volts) for $a = 1$ V, $f_m = 1$ kHz, is: $\phi(t) = 100 \cos (2\pi \times 10^7 t + 4 \sin 2000\pi t)$ across a 50-ohm resistive load.

a) What is the peak frequency deviation from carrier?
b) What is the total average power developed by $\phi(t)$?
c) What percentage of the average power is at 10.000 MHz?
d) What is the approximate bandwidth, using Carson's rule?
e) Repeat parts (a)–(d) for the input parameters $a = 0.75$ V, $f_m = 2$ kHz; assume all other factors remain unchanged.

6.4.2 A certain sinusoid at a frequency of f_m Hz is used as the modulating signal in both an AM (DSB-LC) and a FM system. When modulated, the peak frequency deviation of the FM system is set to three times the bandwidth of the AM system. The magnitudes of those sidebands spaced $\pm f_m$ Hz from carrier in both systems are equal, and the total average powers are equal in both systems. Determine a) the modulation index of the FM system; and b) the modulation index of the AM system.

6.4.3 The output of a given FM modulator with a sinusoidal input is: $\phi(t) = A \cos (\omega_c t + \beta \sin \omega_m t)$. This output $\phi(t)$ is applied to a synchronous detector and RC low-pass filter, with $(RC)^{-1} = \omega_m$. Develop an expression for the average power at the filter output if it is given that the output is one watt when $\beta = 0$.

6.4.4 A FM carrier waveform is sinusoidally modulated.

a) For what values of β is all of the power in the sidebands (i.e., zero average power at the carrier frequency)?
b) Using this information, devise an experimental procedure for the calibration of a given frequency modulator.
c) Other values could be used in calibration. Plot the lowest nonzero value of β for which the average power in a given sideband first goes to zero, versus the harmonic number of the sideband (for $0 < n \leq 6$).

√ **6.4.5** Show that the error in the total average power caused by truncating the bandwidth to the condition $n = \beta$ of a sinusoidally modulated FM signal is $\leq 4.1\%$.

6.5.1 Answer the questions of Problem 6.4.1 for the PM signal $\phi(t) = 100 \cos (2\pi \times 10^7 t + 4 \cos 2000\pi t)$ with the same input parameters.

6.5.2 A transmitter is modulated with a sinusoid of frequency $f_m = 1$ kHz to generate an angle-modulated signal with a peak frequency deviation of 20 kHz. Determine the modulation index β and the bandwidth of the modulated signal for each of the following changes, assuming that all other factors remain unchanged.

a) FM; f_m increased by 4
b) FM; f_m decreased by 4
c) PM; f_m increased by 4
d) PM; f_m decreased by 4

6.5.3 A carrier waveform is phase-modulated by a sinusoidal signal $f(t)$. The peak phase deviation is 1 rad when the peak input amplitude is 0.5 V. Find the bandwidth using Carson's rule, and the ratio of the average power in the carrier and first-order sidebands to that in all remaining sidebands for each of the following inputs. [Assume that $f(t)$ is expressed in volts in each case.]

a) $f(t) = \cos 1000\pi t$
b) $f(t) = 1.2 \cos 300\pi t$
c) $f(t) = 1.9 \cos 200\pi t$

6.5.4 The testing procedure developed for a certain PM system designed to transmit remote sensing information to an orbiting satellite uses the staircase-type periodic modulating signal $\gamma(t)$ shown in Fig. P–6.5.4.

a) Derive an expression for the spectral density, referred to the carrier frequency ω_c, for this modulation as a function of the peak phase deviation α. (The comments in Problem 6.3.9 may be helpful here also.)
b) Sketch the magnitude spectrum for $\alpha = \pi/3$ (60°).

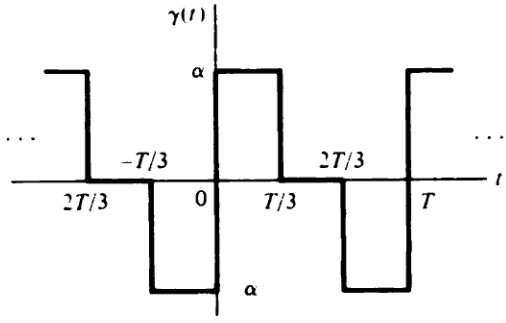


Figure P–6.5.4.

6.6.1 A 1.28-GHz signal is derived from a 10-MHz oscillator using a $N = 2^7$ frequency multiplier chain. For parts (a)–(c), assume that this multiplier chain is ideal and contributes no phase noise of its own.

a) What is the maximum frequency inaccuracy in the 1.28-GHz signal if the accuracy of the 10-MHz oscillator is within one part in 10^9?
b) What is the frequency stability of the 1.28-GHz signal if the output of the 10-MHz oscillator varies ±0.05 Hz?
c) What is the frequency stability of the output if the 10-MHz oscillator varies ±5° per second?
d) What is the frequency stability of the output if the 10-MHz oscillator varies ±5° per second and each ×2 multiplier contributes ±2° phase noise? (Include both the uncorrelated and the completely correlated cases.)

6.6.2 An Armstrong-type FM system to handle an input with maximum frequency content of 3.3 kHz first generates an FM signal with $\beta = 0.20$ (constant), $f_c = 5$ MHz. The desired output is required to have a carrier frequency of 30.0 MHz. Using one ×8 frequency multiplier and one frequency mixer (plus one required bandpass filter to follow each), devise two different system configurations to accomplish the desired objectives. Specify all frequencies and bandwidths in your diagrams.

6.6.3 Assume that the oscillator frequencies f_1, f_2 in Fig. 6.13 are not derived from a common source, and that each has a frequency inaccuracy of one part in 10^8. What is the maximum frequency inaccuracy in the output signal?

6.6.4 A 200-kHz carrier signal is frequency-modulated by a 1-kHz sinusoid so that the peak frequency deviation is 200 Hz. This FM signal is applied to a chain of four consecutive $\times 4$ frequency multipliers. Each frequency multiplier is followed by a BPF; in addition, the first $\times 4$ multiplier is preceded by a BPF. Specify the center frequency and the bandwidth (use Carson's rule) for each of the five filters.

6.7.1 Let the FM signal $\phi(t) = A \cos(\omega_c t + \beta \sin \omega_m t)$ be the input to the RL discriminator shown in Fig. 6.17(b). Let r be the forward resistance of the piecewise linear diode (you may assume infinite diode back resistance), and let R_1, C be the shunt resistance and capacitance, respectively, of the envelope detector to ground. Analyze the operation of this discriminator. Briefly describe the design tradeoffs in (low and high values of) choices in a) R/L; b) $R_1 C$ and r in terms of the pertinent parameters of the FM signal input.

6.7.2 Let the FM signal $\phi(t) = A \cos(\omega_c t + \beta \sin \omega_m t)$ be applied to the input of a system that generates a narrow (assume rectangular, for convenience) pulse of fixed amplitude a_0 and duration τ whenever the amplitude of $\phi(t)$ passes through zero. The output of this "zero crossings detector" is applied to a filter that passes only the fundamental of the pulse train.

a) Show that if $f_c \tau \ll 1$ then the filter output voltage can be made to be proportional to the instantaneous frequency of $\phi(t)$.

b) Investigate possible replacement of the output filter with a RC low-pass filter.

6.7.3 A certain FM demodulation system designed to operate at a carrier frequency ω_c generates two phasors $A_1 \exp(j\theta_1)$, $A_2 \exp(j\theta_2)$, which are adjusted to have complementary phase angles ($\theta_1 = -\theta_2$) at $\omega = \omega_c$. The difference between magnitudes $(A_2 - A_1)$ is made proportional to the instantaneous frequency from carrier. The output voltage of this demodulator is proportional to the phase angle of the resultant of the addition of the two phasors.

a) Show that the output voltage can be made proportional to the instantaneous frequency from carrier of $\phi(t)$ if $(A_1 + A_2)$ is held constant.

b) Investigate the design trade-offs on θ_1.

6.7.4 Estimate the maximum peak frequency deviation that can be used before the error in the demodulator gain in Example 6.7.2 is 1% from a linear characteristic. (It may be helpful to refer to Example 6.6.2.) Using this result, what error was assumed to be tolerable in Example 6.7.2?

6.7.5 For narrowband sinusoidal inputs, a synchronous detector can be used for a phase comparator. Show that if the two inputs to such a comparator are $x(t) = A_1 \cos[\omega_c t + \theta_1(t)]$ and $y(t) = A_2 \cos[\omega_c t + \theta_2(t)]$ then the output of the phase comparator, within a gain constant, is

$$z(t) = \tfrac{1}{2} A_1 A_2 \cos \theta(t), \quad \text{where } \theta(t) = \theta_1(t) - \theta_2(t).$$

6.7.6 a) Sketch the phase-comparator characteristic for a symmetric square-wave input at a fundamental frequency f_c and a VCO rectangular wave with one level duration one-fourth the period and with a fundamental frequency f_c.

b) Describe the effect of the rectangular wave on the comparator gain k_c and compare with that of a square wave.

6.7.7 Two periodic functions, $x(t)$ and $y(t)$, both having the same period, are inputs to a phase comparator.

a) Using a Fourier exponential series representation for each, find a Fourier exponential series representation for the comparator output.
b) What is the average value of the comparator output if one of the inputs has zero average value?
c) What is the output when $x(t)$ and $y(t)$ have no harmonic content in common? (For example, examine the case in which one input is a symmetric square wave with fundamental f_c, the second is a square wave with fundamental $2f_c$.)

√ **6.7.8** Let $x(t)$ be a symmetric square wave, and $y(t)$ be a sinusoid. The period of each waveform is the same. Using a 64-point FFT,

a) compute $R_x(\tau)$ for one period of the symmetric square wave;
b) compute $R_y(\tau)$ for one period of the sinusoid;
c) compute $R_{xy}(\tau)$;
d) normalize all outputs to the peak value of part (a) and plot over one period.

★ **6.7.9** A first-order PLL with a linear phase comparator characteristic (cf. Fig. 6.22b) is initially adjusted to an operating null point in the presence of unmodulated carrier. The time constant of the PLL is 10 μsec.

a) If the input is the sinusoidally modulated FM signal:

$$\phi(t) = A \cos(\omega_c t + \beta \sin \omega_m t),$$

determine the maximum allowable peak frequency deviation from carrier if the PLL is to recover the modulation waveform without distortion.
b) For demodulation of FM that uses a fixed peak frequency deviation, it is required that the frequency response of the demodulated output be as flat as possible (versus ω_m). Determine the approximate bandwidth of the PLL.
c) The VCO has the property that when its input is changed by 1 V, its frequency changes by 100 kHz. Determine the maximum signal output of the PLL.

★ **6.7.10** a) Determine the design parameters ($k_c k_f$) and τ_1 of a minimum-bandwidth second-order PLL to demodulate the audio for commercial (monophonic) television. Base your design on a sinusoidally modulated FM signal with a peak frequency deviation of 25 kHz and a maximum audio frequency of 15 kHz. Assume a triangular phase-comparator characteristic, $\zeta = 0.707$, and $\tau_1 = \tau_2$.
b) A frequency step of 10 kHz is applied to this PLL. If $k_c = 1$ V/rad, determine the peak phase error, the steady-state phase error, and the steady-state voltage output.

★ **6.7.11** Repeat Problem 6.7.10 for a first-order PLL.

★ **6.7.12** Repeat Example 6.7.4 and Drill Problem 6.7.2 for a sinusoidal phase-comparator characteristic. Linearize the characteristic with the restriction that the maximum allowable percentage error from a linear characteristic be 5%.

6.8.1 Show that Eq. (6.115) can be written as

$$\frac{S_o}{N_o} = \frac{3}{2}\beta^2\left(\frac{S}{N}\right)_{\text{baseband}}$$

6.8.2 In the derivation of Eq. (6.115), f_m represented both the frequency of the sinusoidal modulating signal and the cutoff frequency of the LPF at the discriminator output. Therefore this result holds for the case in which the sinusoidal modulating signal is at the highest frequency within the bandwidth of the LPF. Show that in the case where f_m represents the frequency of the sinusoidal modulating signal, f_M is the cutoff of the LPF (and BPF), $f_m \leq f_M$, then the result given in Eq. (6.115) is multiplied by the factor $(f_m/f_M)^2$.

6.8.3 Suppose that the output of the FM discriminator in Fig. 6.25 were applied to a low-pass filter with frequency transfer function $H(\omega)$, instead of the ideal LPF assumed in Eq. (6.110). Assuming sinusoidal modulation at $\omega = \omega_m$, show that Eq. (6.115) can be used except that N_c must be replaced by N_{eq}:

$$N_{eq} = 3\eta \frac{\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} \omega^2 |H(\omega)|^2 \, d\omega}{\omega_m^2 |H(\omega_m)|^2}$$

6.8.4 A communication system operates in the presence of white noise with two-sided power spectral density $S_n(\omega) = 0.25 \times 10^{-14}$ W/Hz and with total path losses (including antennas) of 100 dB. The input bandwidth is 10 kHz. Calculate the minimum required carrier power of the transmitter for a 10-kHz sinusoidal input and a 40-dB output S/N ratio if the modulation is

a) AM (DSB-LC), with $m = 0.707$ and $m = 1.0$;
b) FM, with $\Delta f = 10$ kHz and $\Delta f = 30$ kHz;
c) PM, with $\Delta\theta = 1$ rad and $\Delta\theta = 3$ rad.
d) Summarize your findings by marking these points on a graph of power required (in kW) versus bandwidth required (in kHz).

6.8.5 A given FM communications system design uses the baseband and transmission conditions given in Problem 6.8.4.

a) Calculate the minimum power and bandwidth required for $\beta = 1$.
b) Suppose that a system cost factor is defined as the product of transmitted power times the cube of the bandwidth. Determine the largest value of β such that this cost factor is within that of a transmitter having 20% greater power and using double the AM bandwidth for the same baseband signal.
c) Repeat part (b) for an increase of 50%, instead of 20%.

6.8.6 A frequency-division multiplexing system uses SSB-SC subcarrier modulation and FM main-carrier modulation. There are 20 equal-amplitude voice-input channels, each band-limited to 3.3 kHz. A 0.7-kHz guard band is allowed between channels and below the first channel.

a) Determine the final transmission bandwidth if the peak frequency deviation is 400 kHz.
b) Compute the degradation in signal-to-noise of input No. 20 when compared with input No. 1. (Assume a white input noise spectral density to the discriminator and no deemphasis.)
c) Repeat part (b) if PM were used.

6.9.1 In the 1950s the Television Allocation Study Organization (TASO) developed the following six-level subjective picture quality scale.

TASO grade	Picture Quality	S/N, dB
1	Excellent	>41
2	Fine	33–41
3	Passable	28–33
4	Marginal	23–28
5	Inferior	17–23
6	Unusable	<17

This scale is for the video (luminance) signal, transmitted using AM; the audio is transmitted using FM. If a given station transmits the audio carrier at −7 dB with respect to the video carrier, estimate if the audio or the video first becomes unacceptable. Support your reasoning.

★ **6.9.2** Make a graph of the signal input, in dBf (1 dBf = 10^{-15} W), versus receiver noise figure, in decibels, required for full quieting in a receiver designed for the commercial FM broadcast band. State your assumptions.

★ **6.9.3** Two given commercial FM receivers are advertised at differing prices. Both have the same IF bandwidth. The more expensive receiver is advertised as requiring 10 dBf (1 dBf = 10^{-15} W) of signal input for full quieting, whereas the cheaper one requires 13 dBf for full quieting. In use, both receivers are connected to antennas; the effective antenna noise temperature in each case is 1400 K. How many decibels of antenna gain can be sacrificed by the owner of the more expensive receiver to give the same S/N performance in both receivers when tuned to the same station?

6.10.1 A 10-kHz sinusoidal signal is to be transmitted using FM in the presence of additive white noise. If the S/N improvement at the demodulator output is required to be 20 dB, determine the required peak frequency deviation if a) no preemphasis/deemphasis is used; and b) the standard preemphasis/deemphasis is used (cf. Fig. 6.33). c) Make a plot of the required peak frequency deviation versus f_m for the conditions in part (b) over the frequency range 5 kHz $\leq f_m \leq$ 15 kHz.

6.10.2 A PM receiver is used to receive a commercial (mono) FM transmission. Describe what will happen if a) no preemphasis is used; and b) the standard (75 μsec) preemphasis is used in the transmission.

6.10.3 A given preemphasis/deemphasis system is shown in Fig. P–6.10.3. The power spectral density of the additive noise is $S_n(\omega) = C|\omega|$, where C is a constant. The magnitude frequency transfer function of the deemphasis filter is designed to be constant for $f <$ 2 kHz and to yield a white output noise spectral density over the frequency range 2 kHz $< f <$ 15 kHz.

a) What is the magnitude frequency transfer function $|H_p(\omega)|$ of the preemphasis filter required to yield no overall net signal distortion over this frequency range?

b) Plot the mean-square S/N improvement versus f_m using this system for $0 < f < f_m$ if (1) $|H_d(\omega)| = 1$ for $f_m \leq$ 2 kHz, and (2) $f_m \leq$ 15 kHz.

c) Assume that the power spectral density of the input signal is of the form $[1 + (f/f_1)^2]^{-1}$, where $f_1 = 2$ kHz. Plot the increase in the average power, in decibels, at the output of the preemphasis filter versus f_m under the conditions of part (b). Reference your graph to the output power at $f_m = 2$ kHz.

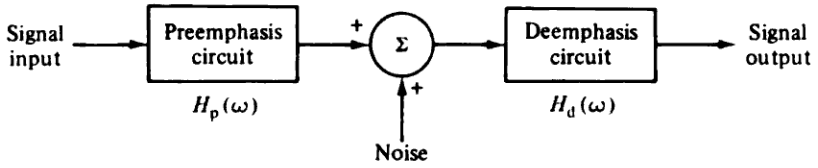


Figure P–6.10.3.

6.10.4 Let the desired power spectral density after preemphasis be (within a constant):

$$\begin{Bmatrix} 1 & \text{for } f < 2 \text{ kHz} \\ f/(2 \text{ kHz}) & 2 \text{ kHz} \le f < 15 \text{ kHz} \end{Bmatrix}$$

Determine the time constant of a first-order preemphasis filter (cf. Fig. 6.31) which will give the best approximation to the desired characteristic for a white spectral input.

6.10.5 In stereo FM receivers the deemphasis circuits are placed after the separation into the L and R signals.

a) Compare the resulting S/N ratio with that of a mono FM receiver if no preemphasis were used in either.

b) Show that stereo FM is 22 dB noisier than monophonic FM if the standard preemphasis circuits are used. [*Hint:* Use Eq. (5.65).]

DESIGN PROBLEMS

Design Problem 6.1 FM Microphone Audio System

A short-range, low-power, small transmitter is desired to send good-quality voice (assume a bandwidth of 10 kHz) from a microphone/amplifier to a high-quality FM receiver for the standard commercial FM band. It is decided that the transmitter will use a varactor diode (see Example 6.6.2) with a capacitance characteristic given by $C = 20/\sqrt{1 + 2V}$ pF, with a recommended reverse-bias dc voltage of 4 V. This is used in a simple LC tuned circuit; the minimum allowable value of L in the design is 0.25 μH. Assume that the FM receiver has an IF bandwidth of 200 kHz and a noise figure of 4 dB; use 1000 K for the antenna noise temperature (to account for noise from interference, etc.) Neglect possible antenna gain/loss and path loss effects.

The transmitter consists of audio amplifier stages and the tuned output power amplifier; the primary power drain is the output power amplifier. To be easily carried on a person and concealed, it should have as low a battery power drain as possible. Parameters to be set in this design are (1) the peak-to-peak voltage applied to the varactor diode, and (2) the amount of output power required for above-threshold operation of the system. Using the help of computational methods, compute the total harmonic distortion (THD) of the system for a 1-kHz sinusoidal input and study the trade-offs in terms of the THD. Assume that the receiver discriminator characteristic is ideal for these computations.

CHAPTER

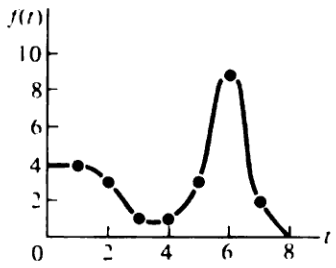
7 Pulse Modulation

In the two previous chapters we have considered applications of continuous-wave (CW) modulation. Recall that in CW modulation the parameters of a continuous sinusoid are varied in proportion to the modulating signal. In amplitude modulation, the amplitude of the sinusoid is varied; in angle modulation the angle and the angular rate are varied. Now we shall consider another category of modulation—that of pulse modulation.

From the sampling theorem, we know that to convey the information contained in a band-limited signal it is necessary to send only a finite number of discrete samples. A low-pass modulating signal that is band-limited to f_m Hz is completely specified by its values at intervals spaced no greater than $(2f_m)^{-1}$ seconds apart. We conclude that instead of transmitting the complete signal in analog form we really need to transmit only a discrete number of samples.

In pulse modulation, these discrete samples are used to vary a parameter of a pulse waveform. For example, one may vary the amplitude, width, or position of the pulse waveform in proportion to the sampled signal.

Consider a given analog signal $f(t)$ that is sampled at the uniformly spaced time intervals t_i, as shown in Fig. 7.1(a). A corresponding table of the sample values appears in Fig. 7.1(b).



(a)

i =	1	2	3	4	5	6	7
$f(t_i)$ =	4	3	1	1	3	9	2

(b)

Figure 7.1 A sampled signal and its table of sample values.

This discrete data is to be transmitted in such a way that when the signal is demodulated at the receiver the first output pulse is a 4, the second a 3, the third a 1, etc., until the entire table of sample values is reconstructed correctly at the receiver. There is no reason, of course, why we must use the same time scale for the pulse train as was used for the modulating signal. For instance, one could sample the signal $f(t)$ using 1000 samples taken over a period of 1 sec and then proceed to transmit this

data at a rate of one pulse per minute for the next 1000 minutes. On the other hand, one could just as well transmit the entire 1000 sample values in 1 μsec. (Such data slow-down and speed-up techniques have actually been used in satellite experiments.) However, unless we specify to the contrary, we shall not alter the time scaling in our discussion of pulse modulation and all signal transmissions will be in "real time."

Note that we may or may not be able to reconstruct the original signal $f(t)$ exactly at the receiver. The sampling theorem provides the necessary conditions for recovering $f(t)$ from the $f(t_i)$. If the sampling permits the determination of $f(t)$ from the transmitting table, then we wish to make sure that it can be determined at the receiver. Therefore correct sampling is a problem to be solved *a priori* at the transmitter and no correction can be attempted at the receiver for an inadequate sampling rate. Our primary concern here will be to devise a system that will use no more bandwidth than necessary to permit the transmission of a table of sample values via pulse modulation and best reconstruction of the table at the receiver.

7.1 PULSE-AMPLITUDE MODULATION (PAM)

In pulse-amplitude modulation (PAM) the amplitude of a train of constant-width pulses is varied in proportion to the sample values of the modulating signal. The pulses are usually taken at equally spaced intervals of time. An example of a PAM signal is shown in Fig. 7.2. Because the generation of PAM has similarities to sampling, we shall quickly review sampling before proceeding.

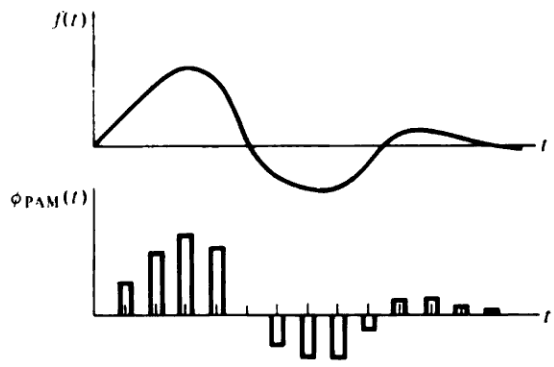


Figure 7.2 Pulse-amplitude modulation.

Consider a low-pass signal $f(t)$ that is band-limited to f_m Hz and multiplied by a periodic train of rectangular pulses $p_T(t)$, as shown in Fig. 7.3. The sampling interval T is taken as the Nyquist interval $(2f_m)^{-1}$ seconds. The sampled signal $f_s(t)$ is the product of $f(t)$ and $p_T(t)$:

$$f_s(t) = f(t)p_T(t) . \tag{7.1}$$

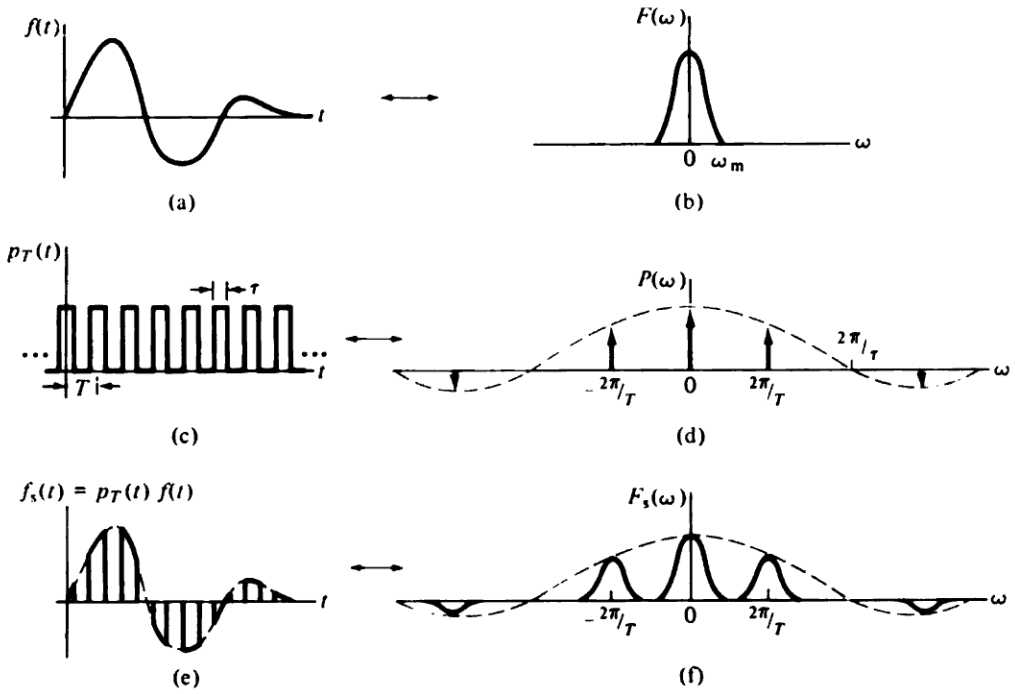


Figure 7.3 Natural sampling of a band-limited signal.

The spectral density of the sampled signal is found by taking the Fourier transform of Eq. (7.1):

$$F_s(\omega) = \frac{1}{2\pi} F(\omega) \circledast P(\omega) . \tag{7.2}$$

The convolution is easy to perform in this case as a result of the impulse functions and is shown in Fig. 7.3(f). The equivalent analytical expression for this result is

$$F_s(\omega) = \frac{1}{2\pi} F(\omega) \circledast \frac{\tau}{T} \sum_{n=-\infty}^{\infty} \text{Sa}\,(n\pi\tau/T) 2\pi\, \delta(\omega - n2\pi/T) , \tag{7.3}$$

$$F_s(\omega) = \frac{\tau}{T} \sum_{n=-\infty}^{\infty} \text{Sa}\,(n\pi\tau/T) F(\omega - n2\pi/T) . \tag{7.4}$$

We conclude from this brief review that the sampling of $f(t)$ results in the generation of spectral replicas at multiples of the periodic sampling rate. The sampling pulses do not have to be rectangular in shape and the particular choice of a pulse form will only alter the shape of the envelope of the spectrum of $F_s(\omega)$. What is important to us here is that each spectral replica generated by the sampling is an exact reproduction of the original spectral density $F(\omega)$ only displaced in frequency. The original signal $f(t)$ can be recovered from the sampled signal $f_s(t)$ by using an ideal low-pass filter.

In the case of natural sampling discussed above, the amplitudes of the pulses varied in proportion to the sampled values of the modulating signal $f(t)$. However, the pulse shapes also varied slightly, as a comparison of Figs. 7.2 and 7.3 will show. Specifically, the slopes of the pulse tops varied with the slopes of the modulating signal at the sample points in the case of natural sampling. In PAM, the pulse tops are flat. We shall investigate what effect this may have on the spectral characteristics.

First we let the sampling pulse width τ become very narrow so that essentially we have an impulse sampler, as shown in Fig. 7.4. Assuming unit area for each impulse, we have, as a special case of Eq. (7.4),

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n2\pi/T). \tag{7.5}$$

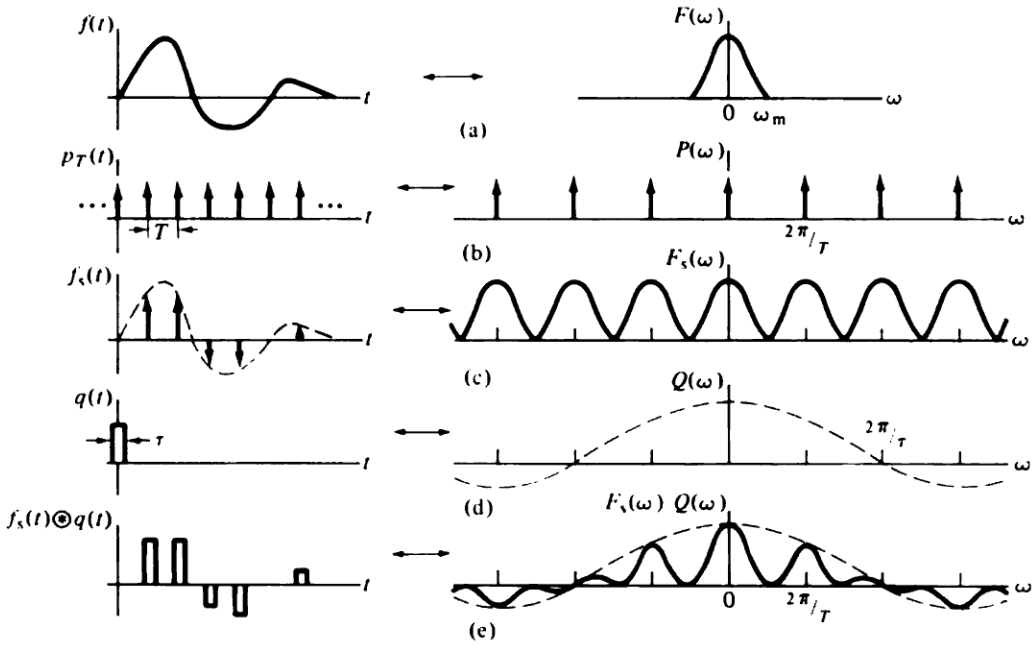


Figure 7.4 Spectrum of a PAM signal.

The impulse-sampler time waveform can be written as

$$p_T(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT). \tag{7.6}$$

The impulse-sampled waveform is, using Eq. (7.1),

$$f_s(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t - nT),$$

$$f_s(t) = \sum_{n=-\infty}^{\infty} f(nT)\,\delta(t - nT), \tag{7.7}$$

where the $f(nT)$ are the instantaneous sample values of $f(t)$. These impulse samples, weighted by the sample values of $f(t)$, are applied to a linear time-invariant filter with unit impulse response $q(t)$. The output of the filter is

$$\begin{aligned} f_s(t) \circledast q(t) &= \sum_{n=-\infty}^{\infty} f(nT)\,\delta(t - nT) \circledast q(t) \\ &= \sum_{n=-\infty}^{\infty} f(nT)q(t - nT)\,. \end{aligned} \tag{7.8}$$

Thus the impulse response of the filter, $q(t)$, can be chosen to approximate the desired output pulse shape. If $q(t)$ is the rectangular impulse response shown in Fig. 7.4(d), Eq. (7.8) represents the desired PAM pulse train. In this pulse train every pulse is rectangular in shape with an amplitude proportional to the sample value of $f(t)$ at the sampling points. The pulses are equally spaced in time.

Now that we are satisfied that we have generated a PAM pulse train, let us examine its spectral density. Using Eq. (7.5) and recalling that convolution in time is equivalent to a multiplication of spectral densities, we have

$$F_s(\omega)Q(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n2\pi/T)Q(\omega)\,. \tag{7.9}$$

This is illustrated in Fig. 7.4(e). But this spectral density is not the same as that which was obtained for the sampled waveforms in Fig. 7.3. In Fig. 7.3 the spectrum consisted of $F(\omega)$ repeated at multiples of the sampling frequency with only a gain variation of each spectral replica. This gain variation was determined by the shape of the sampling pulse, as expressed in Eq. (7.4). In contrast, Eq. (7.9) describes a point-by-point multiplication in frequency so that the spectral density $F(\omega)$ has lost its original shape. This distortion is dependent on the pulse shape; at low frequencies it is not severe if the pulse width is very narrow. Thus we conclude that there is a subtle difference between natural sampling with rectangular pulses and the generation of PAM with flat-topped pulses.

At this point, a question about why we are so interested in flat-topped pulses usually arises. The reason is that we do not need to use the shape of the pulses to convey information, and a rectangular pulse shape is an easy one to generate. When signals are transmitted over comparatively long distances, repeaters are often necessary to filter and amplify the signals before transmission to the receiver or next repeater. In analog CW modulation systems, as repeaters must amplify the signals faithfully, the effects of additive noise are compounded. In the type of pulse modulation system that we are considering here, the information is in the pulse amplitudes at the sample times only. Because pulse *shape* is not important, repeaters can regenerate rather than amplify the pulses. For example, a new pulse can be regenerated if we make the pulse amplitude proportional to the area of the pulse input detected over a fixed width or time interval. This pulse regeneration has some advantages in signal-to-noise ratio.

Next we shall consider the problem of recovering $f(t)$ from the PAM waveform. A possibility is to sample the PAM waveform with a periodic train of very narrow pulses (theoretically, impulses) and then use a low-pass filter to smooth the result. This sampling pulse train must be synchronized with the incoming PAM signal.

Another method of signal recovery, which was used in sampling, is to use low-pass filtering. However, because the low-frequency PAM spectrum is now given by $F(\omega) \times Q(\omega)$, it is not possible to recover $f(t)$ exactly using only a low-pass filter. The spectral weighting which has been introduced in the PAM process is known [that is, $Q(\omega)$] so the possibility of removing this effect does exist. Synthesizing a filter that has a transfer function of $Q^{-1}(\omega)$ will satisfy this goal. It generally is not possible to build such a filter over a wide range of frequencies, particularly if $Q(\omega)$ goes to zero. But because $f(t)$ is band-limited, it is sufficient to synthesize the inverse filter over only a very limited frequency range. We can combine the required low-pass filter with the inverse filter as shown in Fig. 7.5. The transfer function of the resulting filter is

$$H(\omega) = \begin{cases} Q^{-1}(\omega) & |\omega| < \omega_m \\ 0 & \text{elsewhere} \end{cases}. \tag{7.10}$$

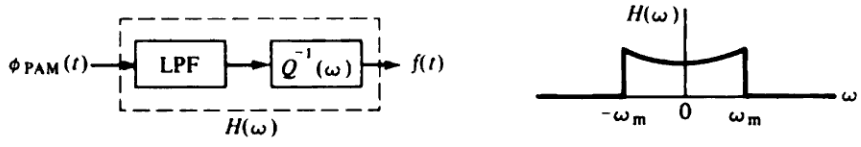


Figure 7.5 Recovering $f(t)$ from a PAM waveform using equalization.

This technique of correcting the frequency response of a system for a known distortion is called *equalization*. Equalization is often used in correcting distortions which are known but over which one has little control. For example, telephone lines are known to introduce both amplitude and phase distortions and these become objectionable when the lines become quite long. Repeater amplifiers spaced along the lines are used not only to amplify the signals to offset signal losses but also to correct for known distortions.

For rectangular pulse shape, the spectral density $Q(\omega)$ is a $(\sin x)/x$ pattern, as shown in Fig. 7.6. Because the signal has been sampled at the Nyquist rate, the sampling period T is related to the maximum frequency f_m by $f_m = 1/(2T)$.

As the pulse width τ is made smaller, the zero crossings in the $(\sin x)/x$ pattern of $Q(\omega)$ move farther out in frequency and thus $Q^{-1}(\omega)$ tends to become flatter within

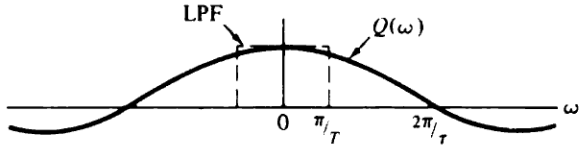


Figure 7.6 Equalization for a rectangular pulse shape.

the assigned bandwidth of the low-pass filter. Because T does not depend on τ, the ratio τ/T is a measure of the flatness of $Q(\omega)$ and $Q^{-1}(\omega)$ within the bandwidth of the low-pass filter. For a rectangular pulse, it turns out that as long as $\tau/T \leq 0.1$, the maximum difference between $Q^{-1}(\omega)$ and the ideal low-pass filter over the required range is less than 1%. In practice, then, the equalization for PAM can usually be neglected as long as $\tau/T \leq 0.1$.

7.2 TIME-DIVISION MULTIPLEXING (TDM)

The use of fairly short pulse widths in PAM signals leaves sufficient space between samples for insertion of pulses from other sampled signals. The method of combining several sampled signals in a definite time sequence is called *time-division multiplexing* (TDM). We shall discuss the principles of TDM here with particular reference to PAM although the principles apply as well to other types of pulse modulation.

Suppose we wish to time-multiplex two signals using PAM. Two alternative methods for accomplishing this are shown in Fig. 7.7. Usually digital logic circuitry is employed to implement the timing operations shown in these diagrams. The use of FET's at lower frequencies, and the diode ring samplers discussed in Chapter 5 at higher frequencies is popular in realizing the sampling operations. The commutator determines the synchronization and sequence of the channels (signals) to be sampled.

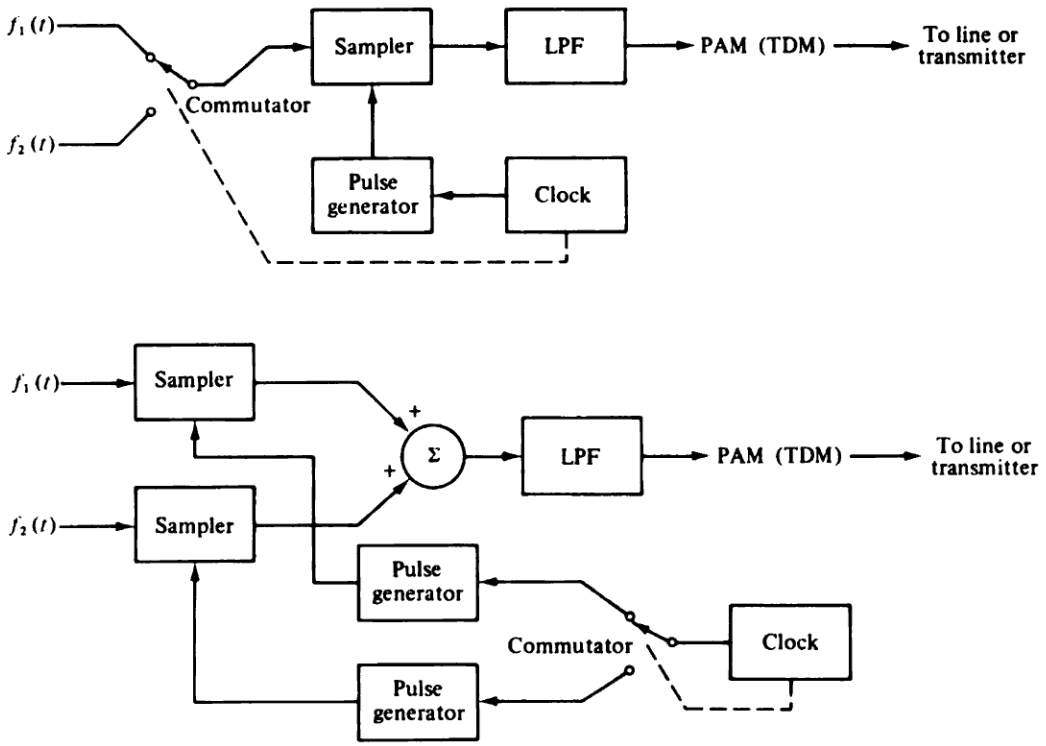


Figure 7.7 Generation of time-multiplexed PAM.

The pulse generator produces the narrow pulses desired to drive the sampler(s). The clock determines the timing of the overall system.

The two alternative methods shown in Fig. 7.7 basically differ only in that in the first system the commutator handles the analog signals before sampling and in the second system it handles the sampler control pulses. The second method is often preferred because it lends itself to digital logic circuitry even though it does require more samplers.

To illustrate the operation of time multiplexing it will be convenient to choose an example. Let us assume that both input signals $f_1(t)$, $f_2(t)$ are low-pass and band-limited to 3 kHz. The sampling theorem states that each must be sampled at a rate no less than 6 kHz. This requires a 12-kHz minimum clock rate for the two-channel system. The time-multiplexed PAM output might appear something like that shown in Fig. 7.8.

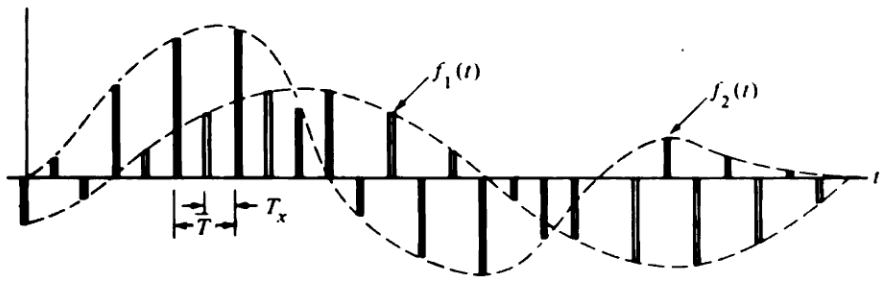


Figure 7.8 Time multiplexing of two PAM signals.

The time-multiplexed PAM signal could now be sent out on a line or used to modulate a transmitter (this will be discussed later in this section). But the bandwidth of this signal is very wide (theoretically infinite). Therefore we should decide just how much bandwidth is really necessary to transmit the information and to then correctly demodulate it at the receiver.

First, let us define T_x as the time spacing between adjacent samples in the time-multiplexed signal waveform (see Fig. 7.8). If all input signals (channels) are sampled equally, it is fairly obvious that $T_x = T/n$, where n is the number of input signals and T is the sampling period for one signal. Now we ask the question: If only the amplitude information is important here (not reproduction of pulse shaping), what is the absolute minimum bandwidth required such that the information in each sampled channel remains independent of that in the other channels? To answer this question, we appeal to the arguments of the sampling theorem. For an ideal filter which band-limits the spectral density to B_x Hz, the sampling theorem states that we must sample at a rate of $(2T_x)^{-1}$ samples per second in order to have independent samples. To prevent any irretrievable loss of information in the composite waveform then requires that the bandwidth B_x of the low-pass filter must satisfy the criterion

$$B_x \geq \frac{1}{2T_x}. \tag{7.11}$$

It should be emphasized that the use of this bound will not yield pulse shapes that closely resemble the pulses generated in the samplers. Pulse shape recognition requires additional bandwidth. What Eq. (7.11) does say is that independent amplitude information is present at the proper times if this condition is satisfied. Recall that the sampling theorem assumed the availability of ideal low-pass filters to satisfy the limiting condition.

A physical interpretation of Eq. (7.11) can be made as follows. After the two signals $f_1(t)$ and $f_2(t)$ in Fig. 7.8 have been sampled and combined, the resultant time-multiplexed signal can be considered as a new sampled signal, $f_3(t)$. The samples of the signal $f_3(t)$ are spaced at intervals of T_x seconds. The sampling theorem states that the highest frequency f_{max} which can be resolved in $f_3(t)$ without ambiguities is

$$f_{max} = \frac{1}{2T_x}.$$

In order to pass all of this possible range of frequencies, then, requires a system with a bandwidth satisfying Eq. (7.11).

We have preferred to find the bandwidth of the time-multiplexed PAM system in the time domain, just as we preferred to use the frequency domain for frequency multiplexing. At this point the unwary reader should be warned against some fallacious reasoning in computing minimum bandwidth for time-multiplexed PAM in the frequency domain. The line of reasoning could go something like this: Each input to the two-channel PAM system is band-limited to 3 kHz. The minimum bandwidth required for each channel is therefore 3 kHz, and if we add the two channels no new frequencies are generated so that we will need only 3 kHz. An extension leads to the ridiculous conclusion that no matter how many of these signals are multiplexed the composite waveform only requires 3 kHz! *This reasoning is wrong* — but why? A moment's reflection shows that this erroneous reasoning is based entirely on spectral magnitude considerations. However, *the phase relationships are essential to time multiplexing and cannot be ignored.* It is generally more convenient to work in the time domain than to work with these phase relations.

Some additional insight into the proper choice of bandwidth can be gained by assuming that the sample pulse widths are so narrow that they can be approximated by impulses. In addition, the filter is assumed to be an ideal low-pass filter with transfer function (see Fig. 7.9):

$$H_x(\omega) = \begin{cases} \pi/W & |\omega| < W, \\ 0 & \text{elsewhere}. \end{cases} \tag{7.12}$$

The impulse response of this filter is

$$h_x(t) = \mathcal{F}^{-1}\{H_x(\omega)\} = \mathrm{Sa}(Wt). \tag{7.13}$$

The effective width of this impulse response is $2\pi/W = 1/B_x$, as shown in Fig. 7.9(c).

The arrival of each impulse in the time-multiplexed waveform produces a replica of the impulse response of the filter, weighted and delayed by the area and timing of the impulse. Using superposition, the overall response is a sequence of $(\sin x)/x$ terms, as shown in Fig. 7.9(d). Because we have chosen the spacing between succesive

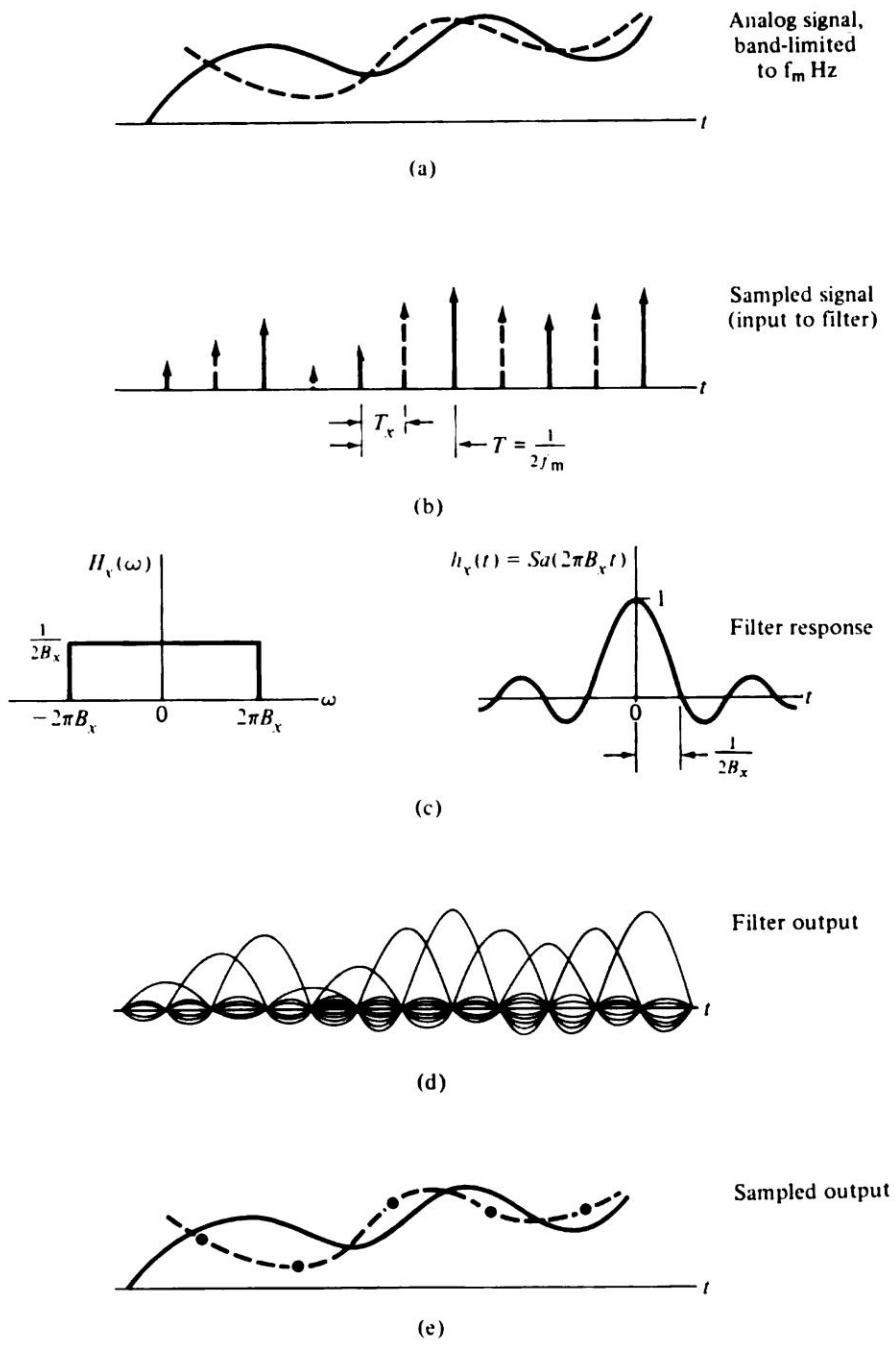


Figure 7.9 Ideal filtering of two impulse-sampled PAM signals.

samples to be $T_x = (2B_x)^{-1}$, contributions from all adjacent channels are exactly zero at the correct sampling instant. Therefore, by sampling the output at the correct instants, one can exactly reconstruct the original sampled values.

The equality sign in Eq. (7.11) then refers to the case in which impulse sampling and ideal filtering are used. Because neither of these conditions holds in practice, the requirement on the bandwidth must be relaxed somewhat. How much is a matter of engineering judgment and depends on the allowable cross-talk between channels. The

amount of signal from an adjacent channel which spills over into the desired time slot gives rise to this *interchannel cross-talk*. Interchannel cross-talk can be minimized by allowing more bandwidth than the minimum required and, to a lesser extent, by operating the different channels at comparable voltage levels.

Let us apply the results of the preceding discussion to the two-channel PAM example that we started out to analyze. Each input signal was band-limited to 3 kHz so that $T = (6\text{ kHz})^{-1} = 166.7\ \mu\text{sec}$, $n = 2$, and therefore $T_x = 83.3\ \mu\text{sec}$. Using Eq. (7.11), we find that the minimum bandwidth needed for the time-multiplexed signal transmission is

$$B_x \geq (166.7\ \mu\text{sec})^{-1} = 6\text{ kHz}.$$

From this result we learn that PAM is just as efficient in conserving bandwidth as SSB, at least when all input signals have the same bandwidth! This efficiency decreases when the bandwidths of the input signals differ. To conserve bandwidth, PAM time-multiplexed systems usually group input signals of comparable bandwidths.

At the receiver the composite time-multiplexed and filtered waveform must be resampled and separated into the appropriate channels. Once the pulses are separated, the normal sampling considerations apply and the analog reconstruction of the signals can be obtained by low-pass filtering. Block diagrams of two possible receivers are shown in Fig. 7.10. Operation of the sample-and-hold circuit shown will be discussed shortly.

Synchronization of the clock and the commutator in the time-multiplex receiver is necessary and can be achieved in various ways. One method used is to periodically send a pulse that is known to exceed the height of any other pulses in the transmission. These can be identified by the receiver and used for synchronization. Another possibility is to use a pilot tone that is shifted in frequency from the time-multiplexed transmission. A more elaborate scheme is to reserve a portion of the time-multiplexed transmission (e.g., one channel) to send some preassigned code which, when identified at the receiver, serves to synchronize the timing. This is considered in Section 7.13.

When a large number of signals are time-multiplexed together in a PAM system, the width of the sampling pulses must be made very narrow in comparison to the sampling interval. Under these conditions the signal power at the final low-pass filter in the receiver decreases to the point where large amplification factors are needed. An attractive alternative to this is to use a sampling switch followed by a capacitor, known as a *sample-and-hold* circuit and shown in Fig. 7.11. The switch closes only when that particular channel is to be sampled. If the source impedance r is small, the capacitor voltage changes to the input voltage within the time τ that the switch is closed. The load impedance R is arranged to be high so that the capacitor retains the voltage level until the switch is closed again. Therefore the sample-and-hold circuit accepts only those values of the input which occur at the sampling times and then holds them until the next sampling time. A low-pass filter is still needed to smooth the output but the requirements on filtering are less severe with the sample-and-hold circuit. Use of the sample-and-hold circuit results in a very efficient, reliable, and relatively noise-free PAM demodulator with no need for high amplification. Note that

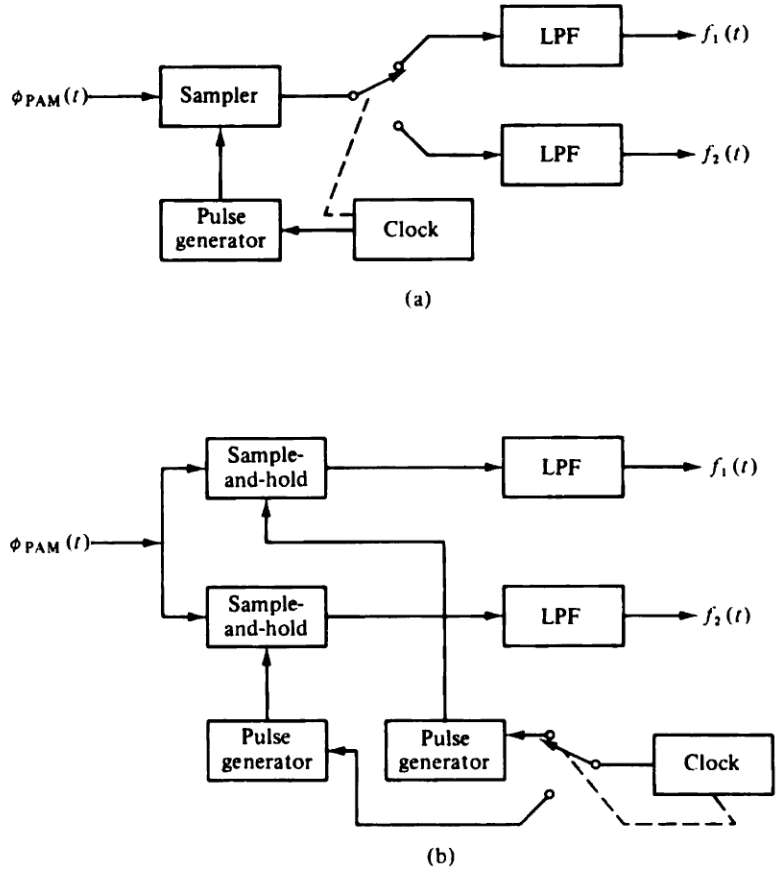


Figure 7.10 Receivers for time-multiplexed PAM.†

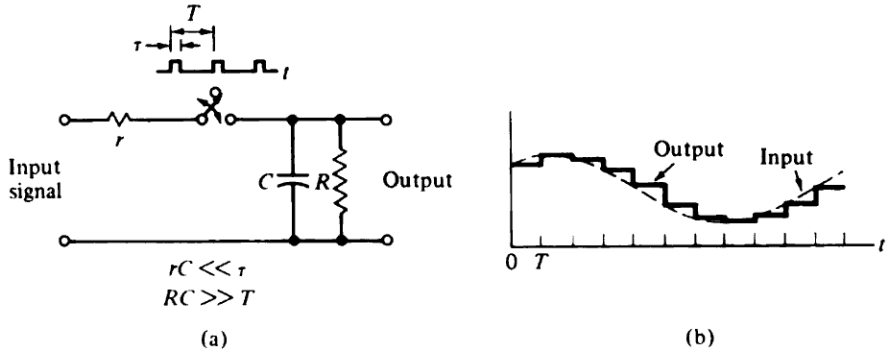


Figure 7.11 The sample-and-hold circuit and its use as a PAM demodulator.

the sample-and-hold circuit is an example of a linear time-varying system. Also note that, at least in this form, this circuit requires equalization to yield good signal fidelity. An alternative to the simple holding circuit is an interpolation algorithm that pro-

† In the second system, the pulse generators can be connected to form a ring. Each pulse generator operates only when it receives simultaneous pulses from the preceding generator in the ring and the clock. This eliminates the need for a separate digital divider (commutator).

vides an accurate smoothed estimate of the correct output based on observations of past sample values.

After time multiplexing and filtering, the pulse-modulated waveform may be transmitted directly on a pair of wire lines. For long-distance transmission, however, it is more convenient to use higher frequencies and transmit the signal using electromagnetic radiation. The PAM spectrum may be translated to these higher frequencies using the CW modulation techniques discussed in previous chapters. If amplitude modulation is used (DSB), the composite modulation is referred to as PAM/AM. If angle modulation is used, the systems are designated as PAM/FM or PAM/PM. A block diagram representation of a PAM/AM system is shown in Fig. 7.12.

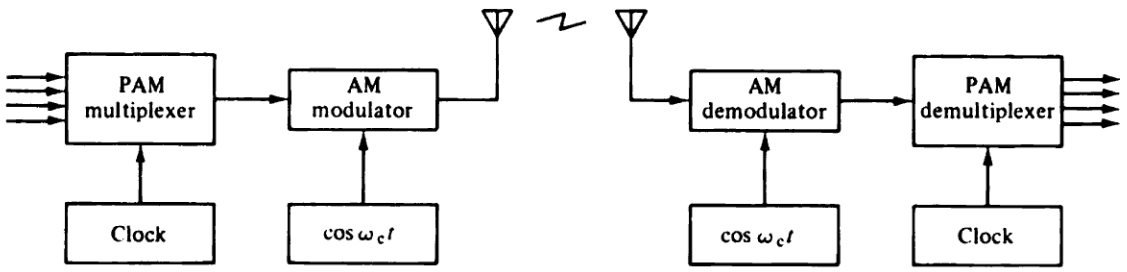


Figure 7.12 A block diagram representation of PAM/AM.

As noted earlier, time-multiplexed PAM is most efficient when all input signals are band-limited to the same frequency. In large multiplexing systems where differing data rates are to be handled, it is fairly common practice to group the input data channels by bandwidths. Each group is pulse-modulated and time-multiplexed and then these groups are frequency-multiplexed using subcarrier frequencies. Finally this composite signal is modulated using CW modulation. An example of a PAM/AM/FM system is shown in Fig. 7.13.

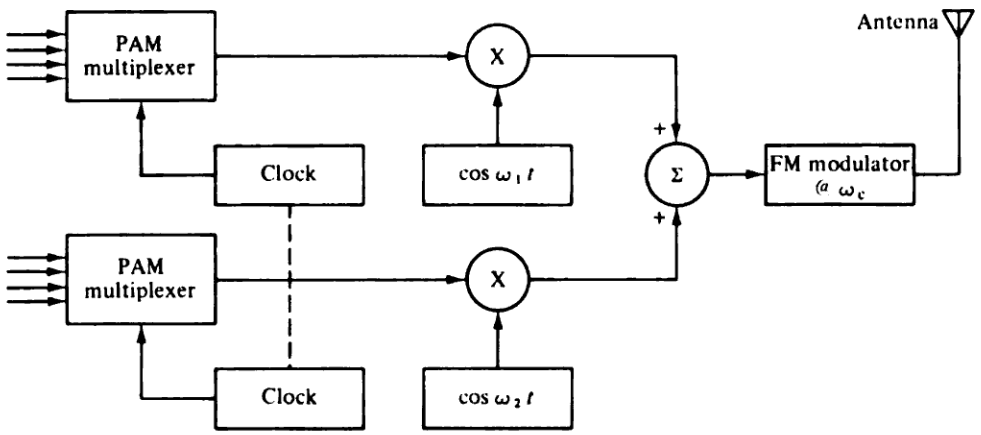


Figure 7.13 Example of a PAM/AM/FM transmitter.

Time multiplexing has been discussed here with particular emphasis on the use of PAM. The same principles apply to other types of pulse modulation.

Advantages in the use of pulse modulation with time-division multiplexing (TDM) include the fact that the circuitry required is digital, thus affording high reliability and efficient operation. This circuitry is simpler than the modulators and demodulators required in frequency-division multiplexing (FDM) systems. The multiplexing of many channels of relatively low-frequency data can be performed very efficiently using TDM when the inputs are all of comparable bandwidths.

Another advantage of TDM systems is the relatively small interchannel cross-talk arising from nonlinearities in the amplifiers that handle the signals in the transmitter and receiver. These nonlinearities produce harmonics that affect higher-frequency channels in FDM systems. Thus the phase and amplitude linearity requirements of the amplifiers in FDM become more stringent as more channels are added. In contrast, in TDM systems the signals from different channels are not handled simultaneously but are allotted different time intervals. Hence the linearity requirements do not become more stringent as the number of channels is increased and the effects of distortion are shared equally by all channels.

Disadvantages of TDM include the fact that pulse accuracy and timing jitter become a major problem at high frequencies so that TDM systems normally operate at clock frequencies below 100 MHz. In addition, time synchronization is required between transmitter and receiver.

EXAMPLE 7.2.1

Channel 1 of a two-channel PAM system handles 0–8 kHz signals; the second channel handles 0–10 kHz signals. The two channels are sampled at equal intervals of time using very narrow pulses at the lowest frequency that is theoretically adequate. The sampled signals are time-multiplexed and passed through a low-pass filter before transmission. At the receiver the pulses in each of the two channels are passed through appropriate holding circuits (i.e., sample-and-hold) and low-pass filters.

a) What is the minimum clock frequency of the PAM system?

b) What is the minimum cutoff frequency of the low-pass filter used before transmission that will preserve the amplitude information on the output pulses?

c) What would be the minimum bandwidth if these channels were frequency-multiplexed, using normal AM techniques and SSB techniques?

d) Assume the signal in channel 1 is $\sin(5000\pi t)$ and that in channel 2 is $\sin(10{,}000\pi t)$. Sketch these signals; sketch the waveshapes at the input to the first low-pass filter, at the filter output, and at the output of the sample-and-hold circuit and output of the low-pass filter in channel 2.

Solution a) In order to sample channel 2 adequately, we must take samples at a 20-kHz rate. Therefore the commutator clock rate $(T_c)^{-1}$ is 40 kHz and the commutator must recycle at a 20-kHz rate.

b) For the composite (interlaced) signal, $T_x = 25\ \mu$sec so that

$$B_x \geq \frac{1}{2T_x} = 20 \text{ kHz}.$$

c) For AM, min $B = 2(10 + 8)$ kHz $= 36$ kHz.
For SSB, min $B = (10 + 8)$ kHz $= 18$ kHz.
Note that SSB is more efficient in terms of the bandwidth required if the signals into the PAM system are not of equal bandwidth. For this reason, low data rate signals are usually time-multiplexed separately from medium data rate signals.

d) See Fig. 7.14.

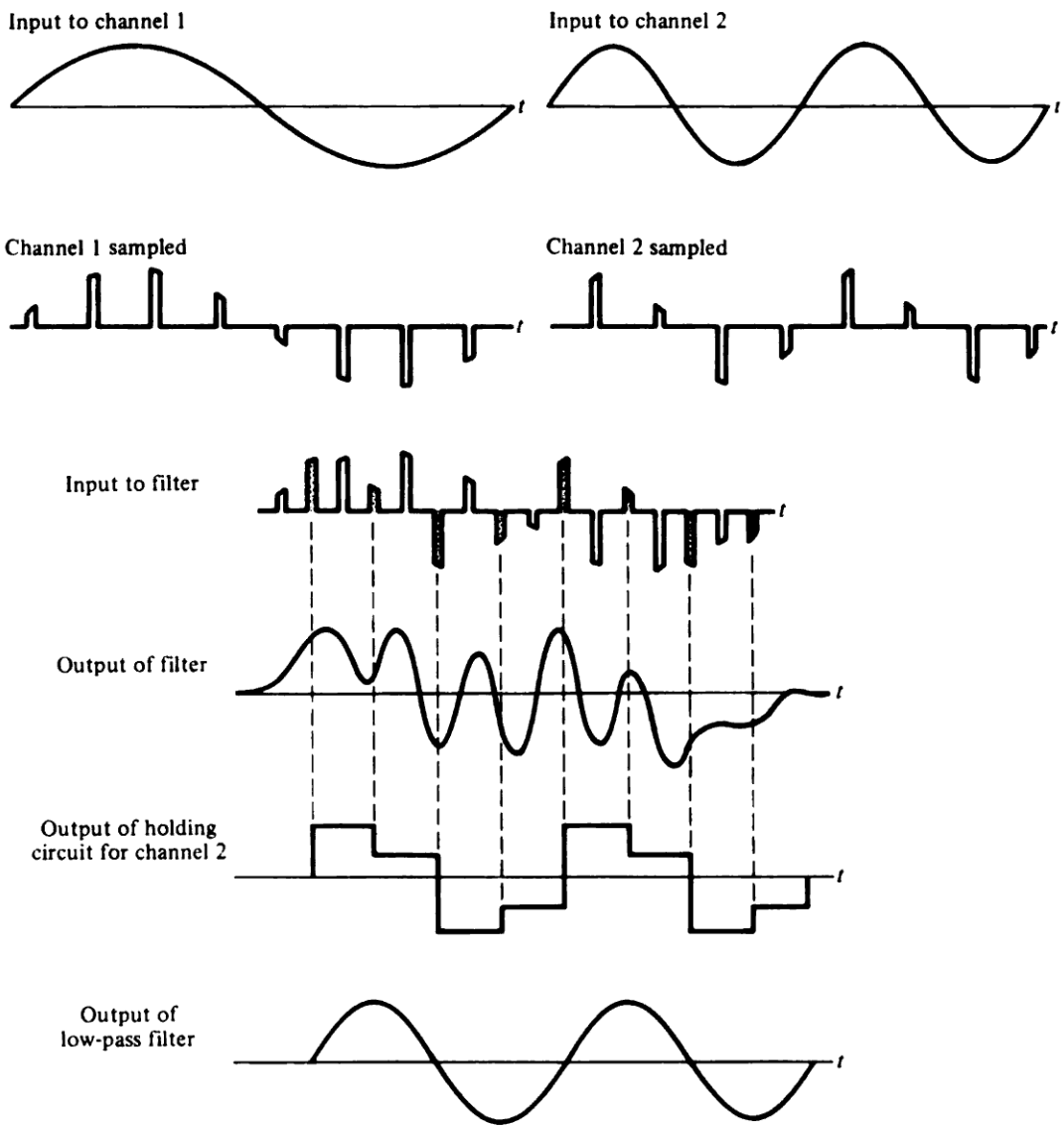


Figure 7.14 Waveforms corresponding to the PAM time-multiplexed system of Example 7.2.1.

DRILL PROBLEM 7.2.1

Two low-pass signals, each band-limited to 4 kHz, are to be time-multiplexed into a single channel using PAM. Each signal is impulse-sampled at a rate of 10 kHz. The time-multiplexed signal waveform is filtered by an ideal low-pass filter (LPF) before transmission.

a) What is the minimum clock frequency of the system?

b) What is the minimum cutoff frequency of the LPF?

c) In the receiver, the signal is impulse-sampled and demultiplexed using the system shown in Fig. 7.10(a). Determine the minimum and maximum acceptable bandwidth of the LPF used in retrieving the analog signals.

ANSWER: (a) 20 kHz; (b) 10 kHz; (c) 4 kHz, 6 kHz.

7.3 PULSE SHAPING AND INTERSYMBOL INTERFERENCE

In the preceding discussion we discussed the use of narrow pulse widths and assumed an ideal low-pass filter with bandwidth B_x. As each pulse passes through this low-pass filter, the corresponding output is a $(\sin x)/x$ waveform, appropriately scaled and delayed by the pulse amplitude and position [Fig. 7.9(d)]. With pulses spaced by $T_x = 1/(2B_x)$ seconds, the received waveform, when observed at the correct sampling instants, results in zero interference between adjacent pulses.

Attainment of the zero-interference condition between adjacent pulses requires precise synchronization between transmitter and receiver. It also demands that the transmission medium (or channel) behaves like an ideal low-pass filter. However, this is physically unrealizable and difficult to approximate in practical systems. Use of a realizable filter characteristic causes the latter assumption to be invalid, and errors in transmitter-receiver synchronization are difficult to eliminate. The result is that each received pulse is affected somewhat by the presence of adjacent pulses. In the case of multichannel pulse modulation, overlap between pulses gives rise to interchannel cross-talk, which in general is called *intersymbol interference* (ISI).

ISI can be decreased by purposely widening the transmission bandwidth. However, this alternative may be unnecessarily wasteful of bandwidth if not carefully controlled. Therefore we try to design received signal waveforms (and hence the transmission filters) to minimize the ISI within as small a transmission bandwidth as possible.

We begin by representing a received PAM signal $y(t)$ as[†]

$$y(t) = \sum_m a_m x(t - mT) + n(t), \tag{7.14}$$

[†] We assume the single-channel case here for convenience, but the results are readily extended to the multichannel TDM case.

where m is an integer and T is the sampling period. Here $x(t)$ represents the pulse waveform, a_m represents the sampled values of the input, and $n(t)$ is the additive noise. At a given time, $t = kT$, we have

$$y(kT) = a_k x(0) + \sum_{m \neq k} a_m x[(k - m)T] + n(kT). \tag{7.15}$$

In Eq. (7.15), the first term represents the kth sampled value of the input that has been transmitted. The second term in Eq. (7.15) arises from the overlap of other pulses adding to the desired pulse $a_k x(t - kT)$ at the kth sampling time. Thus this term represents the ISI present. The last term in Eq. (7.15) represents additive noise at the time $t = kT$.

The ISI given by the second term in Eq. (7.15) can be eliminated by choosing the received pulse shape $x(t)$ to satisfy the following condition:[†]

$$x[(k - m)T] = \begin{cases} 1 & m = k \\ 0 & m \neq k \end{cases}. \tag{7.16}$$

Note that the impulse response of the ideal low-pass filter in the preceding section (see Fig. 7.9) satisfies this condition. The uniformly spaced zeros in the impulse response at all nonzero multiples of the sampling periods yield zero ISI in this case. Because that filter is not physically realizable and is even difficult to approximate in practice, we modify the sharp magnitude-cutoff characteristic to attain a more gradual "roll-off" in frequency that does not affect the desirable uniformly spaced zeros in the impulse response. This is possible if the adjusted filter magnitude response has odd symmetry about the low-pass cutoff frequency.[‡] Satisfaction of this requirement gives rise to a class of waveforms known as *Nyquist waveforms*. In other words, all Nyquist waveforms have uniformly spaced zeros at all multiples of a basic interval except one (at the center).

One type of Nyquist waveform that is popular in both analytical work and practice is that of the raised cosine. A raised-cosine characteristic consists of a flat, or constant, magnitude at low frequencies and a roll-off portion that has a sinusoidal form with odd symmetry about the cutoff frequency. The raised-cosine characteristic can be expressed in equation form as[§]

$$X(\omega) = \begin{cases} T & 0 \leq |\omega| \leq (1 - \alpha)W \\ \dfrac{T}{2}\left\{1 - \sin\left[\dfrac{\pi}{2\alpha W}(|\omega| - W)\right]\right\} & (1 - \alpha)W \leq |\omega| \leq (1 + \alpha)W \\ 0 & |\omega| > (1 + \alpha)W \end{cases} \tag{7.17}$$

[†] This assumes the convenient normalization: $x(0) = 1$.

[‡] M. Schwartz. *Information Transmission, Modulation, and Noise*, Third ed. New York: McGraw-Hill, 1980, p. 181.

[§] Note that this characteristic could be expressed in terms of $\cos^2(\bullet)$.

where $W = \pi/T$. The parameter α is the excess bandwidth used divided by the minimum Nyquist bandwidth. Thus the case for $\alpha = 0$ coincides with the ideal low-pass filter used in the preceding section. The corresponding impulse response is

$$x(t) = \left(\frac{\sin Wt}{Wt}\right)\left(\frac{\cos \alpha Wt}{1 - (2\alpha Wt/\pi)^2}\right). \tag{7.18}$$

The first term in the right-hand side of Eq. (7.18) is the $(\sin x)/x$ waveform that we had used previously for the ideal low-pass filter, and it retains the original zero crossings of that waveform. The second term is a result of the more gradual spectral roll-off.

Plots of $x(t)$ and $X(\omega)$ are shown in Fig. 7.15 for three values of α. The case for $\alpha = 1$ is known as the full-cosine roll-off characteristic; its frequency transfer function is

$$X(\omega) = \begin{cases} \dfrac{T}{2}\left(1 + \cos \dfrac{\pi\omega}{2W}\right) & |\omega| \le 2W \\ 0 & \text{elsewhere} \end{cases}. \tag{7.19}$$

Note that for the $\alpha = 1$ case, $x(t)$ has zeros halfway between sampling times in addition to the sampling zeros. The amplitudes of the oscillatory tails of $x(t)$ are smallest when $\alpha = 1$ (i.e., the most gradual roll-off). Therefore this type of response results in lower ISI for a given timing error in sampling, even though it requires the transmission of some spectral components that are up to twice those required for the ideal low-pass case. The parameter α is called the *roll-off factor*.

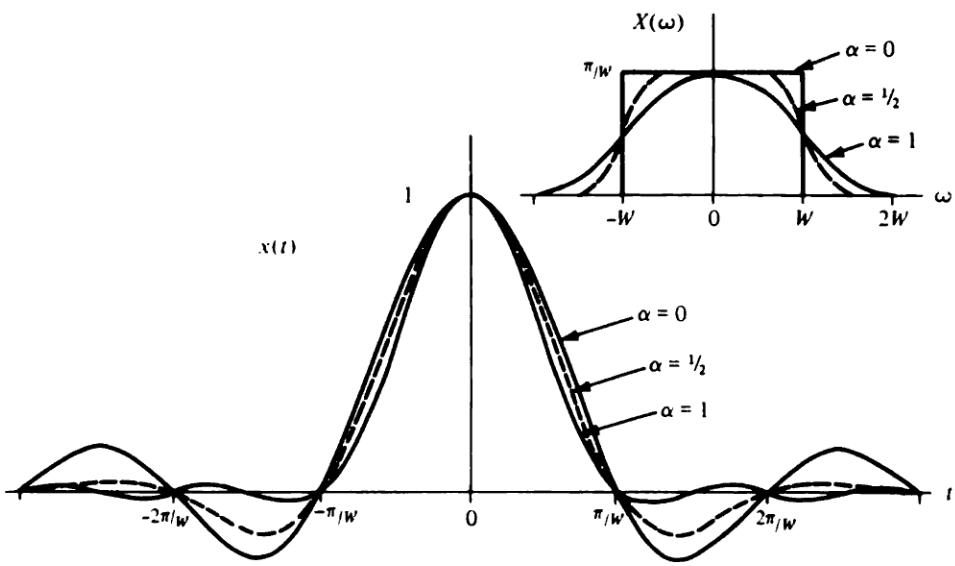


Figure 7.15 Raised-cosine pulse shaping (in frequency) and resulting time responses.†

† The ratio π/W can be written as $1/(2B)$ to make this agree more closely with Fig. 7.9(c).

It is instructive to learn why the raised-cosine waveforms are so effective in yielding zero ISI conditions. Consider the case shown in Fig. 7.16(a). Taking a derivative with respect to ω, we obtain the graph shown in Fig. 7.16(b). An important observation is that this graph can be decomposed into a convolution of a pair of oppositely directed impulse functions at $\omega = \pm W$ and a half-cosine (low-pass) frequency transfer function, as shown in Fig. 7.16(c). Using the convolution property that an integration of one function is equivalent to an integration of the overall result, we compensate for the derivative taken and arrive at the end result shown in Fig. 7.16(d). It is easy to show that the two plots in Fig. 7.16(d) give the two terms, respectively, in the impulse response given in Eq. (7.18).

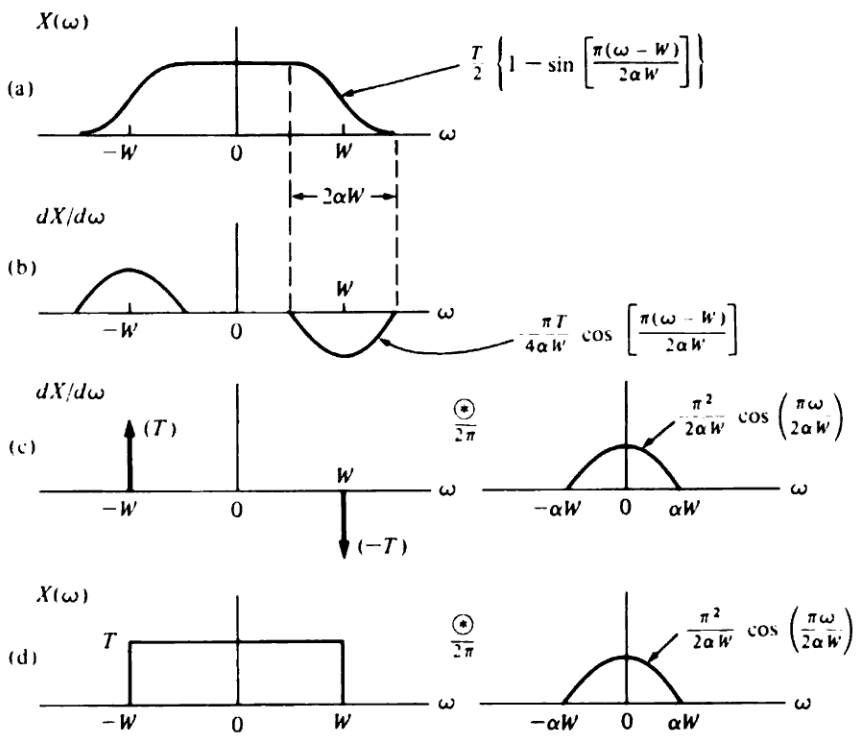


Figure 7.16 Decomposition of raised-cosine pulse shaping.

Note that the decomposition of the derivative of $X(\omega)$ into functions in which one is a pair of oppositely directed impulse functions at $\omega = \pm W$ guarantees that the characteristics of Eq. (7.16) are satisfied (within a constant). As a result of the complex conjugate symmetry property of the Fourier transform, the low-pass frequency transfer function must have even symmetry. Thus the convolution result has odd symmetry about the ideal low-pass filter cutoff frequency.

In summary, the bandwidth occupied for a raised-cosine-type transmission characteristic varies from a minimum of $B = 1/(2T)$ Hz ($\alpha = 0$) to a maximum of $B = 1/T$ Hz ($\alpha = 1$). The larger values of α lead to faster decaying pulses so that receiver synchronization will be less critical and modest timing errors will not cause large

amounts of ISI, but at the expense of added bandwidth. To control the amount of tolerable ISI, then, Eq. (7.11) should be modified to

$$B_x = \frac{(1 + \alpha)}{2T_x}. \tag{7.20}$$

So far we have considered those filters whose impulse response is the required pulse shape for zero ISI. In practice, the inputs are not impulses but finite-width pulses. Designating the Fourier transform of these finite-width pulses by $Q(\omega)$, we find that the magnitude of the frequency transfer function of the required shaping filter is

$$|H(\omega)| = \left|\frac{X(\omega)}{Q(\omega)}\right|. \tag{7.21}$$

As noted in Section 7.1, however, this correction is not significant if the pulse width is much narrower than the sample period.

Finally, note that, strictly speaking, none of the raised-cosine pulse waveforms are physically realizable. However, they can be approximated by causal filters if sufficient time delay (i.e., linear phase) is allowed.

EXAMPLE 7.3.1

Twenty PAM signals, each band-limited to 3 kHz and sampled at 8 kHz, are time-multiplexed prior to transmission. Choose a raised-cosine filter characteristic that will permit transmission of this multiplexed signal within an absolute maximum bandwidth of 120 kHz.

Solution Using Eq. (7.11), we find that the minimum bandwidth with an ideal LPF is $(2T_x)^{-1} =$ 80 kHz. Examination of the raised-cosine filter characteristic reveals that the frequency $(2T_x)^{-1}$ is set at the half-magnitude point. Therefore we set the half-magnitude point at 80 kHz and use $\alpha = (120 - 80)/80 = 0.50$.

7.4 OTHER TYPES OF ANALOG PULSE MODULATION: PWM AND PPM

Recall that to transmit a signal $f(t)$ that is band-limited to f_m Hz it is only necessary to transmit the information about its sample values at $(2f_m)^{-1}$-second intervals. We have discussed one way to accomplish this by using PAM and letting the amplitudes of a train of constant-width, uniformly spaced pulses vary in proportion to $f(t)$. An alternative modulation method is to vary some parameter in the timing of each pulse to convey the information. Pulse-timing modulation may be accomplished in a number of ways but the principles are basically the same and we only consider two specific types here.

One type of pulse-timing modulation uses constant-amplitude pulses whose width is proportional to the values of $f(t)$ at the sampling instants. This type is designated as *pulse-width modulation* (PWM).† Another possibility is to keep both the amplitude and the width of the pulses constant but vary the pulse positions in proportion to the values of $f(t)$ at the sampling instants. This is designated as *pulse-position modulation* (PPM). PAM, PWM, and PPM waveforms for a given $f(t)$ are shown in Fig. 7.17.

The various types of pulse-timing modulation are closely related to one another and one type can be derived from another. In a more general way, the relation of pulse-timing modulation to angle modulation is analogous to the relation of PAM to AM. In fact, one of the methods to generate wideband PM or FM is to first generate PWM or PPM.

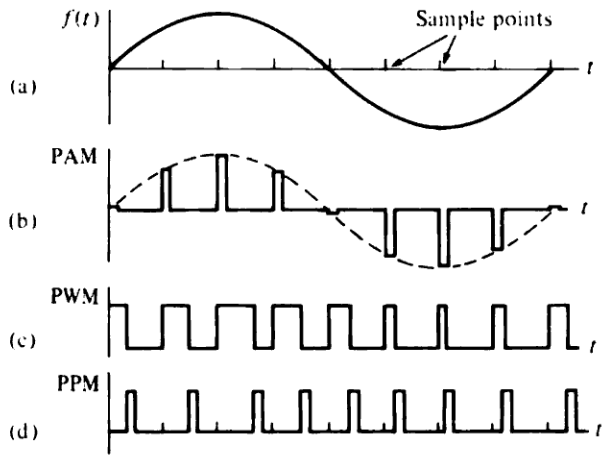


Figure 7.17 An illustration of PAM, PWM, and PPM.

In PWM, the signal $f(t)$ is sampled periodically at a rate fast enough to satisfy the requirements of the sampling theorem. At each sampling instant a pulse is generated with a fixed amplitude and a width that is proportional to the sample values of $f(t)$. A minimum pulse width τ_0 is assigned to the minimum value of $f(t)$. The variation of the pulse width from τ_0 is proportional to $f(t)$ and a modulation constant k_1 is defined in the proportionality. The pulse duration must be shorter than the time slot allocated to a particular sample and an additional guard time τ_g is usually allowed. The modulation in Fig. 7.17(c) is in the timing of the trailing edge of each pulse. Another version of PWM is to modulate both edges simultaneously (i.e., symmetrically).

PWM is a popular choice where the remote proportional control of a position or a position rate is desired. The average value of a PWM waveform varies directly with the modulation and can be used to control a motor by employing efficient switching-type operations. Proportional control can be maintained relatively independent of signal strength over a wide range. FM is also used for remote position control, but PWM systems are generally easier to build and align. Disadvantages of PWM include

† The designation *pulse-duration modulation* (PDM) is also used.

the necessity for detection of both pulse edges and a relatively large guard time. Effects of signal transients introduced in the system may vary with pulse width causing nonuniform performance.

Only the trailing edges of the PWM waveforms discussed above contain the modulating information. Therefore one could convey the information by only sending these timing marks. In PPM these are sent as constant-width, constant-amplitude pulses, as shown in Fig. 7.17(d). The minimum pulse delay is used to designate the minimum value of $f(t)$ and the change in delay is proportional to the modulating signal. The constant of proportionality is the modulation constant k_1. While generally more efficient for communications purposes than PWM, the use of PPM does require a method of regenerating the clock timing. In contrast, both PAM and PWM are "self-clocking"; that is, the clock timing is directly present in the modulated waveform.

Both PWM and PPM are nonlinear modulation methods and hence we cannot apply Fourier analysis directly. As in the case of angle modulation, however, we can obtain some helpful information.

In Chapter 3 we saw that band-limited systems have time responses with finite rise times. In pulse-timing modulation this finite rise-time constraint results in uncertainties in reconstructing the detailed structure of a signal. The uncertainty in specifying or reproducing a given signal amplitude is proportional to the rise time of the system and therefore inversely proportional to its bandwidth. This uncertainty is called the *resolution* of the system.

Without the deleterious effects of noise and distortion, the concept of resolution doesn't have much meaning; a given point on the rise time characteristic of the system can be chosen to indicate signal detection if the response is known exactly. In the presence of noise, however, the following criterion is quite generally used: *the time resolution of a pulse-timing system in the presence of noise is equal to the rise time of the system.*[†]

As a result of measurements on practical systems in the presence of noise, the rule that the two pulses must be spaced by at least the effective width of the system impulse response turns out to be a fairly good criterion. From our knowledge of linear time-invariant systems we recognize that the system rise time is a measure of the effective width of the impulse response.

This criterion can easily be demonstrated in practice by applying two narrow pulses to the input of a low-pass filter, as shown in Fig. 7.18. Here two pulses, each 0.3 μsec in width, are applied to the input of a fifth-order Butterworth low-pass filter

[†] This rule is not quite as arbitrary as it may appear. The rise time of a system is the indefinite integral of the impulse response. The question we are really asking, then, is how close can two impulses of equal weighting be at the input of the system before one cannot tell that there were two and what their separation would be by observing the output. The problem is a classical one first studied in the separation of point images in optical telescopes. The Rayleigh criterion, as formulated in these studies, states that two point images are just resolvable when they are separated by the effective width of the impulse response of the observing system. What we have stated here is essentially the Rayleigh criterion.

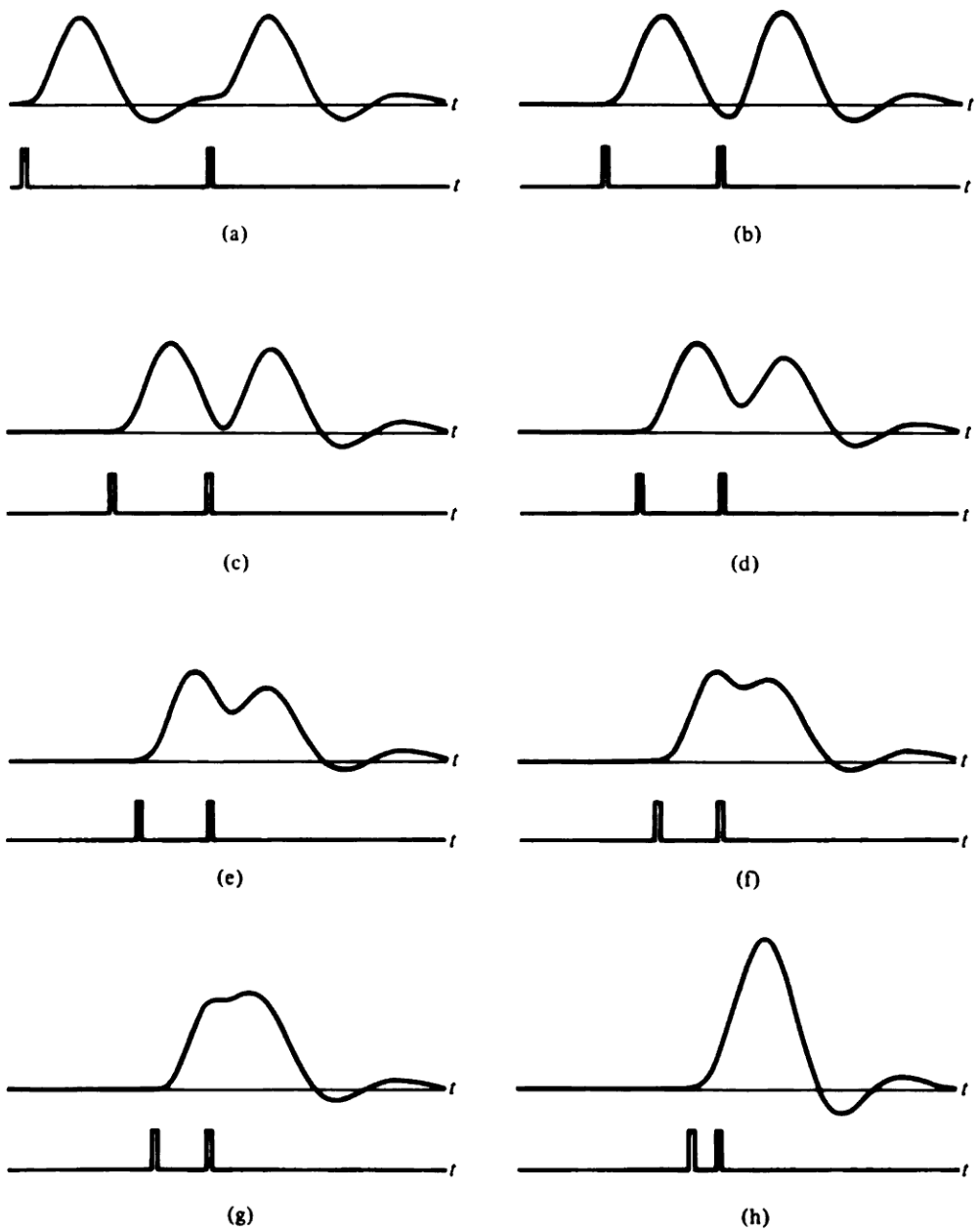


Figure 7.18 Pulse resolution of a low-pass filter.

with a −3-dB frequency of 80 kHz. (The rise time is about 6 μsec and the pulses are thus narrow enough to be considered impulses driving the system.) The input and output filter waveforms are shown for various pulse separations. The position of the first pulse is variable and the position of the second pulse is fixed. Some delay is noticeable in the filter response.

Starting with the waveforms in Fig. 7.18(a), note that the pulses and the corresponding responses are well separated. As the pulses are moved closer together, we

approach a condition in Fig. 7.18(e) in which the pulse separation is equal to the width of the impulse response of the filter. For pulse separations less than this, it becomes difficult to detect from the output response that there are two pulses present. This would be particularly true if some noise were added before the measurement was made.

We have specifically stated the above rule for practical systems. Recall that for an ideal low-pass filter the relationship between rise time and bandwidth is $t_r = 1/B$, where t_r is the elapsed time (in seconds) between the minimum and maximum filter output and B is the filter bandwidth (in Hz). However, for practical filters the bandwidth is usually defined on the basis of the -3-dB criterion. The rise time is defined as the time elapsed for the leading edge of a pulse to go from 10% to 90% of its final value when a step function is applied. In terms of these quantities, a good rule of thumb for practical low-pass filters is $t_r \approx 1/(2B)$. Because our criterion for pulse-edge response used above is based on the response of practical systems, we shall use this latter result for pulse-timing systems. Therefore the rule-of-thumb criterion for minimum detectable timing accuracy $\Delta\tau$ in the presence of additive noise is

$$\Delta\tau \approx 1/(2B) \qquad (\text{low } S/N)\,, \tag{7.22}$$

where B is the bandwidth of the system.

The criterion in Eq. (7.22) turns out to be a pessimistic one for high signal-to-noise ratios. To derive an equation for this case, we assume that the system rise time can be approximated by a straight line. This is a fairly good approximation for an ideal filter (cf. Section 3.13). Using this approximation, we find that each pulse after filtering is trapezoidal in shape (cf. Fig. 7.19). For high signal-to-noise, the additive noise only shifts the trapezoidal waveform in amplitude, and the timing accuracy can be written from the geometry of Fig. 7.19:

$$\frac{\Delta\tau}{t_r} = \frac{n}{A}\,,$$

or

$$\Delta\tau = t_r[\overline{n^2(t)}/A^2]^{1/2}.$$

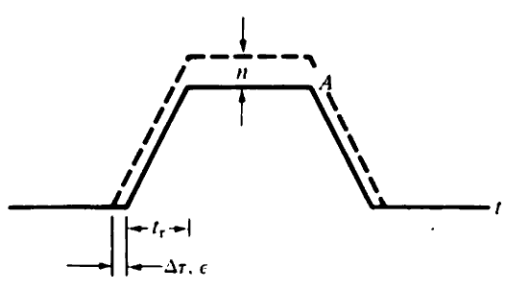


Figure 7.19 Variation in pulse position as result of additive noise.

For the ideal low-pass filter, $t_r \approx 1/B$ so that

$$\Delta\tau \approx \frac{1}{B\sqrt{[S/N]_{pk}}} \qquad (\text{high } [S/N]_{pk}), \tag{7.23}$$

where $[S/N]_{pk}$ is the peak pulse power to average noise power ratio. Note that the "rule-of-thumb" result in Eq. (7.22) corresponds to $[S/N]_{pk} = 4$ (that is, 6 dB). This value is often taken as the detection threshold for pulse timing modulation. We use the threshold value in system calculations unless otherwise specified.

EXAMPLE 7.4.1

A signal that is band-limited to 1 kHz with an amplitude range of -1 to $+1$ V is to be sent with a resolution of ± 0.5 mV using PWM. The minimum pulse width is 1 μsec and the guard time used is equal to the minimum pulse width. Estimate the bandwidth required for an input $[S/N]_{pk} = 20$ dB.[†]

Solution An amplitude-timing diagram for this system is shown in Fig. 7.20. Using this diagram as a guide, we have the following. The sampling time is $(2\text{ kHz})^{-1} = 500\ \mu$sec.

Time available for modulation: $500 - 2 = 498\ \mu$sec.

Modulation constant: $k_1 = 498/2 = 249\ \mu$sec/V.

Timing accuracy required: $\Delta\tau = k_1 \times (\text{signal resolution})$
$= (249)(10^{-3}) = 0.249\ \mu$sec.

Bandwidth: $B \approx 1/(10\ \Delta\tau) = 402$ kHz.

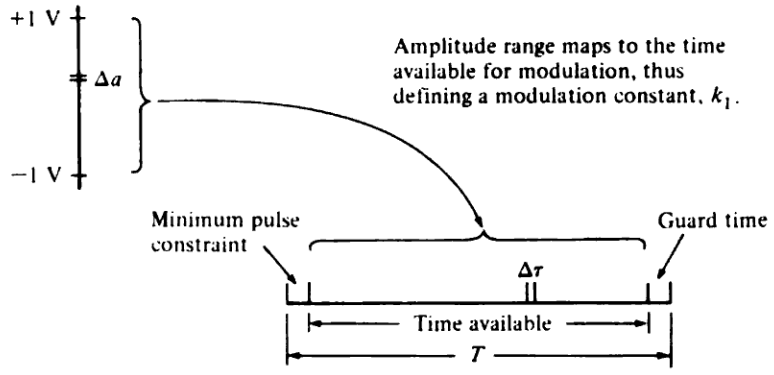


Figure 7.20 Amplitude-timing diagram for PWM system of Example 7.4.1.

[†] The reader may wish to investigate here the bandwidth-S/N trade-off for a given $\Delta\tau$ [cf. Eq. (7.23)].

DRILL PROBLEM 7.4.1

A given five-channel radio-control telemetry system uses PPM.† The time between consecutive samples of channel 1 is 100 msec; half of this time is used for synchronization (see Fig. 7.21). The pulse width is 1 msec and a 1-msec guard time is used. (a) What minimum receiver bandwidth would you recommend if the desired resolution (of the servos) is to be ±2% of the maximum for $[S/N]_{pk}$ = 15 dB? (b) What is the maximum frequency response of each channel?

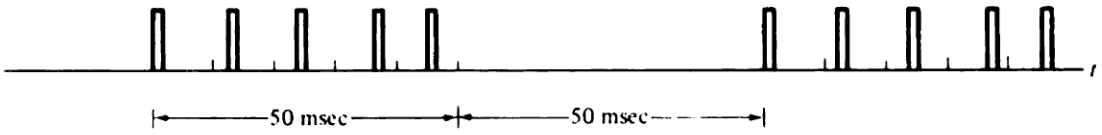


Figure 7.21 The PPM system of Drill Problem 7.4.1.

ANSWER: (a) 556 Hz; (b) 5 Hz.

EXAMPLE 7.4.2

A pulse radar uses pulse timing to determine range. A short burst of high-frequency electromagnetic energy is radiated from an antenna. The receiver amplifies and detects returned energy reflected from objects. Measurement of the time delay between transmission and reception is an indication of the range to the reflecting object. If the reflecting object is relatively small, the reflected waveform is a delayed replica of the transmitted waveform. Develop expressions for the range R and the range resolution ΔR if the measured time delay is τ seconds and the receiver bandwidth (low-pass) is B Hz.

Solution The transmitted pulse travels at the speed of light, c, to the reflecting object and back in the time τ so that

$$R = c\tau/2 = 150 \times 10^6\tau \qquad \text{m}.$$

Therefore 1 μsec of time delay corresponds to 150 meters of range.

† This type of system is commonly used for radio-controlled models. To simplify the receiver design and avoid deriving a clock signal for channel synchronization, each channel reference is triggered by the preceding one. This requires transmission of six pulses for five channels, but otherwise all operating principles remain the same as for conventional PPM.

The range resolution is related to the pulse timing by $\Delta R = (\Delta\tau)c/2$. The pulse timing accuracy is limited by the rise time of the receiver and $B \approx 1/(2t_r)$ so that

$$\Delta R \approx c/(4B) \qquad \text{m}.$$

Thus a receiver bandwidth on the order of 75 MHz is necessary to achieve a range resolution of 1 meter.

Generation of pulse-timing modulation commonly employs various combinations of a sample-and-hold circuit, a precision ramp voltage generator, and a comparator. The block diagram of a typical circuit for generating PWM and PPM is shown in Fig. 7.22(a). Operation of the sample-and-hold circuit was discussed in a previous section. The ramp generator produces a precision ramp voltage that has a peak-to-peak amplitude slightly larger than the maximum amplitude range of the input signals. This ramp voltage is the basis for the amplitude-to-timing conversion and therefore must be accurately known. It is resettable by the clock command. The comparator is a high-gain amplifier intended for two-state operation. If the input signal is higher than a preset reference level, the output is held in one state (i.e., a given voltage level). Whenever the input signal level is less than the reference level, the output

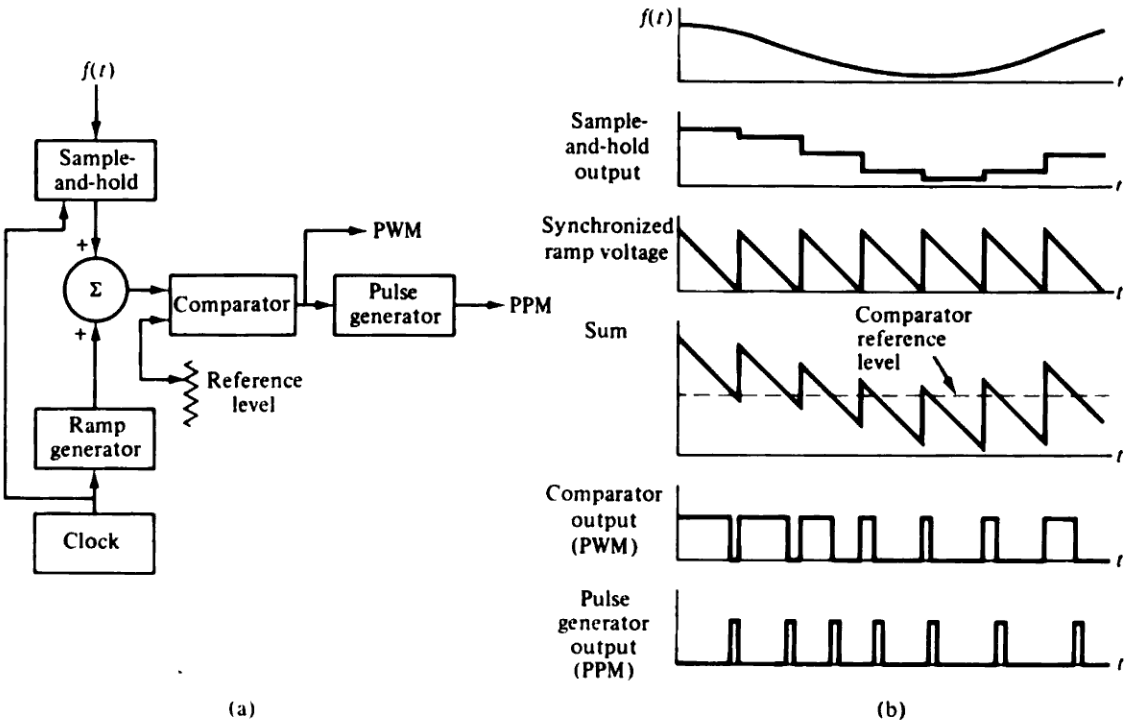


Figure 7.22 A method of generating PWM and PPM.

is held in the other state (voltage level). Which output state is present, then, depends upon whether the input is above or below the threshold (reference level) of the comparator. The transition between the two output states occurs very abruptly in good comparator designs.

Waveforms in the generation of PWM and PPM using this system are shown in Fig. 7.22(b). The voltage reference level of the comparator is adjusted so that there is always an intersection with the sum of the sample-and-hold circuit and the ramp voltage. In this system, the first crossing of the reference level indicates the clock timing and the second crossing generates the variable trailing edge. The constant amplitude of the output pulses is determined by the output voltage levels of the comparator. If the ramp voltage is linear, the output pulse train is PWM.

A convenient way to generate PPM is to use the PWM waveform generated above and then trigger a constant-width pulse generator on those edges of the PWM waveform with a negative slope. This is illustrated in Fig. 7.22(b). This pulse generator produces pulses of fixed width and amplitude whose time-of-occurrence relative to the clock is linearly related to the sampled values of the input signal.

Differing versions of these circuits are used in practice. Sometimes the sample-and-hold operation is deleted if the sampling rate is high. Note that modulation of the leading edge of the PWM waveform can be effected by reversing the slope of the ramp voltage. If a symmetric ramp voltage is used, both pulse edges in the PWM waveform are modulated symmetrically.

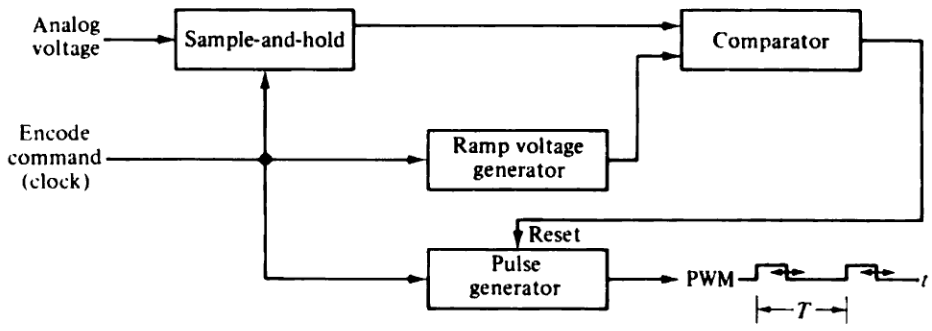


Figure 7.23 A popular method for generation of PWM.†

A popular circuit for modulation of the trailing edges in PWM is shown in Fig. 7.23. Here the ramp generator and the pulse generator are both started by the clock. When the ramp voltage exceeds the input voltage level, the comparator initiates the reset command to the pulse generator. The resulting waveform is a train of pulses whose leading edges are governed by the clock and whose trailing edges are governed by the amplitude of the input signal. Again, PPM is easily derived from this PWM waveform.

Some means of synchronization must be provided when several channels of pulse-timing modulation are transmitted using time multiplexing. A common method of

† The pulse generator is not really necessary but is used as a buffer stage.

implementing this is to periodically send pulses that are wider than any used for modulation. Use of an integrator and a comparator in the receiver can then distinguish the synchronization pulses and use them to synchronize the commutators. This scheme is particularly advantageous when using PPM because the change in pulse width can be made large to easily distinguish the synchronization pulses.

7.5 SIGNAL-TO-NOISE RATIOS IN ANALOG PULSE MODULATION

The pulse modulations we have discussed in preceding sections employ a variation of some parameter (e.g., amplitude, position, width, etc.) in a train of pulses in response to the amplitude of an input signal. This variation is an analog one—i.e., continuous in amplitude—even though the information is sent at discrete time intervals. Thus we classify these generally as *analog pulse modulation*.

Here we investigate the performance of analog pulse modulation systems in the presence of additive noise. In many respects, the modulation performance of PAM is analogous to that of AM, while the performance of PWM and PPM is analogous to that of FM and PM. Therefore it is not surprising that the signal-to-noise performance of PAM is the same as that for SSB-SC and that of PWM and PPM are closely related to those obtained for FM and PM.

7.5.1 PAM

Noise is added in the transmission of the PAM signal, as illustrated in Fig. 7.24(a). But the noise occurring between pulses adds noise power to the transmission without any increase in signal power. To avoid this, a synchronized gating circuit is used in the receiver to accept samples only when the signal is known to be present.

We shall assume that the signal and the additive noise present in the input to the PAM receiver are band-limited and that the conditions of the sampling theorem are

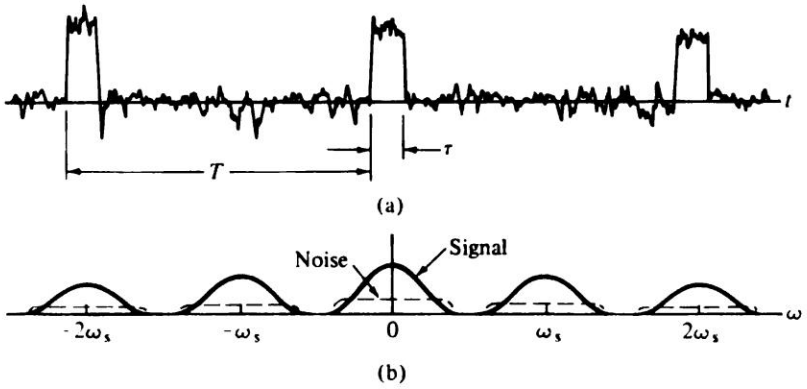


Figure 7.24 (a) Time waveform of a PAM signal with additive transmission noise; (b) spectrum of PAM with band-limited input signal and noise.

satisfied. Because the PAM receiver is linear, we can apply the signal and the noise separately, measure their powers, and then combine. The sampling and low-pass filtering at the receiver reproduce the band-limited signal and noise spectra within a constant, as shown in Fig. 7.24(b). Thus we can write

$$\overline{s_o^2(t)} = k\overline{s_i^2(t)},$$
$$\overline{n_o^2(t)} = k\overline{n_i^2(t)},$$

so that

$$\frac{S_o}{N_o} = \frac{S_i}{N_i}. \tag{7.24}$$

This can be extended to time-multiplexed PAM as long as the samples from adjacent channels are independent; i.e., if Eq. (7.11) is satisfied. Note that Eq. (7.24) is the same as that obtained earlier for SSB-SC systems [cf. Eq. (5.80)]. These results are on the basis of average power. The problem of maximizing the peak signal to rms noise ratio for each pulse will be considered in Section 7.11.

7.5.2 Pulse-Timing Modulation

In pulse-timing modulation (such as PWM or PPM), the information is contained in the relative positioning of the pulse edges. While we restrict our discussion to PPM systems here, the same general conclusions hold for other pulse-timing systems.

Although the pulses to convey the information may be generated with extremely short (fast) rise times, after passage through a band-limited system they have rise times that are governed by the bandwidth of the system. We shall approximate this rise time by a linear ramp, as shown in Fig. 7.19, so that the pulse assumes a trapezoidal shape.

The position of the trapezoidal pulse is sensitive to additive noise. If the noise voltage is assumed to vary slowly compared to the rise time of the pulse, the variation in the pulse amplitude, n, may be represented by a shift ϵ in the pulse position, as shown in Fig. 7.19. In fact, this figure shows that the pulse positional error ϵ arising from the additive noise voltage n may be reduced by increasing the slope of the pulse edge. For a given noise power, this requires a decrease in the rise time t_r—thus increasing the system bandwidth—or an increase in the pulse amplitude A—thus increasing the transmitter power. Therefore an exchange of bandwidth and transmitted power toward improvement of the S/N ratio is possible here, much as it was for angle-modulated systems.

We can obtain a simple expression to exemplify these conclusions if we make assumptions similar to those used in our analysis of FM. Specifically, we assume that the average output noise power is measured in the absence of modulation and for a large signal-to-noise ratio. Also, the output signal power is measured under noise-free conditions.

From the geometry of Fig. 7.19, we have

$$\frac{\epsilon}{t_r} = \frac{n}{A},$$

or

$$\overline{\epsilon^2} = (t_r/A)^2\overline{n^2}. \tag{7.25}$$

The output signal amplitude is proportional to the modulating signal $f(t)$ through a modulation constant k so that

$$s_o(t) = kf(t),$$

or

$$\overline{s_o^2(t)} = k^2\overline{f^2(t)}. \tag{7.26}$$

The output noise, when we use Eq. (7.25), is

$$\overline{n_o^2(t)} = \overline{\epsilon^2} = (t_r/A)^2\overline{n_i^2(t)}. \tag{7.27}$$

We also have

$$\overline{s_i^2(t)} = A^2(\tau/T), \tag{7.28}$$

and for the ideal filter (giving a nearly linear rise time),

$$B \approx 1/t_r. \tag{7.29}$$

Combining Eqs. (7.26)–(7.29), we have

$$\frac{S_o}{N_o} = \frac{k^2\overline{f^2(t)}}{\tau/T}B^2\frac{S_i}{N_i}, \tag{7.30}$$

or

$$\frac{S_o}{N_o} \propto B^2\frac{S_i}{N_i}. \tag{7.31}$$

Therefore the S/N improvement in a PPM system is proportional to the square of the bandwidth, just as it was for angle-modulation systems.

As in the case of angle modulation, however, the output signal-to-noise ratio in pulse-timing systems cannot be improved indefinitely. The noise power accepted by the receiver increases with the bandwidth, eventually becoming comparable to the signal level, violating our assumption of a large signal-to-noise ratio, and taking over the system. This threshold effect is found to take place for input signal-to-noise ratios on the order of 6 dB.

It is instructive to make a comparison of the S/N expressions for FM and PPM. Recall that for FM we obtained [cf. Eq. (6.115)]

$$\frac{S_o}{N_o} = 3\beta^2\frac{S_c}{N_c}.$$

For the wideband FM condition $\beta \approx B/(2f_m)$ so that

$$\frac{S_o}{N_o} \propto \left(\frac{B}{f_m}\right)^2 \frac{S_c}{N_c} \qquad \text{wideband FM}. \tag{7.32}$$

For PPM, the modulation constant k is proportional to the time allotted between pulses, which in turn is governed by the sampling theorem. Therefore $k \propto 1/f_m$ and we can rewrite Eq. (7.30) as

$$\frac{S_o}{N_o} \propto \left(\frac{B}{f_m}\right)^2 \frac{S_i}{N_i} \qquad \text{PPM}. \tag{7.33}$$

In both cases, the output voltage signal-to-noise ratio is proportional to the ratio of the transmission bandwidth B to the signal bandwidth f_m. The larger the ratio B/f_m, the greater the S/N improvement as long as operation is above the threshold level. Thus both FM and PPM are examples of analog signal transmission methods in which the rms noise improvement is linearly proportional to the ratio B/f_m.

7.6 PULSE-CODE MODULATION (PCM)

The preceding types of pulse modulation made use of the discrete time samples of analog signals. In these cases, the transmission is composed of analog information sent at discrete times. The variation of pulse amplitude or pulse timing is allowed to vary continuously over all values. A further refinement is to quantize the sampled analog signal into a number of discrete levels. This is sometimes done in PAM systems and is referred to as "M-ary PAM," where M designates the number of levels used. We now not only quantize the sampled analog signal into a number of discrete levels but also use a code to designate each level at each sample time. This type of modulation is called *pulse-code modulation* (PCM).

Suppose, for example, that we wish to quantize one-half a cycle of a 1-V (peak) sinusoid using eight discrete levels, as shown in Fig. 7.25. At each sample time we must decide which of these eight levels is the best approximation to the sinusoid. We choose the closest value (some other criterion can be chosen as well, depending on the particular circuit), and use this value until the next sample time, etc. Obviously, this process of *quantization* introduces some fluctuations about the true value; these fluctuations can be regarded as noise. An increase in the number of quantizing levels assigned will tend to decrease this quantization noise. The quantization may be linear (i.e., uniformly spaced levels), as shown in Fig. 7.25(a), or it may be nonlinear (i.e., nonuniform level spacings). Figure 7.25(b) illustrates a case in which the half-sinusoid is quantized using a square-law dependency for the level spacings. This type of quantization could be used to advantage, for example, in systems where the low levels are of less importance than the higher levels. Unless otherwise specified, quantization using equal level increments (i.e., linear) is usually assumed.

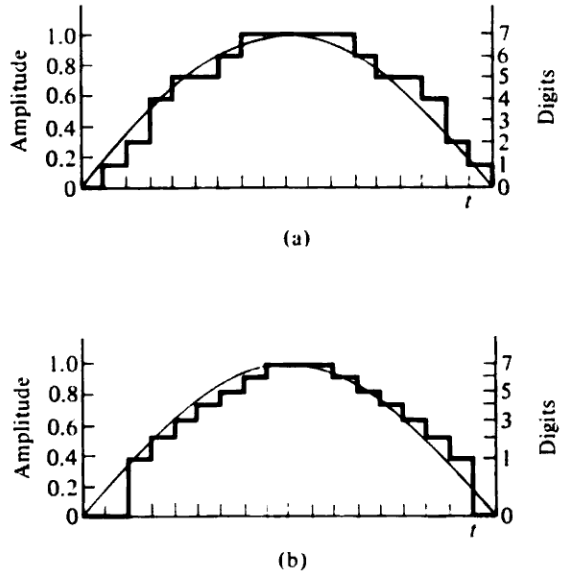


Figure 7.25 Quantization of half-sinusoid: (a) linear; (b) nonlinear.

The next step is to assign a digit to each level in such a way that each level has a one-to-one correspondence with the set of real integers. This is called *digitization* of the waveform. The process of digitization has reduced the waveform to a set of digits at the successive sample times, resulting in a completely digital modulation system. The digits are expressed in a coded form. The most common code used for this purpose is a binary code (i.e., a code using only two possible pulse levels). A binary code commonly used for an eight-level system is illustrated in Fig. 7.26.

Thus instead of transmitting the individual samples at the sample times, a pulse code or pattern is sent at each sample time to convey the information in quantized form. Systems making use of the transmission of digitized (i.e., quantized and coded) signals are called pulse-code modulation (PCM) systems.

Digit	Binary code	A binary pulse code
0	000	
1	001	
2	010	
3	011	
4	100	
5	101	
6	110	
7	111	

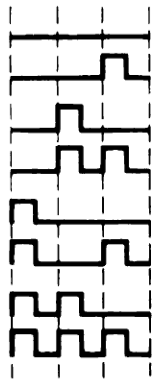


Figure 7.26 A binary pulse code.

Note that if the quantized samples were transmitted directly, the resulting system would simply be a quantized PAM (M-ary PAM) system. It is the assignment of the numbered signal levels which can be coded that gives PCM a distinct advantage over other types of modulation systems. Although we shall restrict our attention to the binary code here, one can also use pulse codes with more than two levels (called M-ary PCM).

In a binary code, each pulse can represent one of two possible states, e.g., a "1" or a "0". This basic quantum unit for conveying information is called a binary unit, or *bit*. For example, the eight-level binary code used above is referred to as a three-bit code. A four-bit code corresponds to 16 quantization levels, etc.

$2^3 = 8$
$2^4 = 16$

DRILL PROBLEM 7.6.1

Determine the number of quantizing levels that can be used if the number of bits in a given binary code is (a) 5; (b) 8; (c) x.

ANSWER: (a) 32; (b) 256; (c) 2^x.

PCM systems are becoming increasingly more important as a result of certain inherent advantages over other types of modulation systems. Some of these advantages are the following:

1. In long-distance communications, PCM signals can be completely regenerated at intermediate repeater stations because all the information is contained in the code. Essentially a noise-free signal is retransmitted at each repeater. The effects of noise do not accumulate and one need be concerned only about the effects of transmission noise between adjacent repeaters.
2. Modulating and demodulating circuitry is all digital, thus affording high reliability and stability, and is readily adapted to integrated-circuit logic design.
3. Signals may be stored and time scaled efficiently. For example, PCM data may be generated in an orbiting satellite once every minute during a 90-minute orbit and then retransmitted to a ground station in a matter of a few seconds. Digital memories can perform the required storage very efficiently.
4. Efficient codes can be utilized to reduce unnecessary repetition (redundancy) in messages. For example, if one wishes to send "A Merry Christmas and a Happy New Year" to a distant friend via Western Union, it is much more efficient to assign a code (i.e., a number) to this redundant message and then send the code (number). At the receiving station the decoder recognizes the code and types out the message.
5. Appropriate coding can reduce the effects of noise and interference. The communication-systems designer therefore has added versatility in the design of a PCM system to meet given performance criteria.

Now we shall devote some attention to the practical questions of generating PCM. The central operation is that of the analog-to-digital (A/D) converter; i.e., encoding analog signals into digital codes. We briefly describe several techniques for accomplishing this operation.

The first method is called the ramp encoder and uses the same basic principle as was used in the generation of PWM: generation of a linear ramp, which is compared with the input signal level. Instead of generating a pulse whose width is controlled by the time it takes for the ramp to reach the input-signal level as in PWM, however, we now operate a binary counter during this time and then read out the digital count. A simplified block diagram of the ramp encoder is shown in Fig. 7.27.

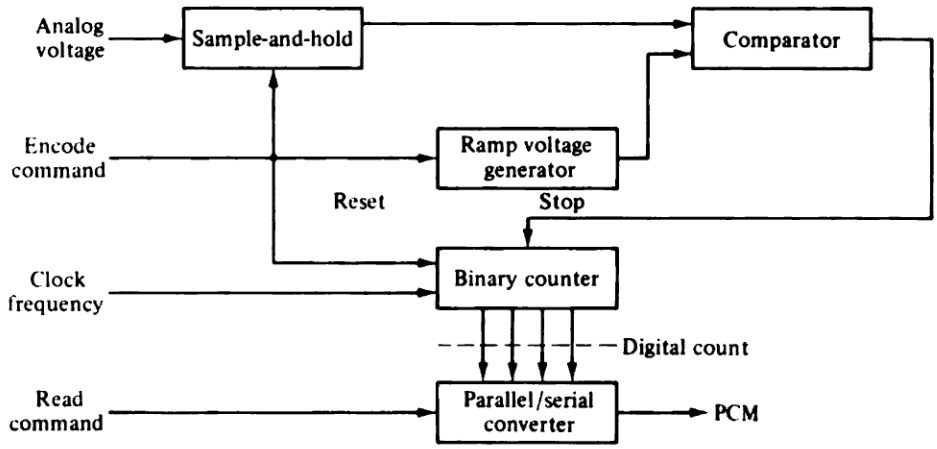


Figure 7.27 A simplified diagram of the ramp encoder for generation of PCM.

The ramp encoder uses relatively few precision components and is widely used. Its chief limitation is in the required linearity of the ramp voltage generated. Speed of operation is generally limited by the speed of the binary counter.

A popular version of the ramp encoder in instrumentation-type systems is the dual-slope (dual-ramp) encoder. In the dual-slope encoder, an integrator is inserted between the sample-and-hold and the comparator. The ramp voltage generator is not needed. As the output of the sample-and-hold is integrated, the slope of the integrator output is proportional to the input voltage. When the integrator output reaches a preset threshold, the input to the integrator is switched from the sample-and-hold to a precision voltage source with polarity opposite that of the signal. Conversion is completed when the integrator output returns to its initial starting level. Advantages of this method include good linearity and accuracy, excellent noise rejection, simplicity, and economy. A major disadvantage is the relatively long conversion time required.

A second method is the feedback encoder, shown in simplified block diagram form in Fig. 7.28. The feedback encoder operates by successively comparing the analog voltage input with a series of trial voltages. Successive trials are governed by the outcomes of previous decisions as to whether the trial voltage is greater or less than

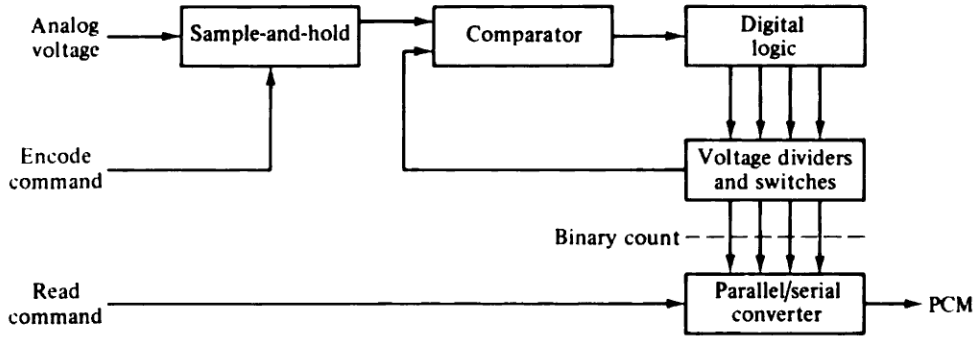


Figure 7.28 A simplified diagram of the feedback encoder for generation of PCM.

the level of the analog input. Usually the routine is designed to take one-half the remaining voltage interval for each trial so that the trials converge rapidly, following the binary code sequence to describe the input level. (This procedure is analogous to that of an analytical balance used for weighing objects, with the trial weights in this case arranged in powers of two.) The trial voltages are generated from a series of voltage dividers (a "ladder" network) with switches controlled by a digital-logic configuration. The code corresponding to the input voltage level can be read from the position of the switches in the voltage-divider network.

The accuracy of the feedback encoder is dependent on the accuracy of the trial voltages; speed is limited by the speed with which the voltage-divider switches can be operated. The feedback encoder requires more precision components than the ramp encoder, but it also allows a more accurate encoding.

A method of decreasing the conversion time is to apply the output of a sample-and-hold to many comparators operating in parallel, as illustrated in Fig. 7.29. Each comparator has its own reference. Outputs of all comparators are compared by digital logic circuits to determine the binary code output for each sample value. This parallel or *flash* encoder can be operated at extremely high speeds at the expense of providing $n - 1 = 2^m - 1$ comparators for m-bit quantization. The high parts count can be reduced somewhat by using a $(m/2)$-bit flash encoder feeding a high-speed D/A converter (DAC). The output of the DAC is subtracted from the input voltage, and the

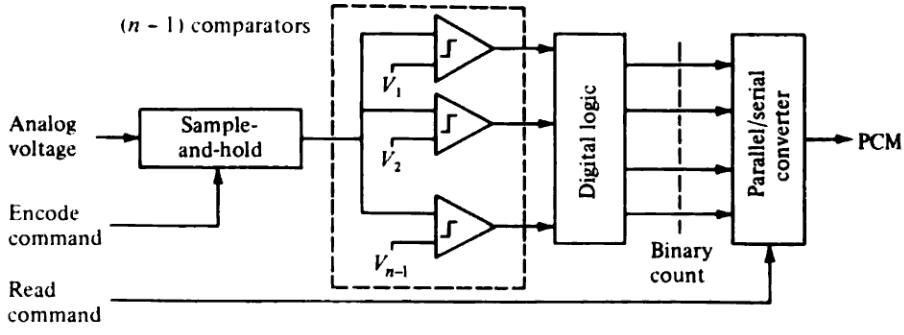


Figure 7.29 A parallel or flash encoder for generation of PCM.

difference between the two is converted by a second $(m/2)$-bit flash encoder. This feedback arrangement increases the conversion time somewhat, which may be a reasonable compromise if extremely high speeds are not required. A fourth method of A/D conversion will be covered in Chapter 9, where delta modulation methods are discussed.

We shall postpone discussion of the optimum reception of PCM in the presence of noise until later. If transmission noise is not a problem, or if constraints on the receiving system complexity do not allow an optimum receiver, a suboptimal design can be used. The block diagram of a suboptimal PCM receiver is shown in Fig. 7.30 (also see Problem 7.6.3). It uses some of the same ideas as the feedback encoder without the necessity for the feedback. Resistive dividers are operated by switches that are controlled by the binary code. By weighting the resulting voltage levels in the same proportions as the weights of the code pulses, we produce a composite voltage, which reproduces the original quantized voltage. This receiver design is simple and widely used in instrumentation systems. It is generally not used on long-distance communication systems, which attempt to maintain as low an error rate as possible.

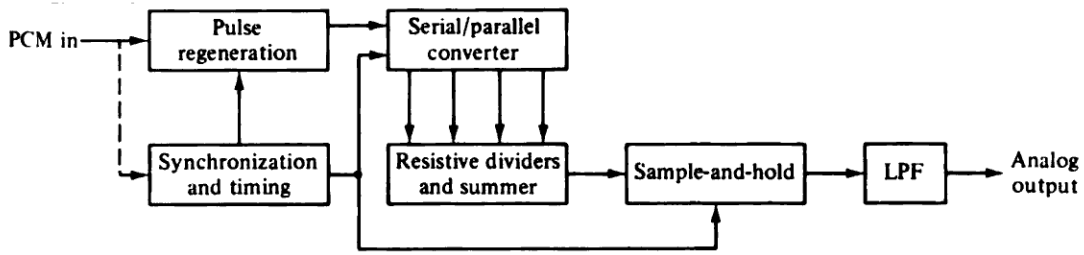


Figure 7.30 A suboptimal PCM decoder.

There are several levels of synchronization necessary in PCM systems. The first one is the clock rate, which can be sent as a separate frequency-multiplexed single-frequency tone or by using a phase-locked loop with a long time constant on the incoming PCM. Frame synchronization is necessary to establish when each pulse code is to begin. This is often accomplished by periodically sending a preassigned code that cannot be duplicated by any input level. When detected, this establishes the beginning of the frame. A divider on the clock frequency helps to maintain frame synchronization until the next synchronization arrives. A third level of synchronization is that needed to identify which channel is being sent when several channels of information are time-multiplexed. This is often combined with the frame synchronization.

In PCM telephone systems in the United States, 24 eight-bit voice channels are time-multiplexed to give 192 bits per frame.[†] One extra bit is inserted to give frame synchronization, yielding 193 bits/frame. Each input channel is sampled at an 8-kHz rate resulting in a frame interval of: $(8\text{ kHz})^{-1} = 125\ \mu\text{sec}$ and a clock rate of: $(193)(8\text{ kHz}) = 1.544 \times 10^6$ bits per second. Signaling (such as dial pulses) bits can

[†] See, for example, S. D. Personick, "Digital Transmission Building Blocks," *IEEE Communications Magazine*, vol. 18 (January 1980): 27–36.

be transmitted at a much slower rate and are either transmitted using one bit from each channel every sixth frame (yielding an effective $7\frac{5}{6}$-bit channel operation) or using a single data channel (yielding an effective 23-channel system). A diagram of the TDM format for one frame in this system is shown in Fig. 7.31.

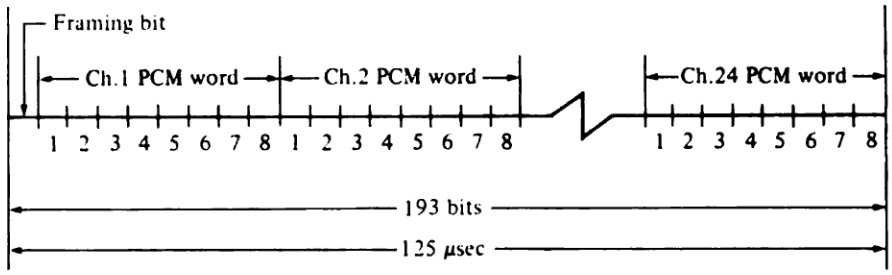


Figure 7.31 TDM frame format for the T1 system.

This 24-channel multiplexer is used as a basic system, known as T1, in the modular T-carrier TDM/PCM telephony system designed by the American Telephone and Telegraph Company. The T1 system is designed primarily for short-distance, heavy usage in metropolitan areas. In this system two twisted wire pairs are terminated at each end in a "channel bank" that combines the 24 voice channels using TDM/PCM as described above. Each wire pair carries data transmissions at a rate of 1.544 megabits per second (Mbps). Digital repeaters are spaced at approximately 1-mile intervals (35-dB loss spacings) along the line.

Maximum length of the T1 system is limited to about 50 to 100 miles because of timing jitter. It is designed to be compatible with existing PAM telephone systems. Transmissions that are digital (e.g., digital computer links) can be handled by replacing the channel bank with a data terminal.

The overall T-carrier system is made up of various combinations of lower-order T-carrier subsystems in order to accommodate the requirements of voice channels, Picturephone® service, and commercial television network programming. A block diagram of the overall T-carrier system concept is shown in Fig. 7.32 and the operating specifications of the various levels are listed in Table 7.1.

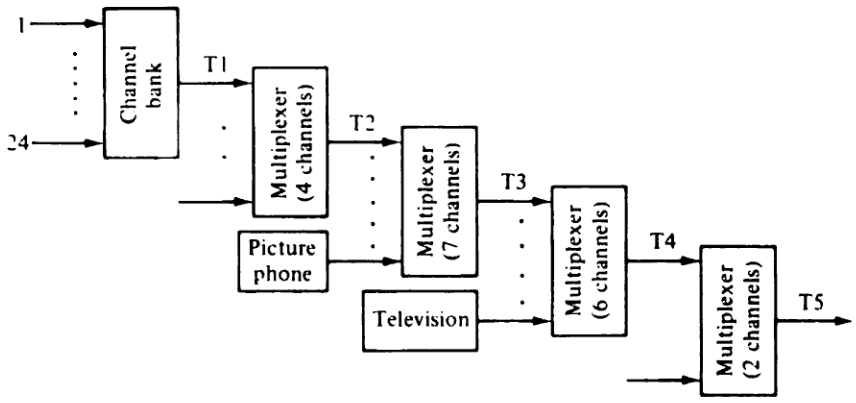


Figure 7.32 The T-carrier TDM/PCM telephony system.

TABLE 7.1 Specifications for the T-carrier TDM/PCM Telephony System

System[a]	Rate (Mbps)	MUX[b]	System capacity: Voice Channels	Picture-phone	TV	Medium	Repeater Spacing (miles)	Maximum System Length (miles)	System Error Rate
T1	1.544	T1	24	—	—	Wire pair	1	50	10^{-6}
T2	6.312	4(T1)	96	1	—	Coax	2.5	500	10^{-7}
T3	44.736	7(T2)	672	7	1	Coax	[c]	[c]	[c]
T4	274.176	6(T3)	4032	42	6	Coax	1	500	10^{-6}
T5	560.160	2(T4)	8064	84	12	Coax	1	500	4×10^{-7}
WT4	18,500.0	58(T4)	233,000	2436	348	60 mm circ. waveguide	25	4000	10^{-8}

[a] The corresponding channel bank designation uses "D" in place of "T."
[b] Abbreviation commonly used for multiplexer.
[c] Multiplexing levels only; not intended for transmission.

As in other types of pulse modulation, the maximum attainable rates for PCM are limited by switching speeds. After the data have been encoded, final transmission is often accomplished after a second modulation process. The PCM signal can be used to modulate a high-frequency carrier using AM, FM, or PM techniques. It is also possible to have a multichannel PCM system in which each of the pulse sequences modulates a subcarrier. These subcarrier signals are, in turn, frequency-multiplexed and then used to modulate a carrier.

Before the PCM signal is applied to a modulator, however, it may be put into several different forms. Choice of these different ways of conveying the binary code information will depend somewhat on the type of modulation and demodulation employed and other constraints on bandwidth, receiver complexity, etc. Figure 7.33 illustrates some of the more commonly used PCM representations. We briefly describe each of those shown.

The return-to-zero (RZ) method represents a 1 by a change to the 1 level for one-half the bit interval, after which the signal returns to the reference level for the remaining half-bit interval. A 0 is indicated in this method by no change, the signal remaining at the reference level. This is the representation used in the binary code shown in Fig. 7.26.

In the return-to-bias (RB) method, three levels are used: 0, 1, and a bias level. The bias level may be chosen either below or between the other two levels (it may be a zero reference, as shown in Fig. 7.33). The waveform returns to the bias level during the last half of each bit interval. The RB method has an advantage in being self-clocking; that is, the clock frequency is easily found from the magnitude of the pulse code. However, the average value of the waveform depends on the particular proportion of ones and zeros present. The RB representation also takes more bandwidth than necessary and it uses three levels. The succeeding representations are designed to overcome some of these drawbacks at the expense of increased complexity.

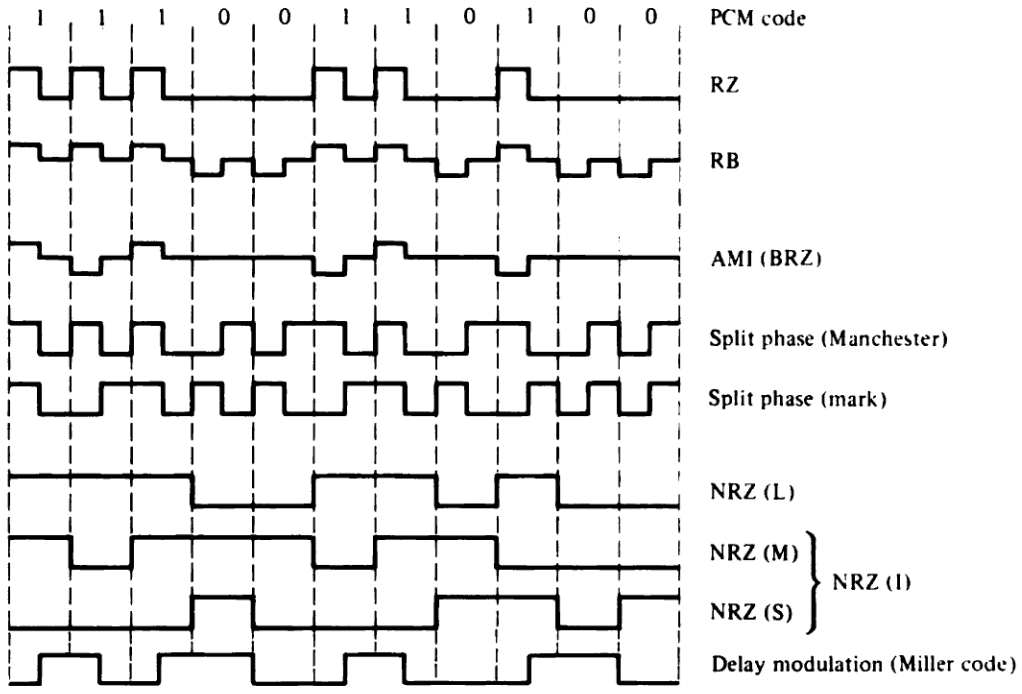


Figure 7.33 Some methods of representing binary data.

In Alternate Mark Inversion (AMI) the first binary one is represented by $+1$, the second by -1, the third by $+1$, etc. The AMI representation is easily derived from an RZ binary code (and vice versa) by alternately inverting the 1's. It has zero average value and is widely used in telephone PCM systems. This is also referred to as a bipolar return-to-zero (BRZ) representation.

The split-phase binary representations eliminate the variation in average value using symmetry. In the Manchester split-phase method, a 1 is represented by a 1 level during the first half-bit interval, then shifted to the 0 level for the latter half-bit interval; a 0 is indicated by the reverse representation. In the split-phase (mark) method, a similar symmetric representation is used except that a phase reversal relative to the previous phase indicates a 1 (i.e., mark) and no change in phase is used to indicate a 0.

The nonreturn-to-zero (NRZ) representations reduce the bandwidth needed to send the PCM code. In the NRZ(L) representation a bit pulse remains in one of its two levels for the entire bit interval. In the NRZ(M) method a level change is used to indicate a mark (a 1) and no level change for a 0; the NRZ(S) method uses the same scheme except that a level change is used to indicate a space (a 0). Both of these are examples of the more general classification NRZ(I) in which a level change (inversion) is used to indicate one kind of binary digit and no level change indicates the other digit. The NRZ representations are efficient in terms of the bandwidth required and are widely used. Note that use of split-phase and NRZ representations require some added receiver complexity to determine the clock frequency.

In delay modulation (Miller code), a 1 is represented by a signal transition at the midpoint of a bit interval. A 0 is represented by no transition unless it is followed by

another 0, in which case the signal transition occurs at the end of the bit interval. In this method, a succession of 1's and a succession of 0's each are represented by a square wave at the bit rate, but one is delayed a half-bit interval from the other. Delay modulation is insensitive to the initial 180° ambiguity that is present in NRZ and Manchester representations. It is efficient in terms of a required bandwidth without a required good low-frequency response.

Illustrations of sinusoids at the fundamental frequency for alternating levels in a PCM bit stream are shown in Fig. 7.34. The alternating-level sequence represents the most rapidly varying pattern possible in a binary code stream. In these diagrams, T_b represents the bit interval. Note that NRZ codes transmit one bit per level change while RZ codes transmit one bit per pair of level changes. It follows that the NRZ codes (and the Miller code) can transmit two bits per second per Hz (bps/Hz). In contrast, the limit for the other code representations depicted in Fig. 7.33 is 1 bps/Hz. We call the number of bits per second per Hertz the *bandwidth efficiency*. Therefore the potential bandwidth efficiency of a binary PCM code is 2 bps/Hz using an NRZ (or Miller) code representation.

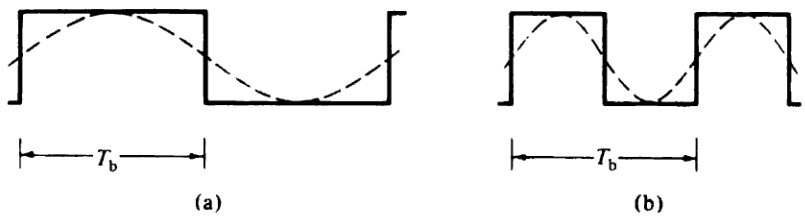


Figure 7.34 Maximum PCM bit rates of (a) NRZ and (b) RZ codes.

Let's next consider the minimum bandwidth required for a binary PCM system using n quantization levels. Let the number of bits per sample be designated by $[\log_2 n]$, where the use of the brackets indicates the next higher integer to be taken. Also, let the sample rate be $1/T$; then $[\log_2 n]/T$ is the number of bits per second to be sent.

From this discussion, we can see that the bandwidth efficiency for the NRZ code representation is 2 bps/Hz, so that the minimum bandwidth is

$$B \geq \frac{1}{2}\left(\frac{[\log_2 n]}{T}\right). \tag{7.34}$$

Extending this result to the time-multiplexed case, we have for the NRZ bandwidth

$$B_x \geq \frac{[\log_2 n]}{2T_x} \qquad \text{(NRZ)}. \tag{7.35}$$

Because the bandwidth efficiency for the RZ code representation is 1 bps/Hz, the minimum bandwidth required for RZ is

$$B_x \geq \frac{[\log_2 n]}{T_x} \qquad \text{(RZ)}. \tag{7.36}$$

A comparison of Eqs. (7.35), (7.36) with Eq. (7.11) reveals the following convenient rule: The minimum bandwidth required for a binary NRZ code representation is $[\log_2 n]$ times that required for a PAM system operating with the same sample rate and the same number of channels. The minimum bandwidth required for a binary RZ code representation is double this amount. These results can easily be extended to those systems that use more than two levels for allowable code representation levels. A logarithm to a different base accounts for the change, as shown in the following drill problem.

DRILL PROBLEM 7.6.2

Determine the minimum increase in bandwidth over that required by a PAM signal for the transmission of a PCM signal that has been quantized to 64 levels. Assume that each pulse in the code is allowed to take on the following number of levels: (a) 2; (b) 3; (c) x.

ANSWER: (a) 6; (b) 4; (c) $\log_x 64$.

Several channels using PCM codes such as those shown in Fig. 7.33 may be time-multiplexed and the resulting bit stream sent on a transmission line. If electromagnetic propagation is to be used, the bit stream is used to modulate a high-frequency CW carrier. One choice of modulation is that of amplitude-shift keying (ASK) in which the amplitude of a CW carrier is switched between two (or three) values in response to the PCM code. Another choice is to shift the frequency of the CW carrier, and this is called frequency-shift keying (FSK). A third possibility, called phase-shift keying (PSK), is that of shifting the phase of the CW carrier in response to the ampli-

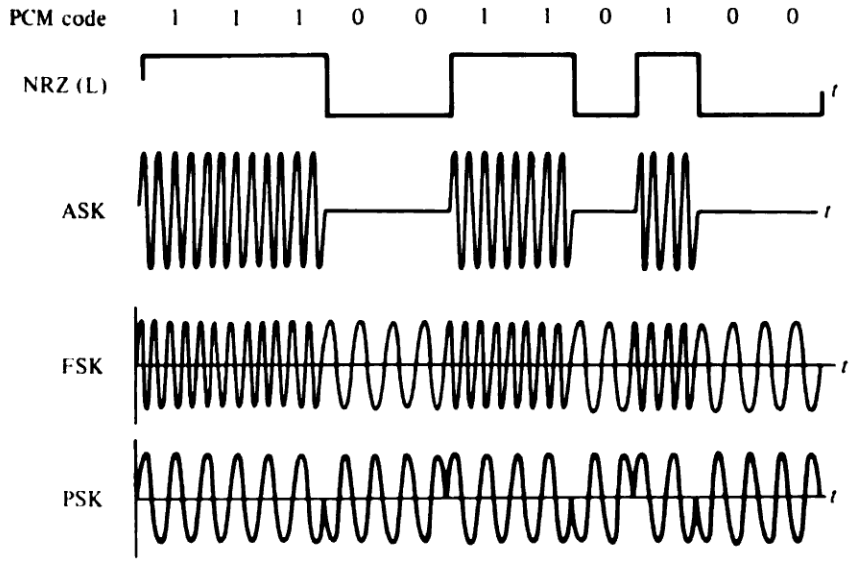


Figure 7.35 Idealized ASK, FSK, and PSK waveforms.

tude of the PCM code. These are illustrated in Fig. 7.35 for the PCM code used in Fig. 7.33. The S/N performance of these modulation methods will be discussed in Chapter 10.

7.7 FIBER OPTIC COMMUNICATION SYSTEMS

Although we shall examine modulation systems in Chapter 10 in which a carrier waveform is turned on and off in response to a binary code, we briefly describe here the specific application of fiber optics for that purpose. The use of fiber optics in communication systems is of rapidly increasing importance to the transmission of digital data.

The optical spectrum occupies that portion of the electromagnetic spectrum from a wavelength of about 50 nm (ultraviolet) to about 100 μm (far infrared). (This corresponds to frequencies of 6×10^{15} Hz to 3×10^{12} Hz, respectively; in the optical portion of the spectrum it is customary to specify the band of interest in terms of wavelength instead of frequency.) Because the usable information bandwidth of a bandpass communication system is typically limited to a fraction of the carrier frequency, optical systems offer an advantage of very wide potential bandwidths. For example, a system in which $B/f_c = 0.1\%$ operating at $f_c = 1$ GHz offers 1 MHz of potential bandwidth, while at the optical frequency of 10^{14} Hz such a percentage yields 100 GHz of bandwidth! Also, because optical fibers are dielectrics, transmission through fibers is immune to electromagnetic interference and differences in electric potentials. Early systems suffered from very high transmission losses resulting from impurities in materials, but these problems have largely been solved.

The block diagram of a basic fiber optic communication system is shown in Fig. 7.36. The transmitter consists of a light source and associated electrical drive circuitry. The transmission path consists of an optical fiber with its sheathing and environmental protection, and light couplers and connectors. The receiver consists of a photodetector, plus signal amplification and restoration circuitry.

The optical fiber is the central part of the optical communication system. An intuitive picture of the propagation in an optical fiber can be obtained using ray (geometric) optics concepts of light reflection and refraction. However, a more complete understanding of the optical wave propagation mechanisms depends on concepts from electromagnetic theory.

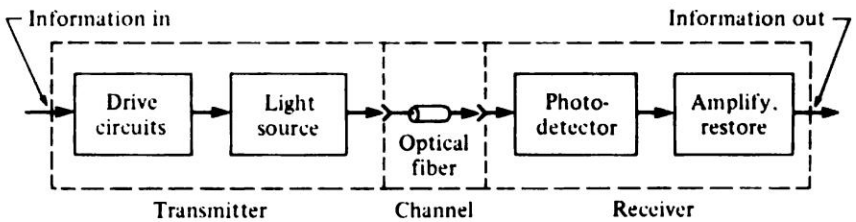


Figure 7.36 Block diagram of an optical fiber communication system.

In free space, light travels at a velocity of $c = 3 \times 10^8$ m/sec. Upon entering a dielectric it travels at a velocity v, which is a characteristic of the material (and is less than c). The ratio of the velocity in free space to that in the material is called the index of refraction, $n = c/v$.

In general, when a light ray is propagated toward a boundary separating two different media, part of the ray is reflected back and part is refracted (bent) as it enters the second material. Refraction of light at the interface is a result of the difference in the speed of light (refractive indexes) in the two materials. The relation at the interface, known as Snell's law, is (see Fig. 7.37):

$$n_2 \cos \theta_2 = n_1 \cos \theta_1 \qquad \text{(refraction)}, \tag{7.37}$$

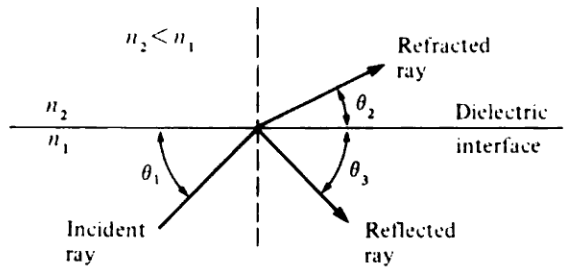


Figure 7.37 Refraction and reflection of a light ray at an interface between two materials.

where θ_1, θ_2 are the respective angles of incidence. In contrast, the reflected ray returns at the same angle as that of the incident ray, or

$$\theta_3 = \theta_1 \qquad \text{(reflection)}. \tag{7.38}$$

But as θ_1 is increased, and $n_1 > n_2$, it is possible to reach a condition in which $\theta_2 = 0$, $\theta_1 \neq 0$. Beyond this point no refraction is possible and all of the light rays are internally reflected within that medium with the higher refractive index. The angle of incidence at which this begins to occur is called the critical angle, θ_c. Using $\theta_2 = 0$ in Eq. (7.37), the critical angle is given by

$$\theta_c = \cos^{-1}\left(\frac{n_2}{n_1}\right). \tag{7.39}$$

For example, $n_1 = 1.50$ for glass and $n_2 = 1.00$ for air, so that the critical angle for this interface is: $\theta_c = 48.19°$. Any light in the glass incident on the interface at an angle $\theta < 48.19°$ is totally reflected back into the glass.

Although many different configurations are possible, the most widely used structure for an optical fiber is that of a solid dielectric cylinder, called the core of refractive index n_1, surrounded by a solid dielectric cylindrical shell, called the cladding of refractive index n_2 ($n_2 < n_1$). This configuration is called a step-index fiber, and is

illustrated in Fig. 7.38(a).† The cladding serves to add mechanical strength and protection to the core, as well as serving as its dielectric boundary. An important function of the cladding, for n_1 and n_2 nearly the same, is that the critical angle is very small (i.e., all rays are nearly parallel to the fiber axis). This makes the difference between path lengths of the central ray and the most oblique ray very small. The result is a reduction in the disparity of the time delays between different rays, and thus a reduction in the dispersion of a transmitted pulse. Another important factor is that light, which is supposedly confined to the fiber according to ray optics, actually propagates in both core and cladding according to electromagnetic theory. Therefore the cladding material must be a low-loss material, and of comparable quality to that in the core. The ease with which light can be coupled into the fiber for transmission is related to the numerical aperture, which for a step-index fiber is defined as

$$\sin \theta_0 = n_1 \sin \theta_c = \sqrt{n_1^2 - n_2^2}, \tag{7.40}$$

where θ_0 is the angle of acceptance of the fiber. This is illustrated in Fig. 7.38(a).

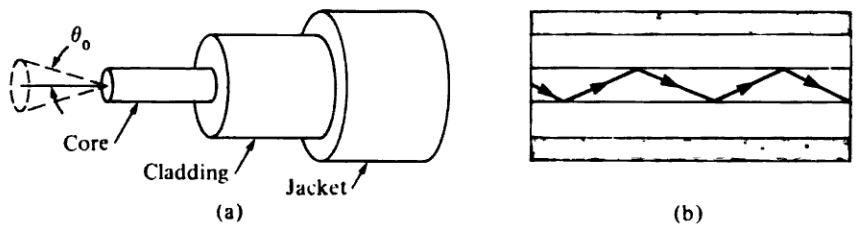


Figure 7.38 (a) Basic construction and (b) propagation in a step-index fiber optic cable (not to scale).

With this brief introduction, we see that light can be guided through an optical fiber by means of dielectric boundaries having lower refractive indexes than the core. This is illustrated in Fig. 7.38(b). Losses in a well-designed fiber are primarily a result of impurities, and continuing material research is yielding fibers with losses considerably less than 1 dB/km at some wavelengths.

While ray optics gives an intuitive picture of propagation in the optical fiber, the electromagnetic theory of guided waves shows that different modes of propagation are possible, dependent on the geometry and relative dimensions of the fiber. Multimodal optical fibers have dimensions large enough that several different modes are possible. Light is reflected off the core-cladding boundary to produce multiple paths of different lengths, and therefore slightly different time delays. This causes a dispersion or spreading out of the transmitted pulse, and this limits the maximum transmis-

† In a second type of optical fiber, called a graded-index fiber, the refractive index is made to vary as a function of the radial distance from the center of the fiber. This gradation in refractive index is effective in reducing the dispersion of light within the fiber. The central ray travels a shorter distance with a smaller velocity (because n is large), while the oblique ray travels a longer distance at a higher velocity (because n is small). In this way, differences in distance are compensated by opposite differences in velocity, and the two rays arrive at (approximately) the same time.

sion bit rate that can be achieved. In contrast, a single-mode fiber has a relatively small core diameter and allows only a single mode of propagation. Consequently there is less light dispersion, and higher transmission bit rates are possible.

The sources used for fiber optic communication systems are light-emitting diodes (LEDs), which produce noncoherent light, and semiconductor laser diodes, which are coherent (i.e., highly monochromatic). The LED is a popular choice because it is rugged and inexpensive, but it has the characteristic of emitting less power and cannot be switched off and on as fast as the laser diode. Laser diodes have better light coupling characteristics and are preferred for wide-bandwidth applications.

After an optical signal has been launched into the fiber, it is attenuated and distorted with increasing distance as a result of scattering, absorption, and dispersion effects. A significant factor in an operational fiber optic system is the loss due to interconnection of the fibers. The fiber-to-fiber coupling efficiency is proportional to the numerical aperture of the fibers and to their alignment.

The receiver in the fiber optic communication system consists of a photodetector, and associated circuitry for amplification and signal restoration. Photodetectors are square-law devices that convert optical power into electric current. Two types used in fiber optic communication systems are the *pin* photodiode and the avalanche photodiode (APD). Both device types are efficient and respond rapidly, but the avalanche photodiode offers greater sensitivity as a result of its internal gain mechanism (avalanche effect).

After an optical signal has traveled a certain distance in a fiber, the signal becomes attenuated and distorted to such an extent that a repeater is needed to amplify and reshape the signal. An optical repeater consists of a receiver and a transmitter. These repeaters are placed at intervals throughout the system, as dictated by the expected losses. The link power budget for an example of a fiber optic communication system operating at 417.792 Mbps is shown in Table 7.2. In this system, repeaters are required every 40 km (25 miles) to stay within the design margins.

TABLE 7.2 Link Power Budget for the LaserNet 1.3 μm Fiber Optic Communication System.[a]

Average effective transmitter power	−3.5 dBm
Receiver sensitivity	−36.0 dBm
Available margin	32.5 dB
Losses	
Fiber attenuation (40 km × 0.5 dB/km)	20.0 dB
Chromatic dispersion/reflection	1.5 dB
Splices, 15 @ 0.1 dB/splice	1.5 dB
Connector losses	3.0 dB
System design margin	5.0 dB
Excess margin	1.5 dB

[a] From: C. M. Siperko, "LaserNet—a Fiber Optic Interstate Network (Planning and Engineering Considerations)," *IEEE Communications Magazine*, vol. 23 (May 1985): 31–45.

Fiber optic systems are rapidly gaining acceptance for terrestial communication systems. It is expected that they will be popular choices for data communications in new systems designs.

7.8 USE OF PARITY AND REDUNDANCY IN PCM

In binary pulse-code modulation, each sample of the signal is represented by a codeword of, say, k bits. After these bits are transmitted, the receiver must recognize each codeword in order to reconstruct the samples. However, errors may occur in transmission as a result of noise. One way to increase the reliability of communication is to add extra bits, at the expense of added transmission time or bandwidth, in order to detect and, possibly, to correct the errors. We consider two simple and effective ways to accomplish greater reliability; more advanced methods are considered in Chapter 9.

One simple and effective error-detecting code can be formed by adding an extra binary digit at the end of each codeword. This extra bit is called a *parity-check* bit. For a binary code, the check bit is often chosen to make the number of 1's in each codeword an even number. (If the number of 1's is already even, a 0 is added to signify that fact.) Such a code is said to have *even* parity. Similarly, for *odd* parity the added parity bit is chosen to make the total number of 1's in the codeword (including the parity bit) an odd number. Let's assume we are using even parity: In a seven-bit codeword (i.e., $k = 7$), the message 0100110 becomes 01001101, and 0101101 becomes 01011010.

The extra bit added to each codeword is said to be redundant because it does not convey any information concerning the value of the sampled signal. It is used only to protect against possible errors by verifying the odd or even number of 1's. The code-rate efficiency of a system is defined as the ratio of the number of bits used to convey signal information divided by the total number of bits sent. Thus the code-rate efficiency decreases as more redundancy is added.

The single parity-check code has a code-rate efficiency of $k/(k + 1)$ and is widely used in digital computation and recording. It offers a simple and effective method for error detection when the possibility of making a bit error is low. When an error is detected, a retransmission can be requested. In other cases, those codewords with detected errors may simply be discarded. Note that an error in two bits in a codeword is not detectable using a single parity-check code.

We may wish to add enough redundancy to the code in order not only to detect single errors but also to correct them. As an example of an error-correcting code, suppose that in transmitting binary PCM we transmit a sequence of three 0's for every 0 and a sequence of three 1's for every 1. Because the two added digits are completely redundant, this repetition code has a code-rate efficiency of $\frac{1}{3}$.

If the possibility of making more than a one-bit error in a codeword is low, we can use a "majority decision" rule to decode the received bit stream. Thus if we receive 001 or 010 or 100, we decode the message as 000; and if we receive 011, 101, or

110, we decode the message as 111. However, if two bit errors occur in a codeword, we would make the wrong decision. Thus the single-bit error correction code is most useful when the possibility of making a bit error in a codeword is low. Repetition codes have good error-detecting capabilities at the expense of transmitting many redundant symbols. The intentional use of redundancy to advantage, whether for error detection or for error correction, is known as *error-control coding*.

To appreciate some of the underlying principles of error-control coding, we introduce the concept of codeword separation or distance. To do this, we define the *Hamming weight* $w(s_i)$ of a codeword s_i as the number of ones in that codeword. The *Hamming distance* $d_{ij} = d(s_i, s_j)$ between codewords s_i, s_j is defined as the number of positions in which s_i, s_j differ. An equation for the Hamming distance is

$$d_{ij} = w(s_i \oplus s_j) \tag{7.41}$$

where $\oplus$ denotes modulo–2 addition. For example, the Hamming distance between the two codewords in the repetition code above is 3, and the decoder rule could have been stated as "choose that codeword with minimum Hamming distance from the received codeword."

It is obvious that a code having all distinct codewords must have a minimum Hamming distance of at least 1. Not quite as obvious is the fact that a code which permits detection of up to e errors per word or correction of up to e errors per word must consist of codewords for which:

$$d_m = \begin{cases} e + 1 & \text{error detection} \\ 2e + 1 & \text{error correction} \end{cases} \tag{7.42}$$

where d_m is the minimum Hamming distance between codewords.

DRILL PROBLEM 7.8.1

A given code consists of the codewords: 0000000, 0011110, 0101101, 0111000, 1001100, 1011001, 1101010, 1110100. If 1011011 is received, what is the decoded codeword based on minimum Hamming distance?

ANSWER: 1011001.

DRILL PROBLEM 7.8.2

A single parity-check code used in computer listings of registration numbers, etc., is found by doubling alternate digits, beginning from the right, and adding the digits modulo–10. (Numbers greater than 10 are counted as two digits.) The parity-check digit is found by taking the tens complement (why?) of this sum. (a) Determine the parity-check digit for the numbers 837412 and 834712. (b) In addition to single-error detection, what type of error is being detected?

ANSWER: (a) 6, 2; (b) Interchange of two adjacent digits.

7.9 TIME-DIVISION MULTIPLEXING OF PCM SIGNALS

In the case of the time-division multiplexing of PAM signals, a number of analog signals are combined sequentially at a common sampling rate and the composite pulse stream is transmitted. This procedure can be extended to include PWM and PPM signals. Multiplexing of PCM signals presents some different problems that we will investigate in this section. Unless specified otherwise, we assume binary PCM.

Two basic types of transmission of digital data are known as *synchronous* and *asynchronous*. In synchronous transmission, the bit rate of each multiplexer input is fixed and is synchronized to a master clock. Thus within any given time period there is always a discrete specific number of bits, and the bits follow in a very regular predetermined progression. The multiplexer operates at a multiple of the master clock rate to sequentially format all data and then transmit the data block. Efficiency is high because one binary word follows the preceding one, with no special designators necessary to separate the words.

In asynchronous transmission, words are sent one at a time without necessarily having any fixed time relationship between one word and the next. Many data terminals—such as teletypewriters and other keyboard terminals—transmit alphanumeric symbols (i.e., letters of the alphabet, decimal numbers, and various control symbols) asynchronously, and we refer to the resulting data words as *characters*. In asynchronous transmission, the receiver has to reestablish synchronization for every character. This is accomplished by beginning each character with a start pulse. Thus the receiver clock is resynchronized at the beginning of each character, and small variations in the rate between transmit and receive clocks can be tolerated. A stop pulse is transmitted at the end of each character.

Although various codes could be used, the majority of data transmission systems use one of two international standardized codes, possibly with some minor deviations. These two codes were established by the Consultative Committee on International Telegraphy and Telephony (CCITT), a committee operating with the International Telecommunications Union and based in Geneva. The first of these is CCITT International Alphabet No. 2. This five-bit code, also known as the Baudot code, is widely used for Telex transmissions around the world. The second is the CCITT International Alphabet No. 5. The U.S. ASCII code (American Standard Code for Information Interchange) is a modification of this seven-bit code (Reference 1).

Most modern data terminals and associated devices in the United States have standardized a 10-bit word for each character, each word consisting of a start bit (always represented by 0), seven data bits using the ASCII code, a parity-check bit, and a stop bit (represented by a 1). An example of a typical asynchronous terminal character is shown in Fig. 7.39. Because start and stop bits are not required in synchronous transmission, these extra bits are the overhead required for the asynchronous transmission.

The purpose of the multiplexer (often abbreviated MUX) is to divide the information-handling capacity of a communications system between a number of terminals. In this respect, it is "transparent"—a MUX does not do anything to alter the data as it is handled.

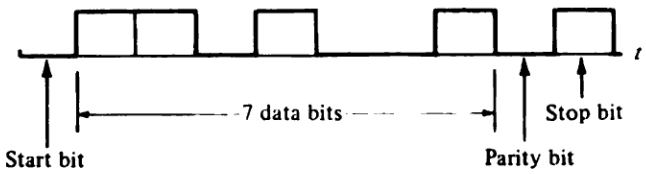


Figure 7.39 A typical 10-bit asynchronous terminal character using the ASCII code.

A *concentrator* is a computer-based device with some similarities to the multiplexer in that it combines the data from a number of terminals onto a high-speed line for transmission. However, it can alter the forms of the data streams prior to merging them. The concentrator, acting as an intelligent terminal, often is capable of assembling complete messages or blocks of messages before transmission, performing code, speed, and format conversions, and improving on line utilization by statistically averaging the traffic from the inputs onto a high-speed line. It can also compress data so that fewer bits are required to convey the desired information, and automatically detect and correct errors. Recently some intelligent multiplexers, generally microprocessor controlled, have been designed to assume some of the functions of concentrators (e.g., statistical multiplexing, error correction, etc.). For our purposes here, we consider only the nonintelligent multiplexer.

There are three major classes of nonintelligent multiplexers in use. The first is used for completely synchronous transmissions. The object of such multiplexers is to combine a number of synchronous inputs and feed out the combined bit stream sequentially at a higher (synchronous) bit rate. The bit rates of all inputs and the multiplexing rate must be governed by a master clock. The efficiency of such multiplexing systems is high, but provisions must be made to supply a master clock signal to all portions of the network being serviced. Long-range plans for telephone service in the United States include plans for a completely synchronous system with a master clock located at Hillsboro, Missouri.

A second class includes those multiplexers designed for almost-synchronous (quasi-synchronous) systems, in which the clock rates are nominally the same for all inputs, but in which small variations about the clock frequencies may exist. In general, multiplexers designed for this type of service operate at high bit rates and are integral parts of larger data transmission services. They multiplex on a bit-by-bit basis and, therefore, are called bit-interleaved multiplexers. To keep high transmission efficiencies without elaborate control procedures, these multiplexers use synchronous timing methods that run slightly fast and then, when necessary, insert bits (a process known as *bit stuffing*) to keep the input rates up to the desired rate.

A third major class of multiplexers includes those designed to combine a number of low-speed asynchronous transmissions into one higher-speed multiplexed signal. Generally these multiplexers are designed to handle the outputs of data terminals and to transmit the data over voice-grade lines in a commercial telephone system. When used with modems to interface with the lines, they normally transmit at output rates of 1200, 2400, 3600, 4800, 7200, or 9600 bits per second. Choice of output bit rate

depends on the complexity of the modem and whether the voice-grade line used is privately leased and specially conditioned, or dialed as in any ordinary telephone call.[†] Multiplexers in this class usually multiplex on a character-by-character basis; input rates may be as low as 75 bps per device.

Although we have grouped multiplexers into three broad categories, it is not always true that an asynchronous one is slow and a synchronous one is fast. There are terminals that operate over short distances at fairly high asynchronous rates. There are also some terminals in use that operate synchronously at 1200 bps.

Another multiplexing method of increasing importance in communications satellite networks is that of time-division multiple access (TDMA). In TDMA systems, each terminal samples and temporarily stores its user information and then transmits this information in short bursts at high bit rates. All terminals in the network transmit using the same carrier frequency (generally about 6 GHz for the up-link and 4 GHz for the down-link for the communications satellites). The bursts of each TDMA terminal are transmitted synchronously at assigned times relative to the transmissions of other terminals in the network. These bursts arrive at the satellite one at a time in a closely spaced sequence that never overlaps. The bursts are amplified by the satellite transponder, shifted to a different carrier frequency, and then transmitted back to the earth stations. A receiver accepts the burst, temporarily stores the bits and then reads out the bits at a slower rate and distributes the information to the users.

Implementation in TDMA terminals often uses a distributed-processing approach. The relatively slow-speed sampling and buffer-store operations are under the control of a conventional microprocessor (known as the background processor). A high-speed processor (known as the foreground processor) controls the high-speed bursts. Each processor has its assigned functions. Both share some functions to enable the background processor to load and read selected data into and from the foreground processor and to provide the proper synchronization and timing.

TDMA systems offer improved flexibility for the handling and multiplexing of digital data from sources with widely varying data rates. Their use is becoming more attractive as economical and reliable methods of digital storage and control become more readily available.

7.9.1 ★ Bit-interleaved Multiplexers and Bit Stuffing

In bit-interleaved multiplexers, input channels are sampled in succession on a bit-by-bit basis. Synchronization bits are added in a predetermined format to the resulting output data stream. After all inputs have been serviced at least once and the necessary synchronization bits have been added, the multiplexer is ready to begin its next cycle. Such a format is illustrated in Fig. 7.40 for a multiplexer with four inputs.

The shortest period of time in which all multiplexer inputs are serviced at least once is called a *frame*. Location information of the channels is called framing infor-

[†] Some typical characteristics of both basic and conditioned telephone lines are given in Appendix E. The principles of modems are discussed in Chapter 10.

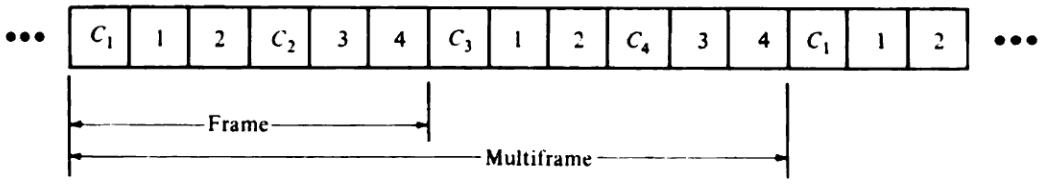


Figure 7.40 A four-channel framing format.

mation. The bits added for framing and synchronization are collectively known as control bits. The shortest period of time in which all inputs to be multiplexed and in which all control bits have been added to the frame is called a *multiframe* or *superframe*. In the simplified framing format of Fig. 7.40, there are four input channels plus two control bits per frame and two frames per multiframe. The control bits are merely overhead and decrease the information-handling capacity of the multiplexer. In this particular case, the throughput efficiency of the multiplexer is $\frac{4}{6} = 67\%$.

All inputs to the multiplexer are assumed to have the same nominal bit rate, R_i. Unless all inputs are completely synchronous, however, small fluctuations can be expected. If no provision is made for accommodating these fluctuations, some information could be lost in the multiplexing process. To prevent this, the multiplexer is operated at a rate slightly higher than the sum of the maximum expected rates of the inputs. Because the multiplexing rate will gradually gain, provision is made for the occasional adding of an arbitrary bit (0 or 1) to the bit stream in order to maintain the multiplexer rate. This operation is known as *bit stuffing* and is illustrated in Fig. 7.41. The stuffed bits carry no information and are added merely to maintain the higher bit rate of the multiplexer.

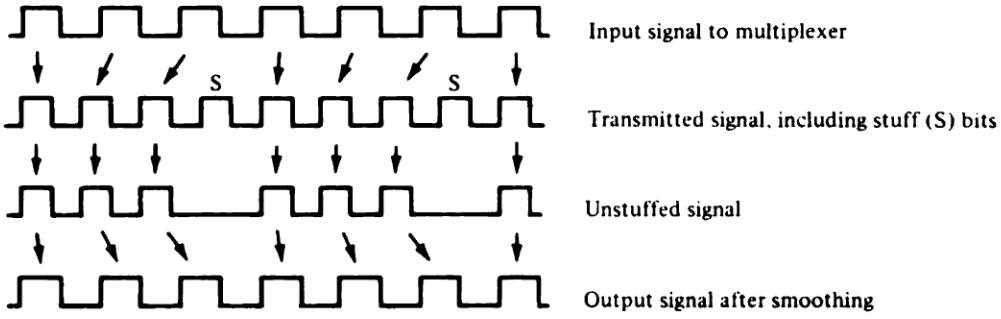


Figure 7.41 Inputs and outputs of a bit-stuffed multiplexer/demultiplexer.†

To smooth out short-term fluctuations in the input bit rate and to signal the need for a stuff bit, each input is applied to a buffer memory, or *elastic store*, as shown in Fig. 7.42. This buffer acts as a reservoir whose fullness is monitored by the buffer control. Whenever the fullness of the buffer memory drops below a given threshold, a stuff bit is inserted into the data stream on command of the synchronizer.

† From J. J. Spilker, Jr., *Digital Communications by Satellite*, © 1977, p. 118. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

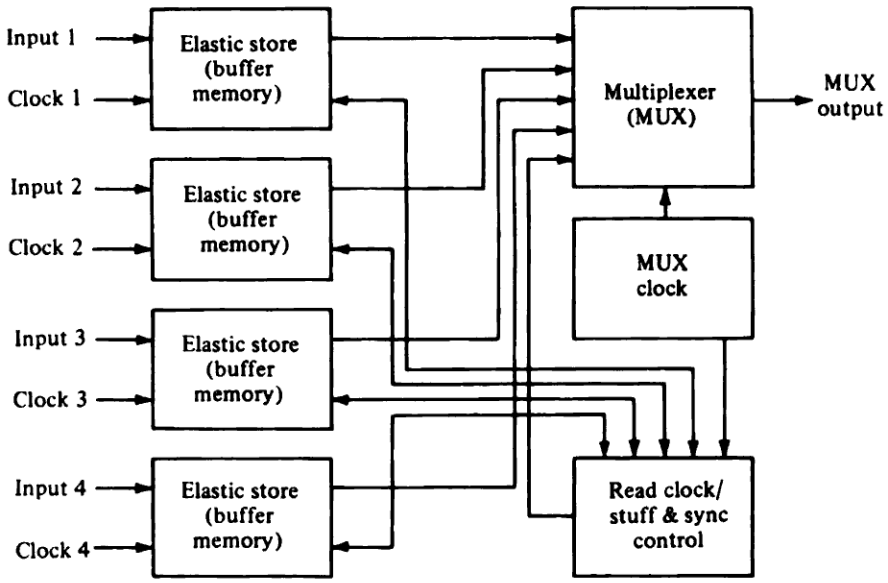


Figure 7.42 A bit-interleaved multiplexer using bit stuffing.

An important function of the demultiplexer at the other end of the line is to remove the stuff bits from the data stream. This requires identification of the stuffed bits in as reliable a manner as possible without using excessive format overhead. Various techniques are used.

Consider the Bell System M12 multiplexer as an example of framing and synchronization methods for bit stuffing.[†] This multiplexer is designed to combine four T1 carrier inputs into one T2 carrier output (cf. Fig. 7.32 and Table 7.1). The M12 multiplexer frame format is shown in Fig. 7.43. This multiplexer uses bit-by-bit interleaving of the four T1 carrier inputs until a total of 48 bits, 12 from each input, have been taken. Then a control bit is inserted and the sequence is repeated. Three types of control bits are used—the M-, F-, and C-series. These provide multiframe indication, frame synchronization, and identification of which of the four inputs has been stuffed. The first row of the format is transmitted first, then the second, and so on.

Order of transmission ⟶

M_0	(48)	C_I	(48)	F_0	(48)	C_I	(48)	C_I	(48)	F_I	(48)
M_I	(48)	C_{II}	(48)	F_0	(48)	C_{II}	(48)	C_{II}	(48)	F_I	(48)
M_I	(48)	C_{III}	(48)	F_0	(48)	C_{III}	(48)	C_{III}	(48)	F_I	(48)
M_I	(48)	C_{IV}	(48)	F_0	(48)	C_{IV}	(48)	C_{IV}	(48)	F_I	(48)

Figure 7.43 Frame format of the Bell System M12 multiplexer.

[†] Bell Telephone Laboratories, *Transmission Systems for Communications*, Fourth ed., 1970, p. 612.

Each row represents a frame, and all four rows represent a multiframe. Only one stuff bit per input channel is allowed in each multiframe.

The subscripts on the M- and F-series control bits in Fig. 7.43 denote the bit (0 or 1) transmitted. The F-series uses an alternating 0, 1 sequence that provides a fundamental frequency whose zero crossings give frame sync. Because the pattern is repetitive, a PLL in the demultiplexer can lock on to it. After doing so, the demultiplexer searches for the 0111 pattern of the M-series. Having found the M-series pattern, the demultiplexer has also identified the multiframe.

Bit stuffing information is supplied by the C-series, and subscripts are used in Fig. 7.43 to designate input channels; thus C_{I} stands for input channel 1, C_{II} for input channel 2, etc. The bit patterns in the M- and F-series are periodic, and detection is tolerant of transmission errors. In contrast, the C-series bit patterns are not repetitive, and they must be transmitted with sufficient redundancy to maintain bit integrity—that is, without leaving in a stuff bit or removing a valid data bit, thus shifting the entire succeeding data stream one bit ahead or behind in time. This loss of bit integrity would cause the remaining bits in the frame to be read out erroneously—probably to the incorrect demultiplexer output port. Therefore this is a much more serious error than a data bit error.

To protect bit integrity, a $\times 3$ repetition code is employed in the M12 multiplexer. The insertion of a stuff bit in any one frame is denoted by setting all three C's in that frame to 1; similarly, three 0's denote no stuffing. Majority-logic decoding is used so that if two out of the three are detected as 1's, a decision is made that a stuff bit has been inserted. Conversely, if two out of the three are detected as 0's, a decision is made that no stuff bit was inserted.

Each multiframe contains $12 \times 6 \times 4 = 288$ bits for each of the four input channels. Because only one stuff bit per input channel is allowed in each multiframe, the maximum fractional change in input bit rate that can be accommodated is 1/288, or 0.35%. Good clocks can be maintained to within 60 ppm (parts per million) or 0.006%, well within the capabilities of this multiplexer. Note that this maximum fractional change represents the change from no stuff bits to all stuff bits. All practical situations lie between these extremes. Let the average number of stuff bits per channel per multiframe be designated as s; then $0 \leq s \leq 1$. To allow for plus or minus deviations from a nominal input clock rate R_{i}, we should choose s near $\frac{1}{2}$ so that bit stuffing will occur, on the average, 50% of the time.[†] For example, for the M12 multiplexer we have four input channels each at nominal rate R_{i}. Allowing for an overhead of one control bit in every 48 data bits and noting that there are 288 data bits in each multiframe, we find the multiplexer output rate R_{o} is

$$R_{\mathrm{o}} = 4R_{\mathrm{i}}\left(\frac{49}{48}\right)\left(1 + \frac{s}{288}\right).$$

Using $R_{\mathrm{o}} = 6.312$ Mbps and $R_{\mathrm{i}} = 1.544$ Mbps (see Table 7.1), we have $s = 0.335$.

[†] It might appear that one would always design for $s = \frac{1}{2}$, which gives the greatest frequency tolerance each way from the nominal rate. However, the rate at which the buffer memories require stuffing may vary, causing jitter in the demultiplexed bit stream. This jitter has a peak-to-peak value equal to s, and therefore, it is desirable to keep s small.

DRILL PROBLEM 7.9.1

Compute the average number of bits stuffed per channel in the Bell System M12 multiplexer if the accuracy on both the input data stream and the multiplexer clock is ±90 ppm and (a) the input rate (R_i) is slow, but the output rate (R_o) is fast; (b) the input rate is fast, but the output rate is slow.

ANSWER: (a) 0.387; (b) 0.283.

EXAMPLE 7.9.1

A given bit-interleaved multiplexer operates at an output rate of 224 Mbps and 145 bits per frame — 143 data bits plus two control bits, labeled the F-series and the S-series.[†] The frame format is shown in Fig. 7.44(a). The frame sync bit (F-series) simply alternates 0, 1. The synchronization bit (S-series) occupies the time slot normally used for channel 73, and the S-bit format is shown in Fig. 7.44(b). The marker word M is the 16-bit word 0110101001011010. Each channel stuff word C uses three-to-one redundancy: 000 not stuffed, 111 stuffed. Nominal input bit rate per channel is the T1 carrier rate of 1.544 Mbps. For this multiplexer, calculate (a) the throughput efficiency; (b) the number of frames in a multiframe; (c) the maximum fractional change in input rate that can be accommodated; (d) the average number of stuff bits per channel when the input is at the nominal T1 carrier rate.

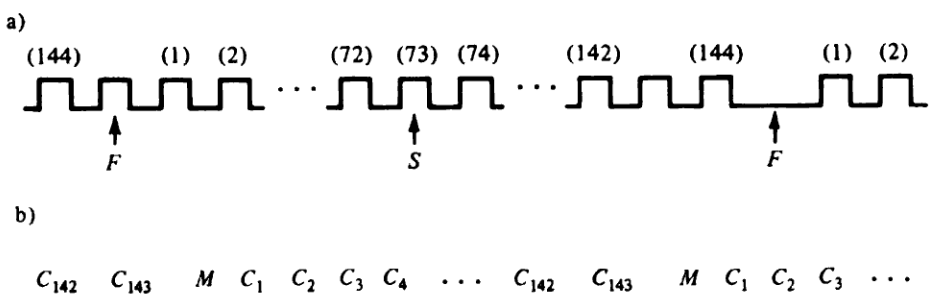


Figure 7.44 (a) Frame format and (b) S-bit format for the multiplexer in Example 7.9.1.

Solution a) Two bits out of 145 are used for overhead so that the throughput efficiency is $(143/145) \times 100 = 98.6\%$.

b) Because two channels (F and S) do not require stuff bits, transmission of one complete S-bit sequence requires $M = 3(143) + 16 = 445$ frames.

[†] This is a simplified version of the multiplexer described in F. J. Witt, "An Experimental 223 Mb/s Digital Multiplexer-Demultiplexer Using Pulse Stuffing Synchronization," *Bell System Technical Journal* (November 1965): 1843–1885; see also J. J. Spilker, *Digital Communications by Satellite*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1977, p. 122.

c) The nominal multiplexed rate of 145 channels of T1 carrier signals is 145(1.544 Mbps) = 223.88 Mbps. Thus the 224 Mbps rate permits a 0.05% fractional change.

d) For this system, $R_o = 145R_i(1 + \delta)$, where δ is the maximum fractional increase in R_i. Because the S-bit sequence contains 445 bits, the maximum fractional change allowed is one bit in 445. But this would require a stuff bit in every multiframe. Letting s be the average number of stuff bits per channel per multiframe, we have

$$R_o = 145R_i\left(1 + \frac{s}{445}\right).$$

Using $R_o = 224$ Mbps and $R_i = 1.544$ Mbps, we find that $s = 0.239$.

7.9.2 ★ Character-Interleaved Multiplexers and Word Stuffing

It is generally simpler to perform the time-division multiplexing of a number of asynchronous terminals on a character-by-character basis, rather than bit by bit. This is particularly true if flexibility to accommodate terminals with different character sizes is desired. Therefore, character-interleaved multiplexers have been adopted widely for the multiplexing of slow-speed asynchronous devices such as data terminals. Each character, usually with a parity check and possibly with an address code, is sent intact as a word and these are interleaved in the bit stream. Synchronous data streams, generally transmitted at higher speeds, are often multiplexed together with the asynchronous data.

The multiplexer rate for each input port is set to run slightly faster (say, 3%) than the maximum character rate expected from the asynchronous terminals. Because individual characters are sent out at a higher rate than received from the terminal, the multiplexer must be able to buffer at least one character from each terminal connected to it. Generally the start and stop bits are dropped from each character before buffering, and then reinserted by the demultiplexer at the other end of the line.

Each of the buffer memories is monitored; and when there is no character to transmit from a given input port, a predetermined control character is transmitted. This procedure is called *word stuffing*. Word stuffing is procedurally like bit stuffing, except that the data are handled on a character-by-character basis. Word stuffing is more tolerant of bit-rate changes than bit stuffing, and can accommodate bit-rate variations all the way from zero input rate to the maximum bit rate of the multiplexer. Throughput efficiency, however, is generally lower than for multiplexers using bit stuffing.

In word stuffing, a predetermined word is assigned to indicate a word stuff. This word must be at least a Hamming distance of 1 (preferably more) from every other word in the alphabet used. Usually this distance is accomplished by adding a parity

bit in the multiplexer. Various techniques are possible; one relatively simple description follows, and a second is included as Problem 7.9.3.

To begin our example of a word stuffing procedure, we assume that an asynchronous terminal transmits 10-bit characters to a multiplexer input port. At the multiplexer, the start and stop bits are dropped, and parity-check bits for even parity are added. Thus each 10-bit character is transmitted as a 9-bit word. Assuming that one character in ten is needed for synchronization, the throughput efficiency of the multiplexer is then $(\frac{8}{9})(\frac{9}{10}) = 80\%$. The stuff word is assigned to one of the permissible characters in the alphabet; but to distinguish it from a data word, we use odd parity for the stuff word. This is sufficient to maintain a minimum Hamming distance of 1 between the stuff word and other words in the alphabet. The demultiplexer searches for this particular word with odd parity to detect the stuff word. One error, of course, is sufficient to cause an error in the output character.

Better protection against stuff word errors requires a greater distance between the stuff word and other words in the alphabet. Therefore a more sophisticated scheme is to note that in adding a parity-check bit to an m-bit word we have generated $2^{m+1} - 2^m$ new words, only one of which is the stuff word. The goal is to choose new data words out of this larger set in such a way that we maximize the distance from the stuff word.

Up to this point, we have assumed that the data format constitutes a frame when the multiplexer—either bit-interleaved or character-interleaved—has serviced all input ports at least once. This is what a conventional multiplexer does. Thus in a conventional multiplexer each input port has its own frame slot that regularly repeats, making available a constant and guaranteed transmission capacity. At any given time, however, some input ports may not have data to send, and the conventional multiplexer is inefficient in reserving time slots when there is no data to send. To remedy this, a version of time-division multiplexing known as *demand multiplexing* allocates available time according to input demand, keeping all slots filled with data as long as there is data waiting to be sent, and interchanging slot allocations, if necessary, to accomplish this.

In demand multiplexing, each input is applied to a buffer memory whose capacity is large enough to store at least several characters. The buffer memory control then monitors the fullness of the buffer memories and instructs the multiplexer to sample from those having the highest occupancy.[†] Every character or block of data to be transmitted is defined by an address attached to it by the multiplexer. Demand multiplexing can increase the efficiency of a transmission system at the expense of higher bit overhead, some variable delays introduced in the data stream, and the risk of

[†] One version of demand multiplexing using long lines (e.g., submarine cables) or satellites is called Time Assignment Speech Interpolation (TASI). Telephone systems have found that subscribers speak only about 40% of the time, on the average, in normal conversations. Therefore some savings can be realized by detecting a talker's speech and then assigning it to an unused channel. That channel remains assigned until the talker is silent and the channel is needed by another speaker. During low-traffic periods, the talker may keep the channel throughout a conversation. During high-traffic periods, however, successive words may go over different routes.

congestion if all terminals serviced should require transmission at the same time. If the latter situation should arise, a set of rules or procedures, with some hierarchy or priorities, must be assigned to handle the situation in an organized fashion.

A trend in modern time-division multiplexers is toward such devices as demand multiplexers, intelligent multiplexers, and concentrators. In fact, such devices can be viewed as intelligent terminals in an overall communication system that may not only multiplex and demultiplex data but also be users of data. As complexity increases, with added control features under the supervision of minicomputers or microprocessors, there are increasing requirements for development of a set of detailed operating procedures. Such a set of rules and procedures, usually involving a hierarchical set of instructions and priorities, is called a *protocol*.

Although the details of a protocol will differ from one system to another and will depend somewhat on the particular minicomputer or microprocessor used, there are some basic features that usually can be identified. These are designed to solve such operating problems as[†]

1. Framing. Determination of which bit groups constitute characters and which groups of characters constitute messages.
2. Error control. Detection of errors, acceptance of correct messages and requests for retransmission of faulty messages.
3. Sequence control. Identification of messages that are retransmitted by the error control system in order to avoid losing messages or duplicating them.
4. Line control. Determination, in the case of a half-duplex line or a multipoint line, of which station is going to transmit and which station(s) is/are going to receive.
5. Time-out control. Solution of the problem of what to do if message flow suddenly ceases entirely.
6. Start-up control. Solution of the problem of what procedure should be followed to get transmissions started in a communication system that has been idle.

Several of these topics have been introduced, with specific reference to multiplexers, in the preceding discussion. It is beyond the purpose of this textbook, however, to pursue the subject further.[‡]

DRILL PROBLEM 7.9.2

A proposed character-interleaved multiplexer uses a standard 10-bit input word consisting of a start bit (0), 7 data bits using the ASCII code, a parity-check bit (even parity) and a stop bit (1). At the multiplexer a parity-check bit derived from the 8-bit data word is added using even parity, resulting in an 11-bit word. The stuff word uses

[†] J. E. McNamara, *Technical Aspects of Data Communication*. Maynard, Mass.: Digital Equipment Corporation, 1977, p. 191.

[‡] See, e.g., D. W. Davies, D. L. A. Barber, W. L. Price, and C. M. Solomonides, *Computer Networks and their Protocols*. New York: John Wiley & Sons, 1977.

one of the permissible characters in the input alphabet, inverts the start and stop bits, and uses odd parity for the multiplexer parity-check bit. (a) What is the minimum Hamming distance between the stuff word and other words in the alphabet? (b) Assuming that one character in ten is needed for synchronization, determine the throughput efficiency of the multiplexer. (c) Devise a stuff-bit decision rule.

ANSWER: (a) 3; (b) 65%; (c) Two-of-three decision rule of (inverted) start and stop bits and the (odd) parity-check bit.

7.10 INTEGRATED SERVICES DIGITAL NETWORK (ISDN)[†]

As a result of the rapidly increasing application of computers and related digital equipment, and the need for communication between such devices, the concepts of networks originally intended primarily for telephone service are being reformulated to provide more efficient transmission of digital data. These communication networks not only must be capable of handling and routing digital data in an efficient manner but also must include provisions for digital telephone service. This section describes briefly some recent advances in this field.

Three basic modes for the operation of communication networks are circuit (line) switching, message switching, and packet switching. In circuit switching, communication is achieved by a connection that is dedicated to the communicating pair for the entire duration of message transmission. During this time the line is completely dedicated and no other users have access to the transmission path, including those possible intervals when no data is being transmitted. Circuit switching is most efficient when there is a relatively constant information transmission rate; it becomes inefficient when transmission occurs in short bursts separated by long intervals with no transmission. Circuit switching can be used for both analog and digital systems, and therefore it has been the one developed first for telephone systems. The other two types listed below lend themselves more specifically to the transmission of digital data.

In the second type, called message switching, an entire message is accumulated in a buffer storage and then sent over the network when a path becomes available. Transmission at that time may be only to the next switching office in the network, where the message is again stored in a buffer until another path is available that routes the message toward its destination, and so on. Message switching provides a more efficient transmission of information, compared with circuit switching, when data occur in short bursts with intervals in which no data are sent. A path is needed only during transmission of data, and there is no need to allocate the entire path at a

[†] R. T. Roca, "ISDN Architecture," *AT&T Technical Journal*, vol. 65 (January-February 1986): 5–17.

given time. However, each message must carry identification concerning such items as source, destination, user (for toll service), and start and stop of information. This auxiliary data is overhead information to the system. Thus message switching has potential to improve the efficiency of a communication network for transmission of computer data at the expense of some system overhead.

As message length increases, the buffer storage capacity required for message switching at each switching office also increases. Because each switching office must wait to relay a message until the entire message is received, the net delay in transmission can become appreciable for long messages. Also, short messages tend to get delayed by the long messages in such a system. A third type of network, called a packet-switching network, is designed to avoid these disadvantages by subdividing messages into packets of shorter length. Each packet is handled as a separate entity by the network, in a manner similar to that used for message switching. However, each packet is now of modest size and can be handled efficiently by the buffer-store-and-forward operations of the network. In some packet-switched systems, different packets of a given message may arrive at a destination by different routes, with different delays, and possibly even in a scrambled order. It follows that in packet-switching networks each packet must be accompanied by overhead bits that are not only similar to those for message switching but also identified in order to correctly piece the message together. Thus, while packet switching provides increased flexibility and can provide an efficient transmission rate for burst-type information, it does so at the expense of increased overhead.

In older telephone systems, the transmission of digital information was accomplished by converting digital data to analog signals through use of modems and then handling the analog signals as if it were voice. Recent work has focused on a communications network that will not only provide for telephone service but which will also provide for a wide variety of digital data services. The Integrated Services Digital Network (ISDN) is such an end-to-end communications network that supports a wide range of services that can be accessed by standardized multipurpose user-network interfaces. ISDN is evolving from a network that is predominantly oriented toward telephone usage, and it is attempting to rectify shortcomings in that system for the transmission of digital data. Today's demands for services are characterized by access for services not only for voice but also for connection of users for digital data transmissions with banks, airlines, stock market companies, and other businesses. The ISDN concept uses a single user-network interface and is designed to integrate voice, data, and image in this interface. Consequently, a single interface can provide integrated access to all services. To accomplish this, the access interface is digital, and separates the information transfer functions (speech, data, image) from the signaling and management functions (ringing, toll recording, line maintenance). A set of general standards for the ISDN framework and services have been established by the International Telegraph and Telephone Consultative Committee (CCITT) to provide some compatibility between systems.

ISDN supports two standard interfaces to connect customer premises equipment to the network: the primary-rate interface (PRI) and the basic-rate interface (BRI). The ISDN architecture supports various information-bearing channel types on these two

interfaces. The basic information-bearing channel, called a B channel, carries voice or data communications and operates at 64 kbps for both interfaces. The D channel carries signaling information about transactions or calls on the B channels that it supports. It may also carry customer data packets. The D channel transmission rate is 16 kbps for the BRI and 64 kbps for the PRI. Traditionally, signaling information in telephone networks was carried "in band," in the same channel as the voice or data communications it controlled. However, this in-band signaling used up to 8 kbps of the 64 kbps channel, limiting the customer to an effective transmission rate of 56 kbps. The new signaling channel uses message-oriented signaling that is not only very efficient but also frees up the full 64 kbps B channel. Thus the B channel can give clear access at the full 64 kbps transmission rate. Additional channels, labeled H0 and H11, support clear access at the higher transmission rates of 384 and 1536 kbps, respectively.

The North American PRI channel structure is composed of 24 time slots. The most common arrangement is the arrangement 23B + D, in which the B and D channels are provided at 64 kbps. The PRI may also contain any combination of B and H0 channels, as well as a single H11 channel. The maximum information transfer rate in the PRI is 1536 kbps, with a sampling clock rate of 1544 kbps. The D channel need not be in the same interface as the B channels it controls. One D channel has sufficient capacity to control the B channels of 20 or more interfaces, depending on customer traffic characteristics. The PRI is expected to be used for major traffic sources, such as those for telephones—including digital private branch exchanges (PBX's)—and other intensive digital information transfer applications, such as local area networks (LAN's) and computer hosts.

The BRI structure uses the arrangement 2B + D, where the B channels are provided at 64 kbps and the D channel is provided at 16 kbps. BRI is expected to be used principally between network or premises switches and terminals. Thus, a D channel could carry control information between a customer and the network, and between equipment at two of the customer's locations. Both circuit-mode and packet-mode services for the B and D channels are specified in the CCITT standards.

Efforts are being made to maintain upward-compatibility in the changes that are being made in the telephone networks. Thus the new system must be compatible with those customers who do not yet have the ISDN services. However, the functionality of the non-ISDN terminating equipment will limit the overall ISDN capabilities in those instances.

7.11 THE MATCHED FILTER

In both PAM and PCM systems we are interested in maximizing the peak pulse signal in the presence of additive noise. We are particularly interested here in the case in which the signal plus additive noise is passed through a linear time-invariant filter. Out of all filters in this class, which one will yield a maximum output?

Let the signal input to the filter be $[f(t) + n(t)]$, where $f(t)$ is the signal and $n(t)$ is the additive noise. The output of the filter is $[f_o(t) + n_o(t)]$ and we wish to

maximize the ratio $|f_o(t_m)|/\sqrt{\overline{n_o^2(t)}}$, where $t = t_m$ is the best observation time (to be set). Actually, as we shall see, it turns out to be more convenient to maximize the square of this ratio.

Let the Fourier transform of $f(t)$ be $F(\omega)$ and let $H(\omega)$ be the frequency transfer function of the desired optimum filter. Then we can write

$$f_o(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)H(\omega)e^{j\omega t}\,d\omega\,,$$

$$f_o(t_m) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)H(\omega)e^{j\omega t_m}\,d\omega\,. \tag{7.43}$$

The power spectral density of the noise is $S_n(\omega)$, so that†

$$\overline{n_o^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_n(\omega)\,|H(\omega)|^2\,d\omega\,. \tag{7.44}$$

Dividing the squared magnitude of Eq. (7.43) by (7.44), we get

$$\frac{|f_o(t_m)|^2}{\overline{n_o^2(t)}} = \frac{\left|\int_{-\infty}^{\infty} H(\omega)F(\omega)e^{j\omega t_m}\,d\omega\right|^2}{2\pi\int_{-\infty}^{\infty} S_n(\omega)\,|H(\omega)|^2\,d\omega}\,. \tag{7.45}$$

At this point we make use of the Schwarz inequality (cf. Example 2.5.2):

$$\left|\int_{-\infty}^{\infty} f_1(x)f_2(x)\,dx\right|^2 \le \int_{-\infty}^{\infty} |f_1(x)|^2\,dx\int_{-\infty}^{\infty} |f_2(x)|^2\,dx\,. \tag{7.46}$$

The equality in Eq. (7.46) holds if, and only if,

$$f_1(x) = kf_2^*(x)\,, \tag{7.47}$$

where k is an arbitrary constant.

Now we let the two functions in Eq. (7.46) be identified with $H(\omega)\sqrt{S_n(\omega)}$ and $F(\omega)e^{j\omega t_m}/\sqrt{S_n(\omega)}$, respectively, so that Eq. (7.46) becomes

$$\left|\int_{-\infty}^{\infty} H(\omega)F(\omega)\,d\omega\right|^2 \le \int_{-\infty}^{\infty} |H(\omega)|^2\,d\omega\int_{-\infty}^{\infty} |F(\omega)|^2\,d\omega\,. \tag{7.48}$$

Substitution of this result into Eq. (7.45) gives

$$\frac{|f_o(t_m)|^2}{\overline{n_o^2(t)}} \le \frac{\int_{-\infty}^{\infty} |H(\omega)|^2\,d\omega\int_{-\infty}^{\infty} |F(\omega)|^2\,d\omega}{2\pi\int_{-\infty}^{\infty} S_n(\omega)\,|H(\omega)|^2\,d\omega}\,. \tag{7.49}$$

† Note that the mean-square value of $n_o(t)$ is independent of t.

For the special case in which the noise is white, $S_n(\omega) = \eta/2$, and we have

$$\frac{|f_o(t_m)|^2}{\overline{n_o^2(t)}} \leq \frac{1}{\pi\eta}\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega = \frac{E}{\eta/2}, \tag{7.50}$$

where E is the energy in $f(t)$ for a 1-ohm load. The equality in Eq. (7.50) holds only if [cf. Eq. (7.47)]

$$H_m(\omega) = kF^*(\omega)e^{-j\omega t_m}, \tag{7.51}$$

or

$$\begin{aligned} h_m(t) &= \mathcal{F}^{-1}\{kF^*(\omega)e^{-j\omega t_m}\} \\ &= kf^*(t_m - t). \end{aligned} \tag{7.52}$$

The constant k is arbitrary and we assume $k = 1$ for convenience.

We conclude from this result that the impulse response of the optimum system is the mirror image of the desired input signal $f(t)$, delayed by an interval t_m. Hence the filter is matched to a particular signal, as conveyed by the terminology *matched filter*.

The result expressed in Eq. (7.51) makes good sense intuitively when applied to the magnitude characteristic of a filter so that: $|H(\omega)| = |F(\omega)|$. (Remember that this result was obtained for the case of additive white noise.) This result states that one should filter in such a way as to attenuate strongly those frequency components in frequency intervals having little relative signal energy while attenuating very little those components where the relative signal energy is high. Recall also that we are filtering for signal recognition in the presence of noise, not for signal fidelity (in which one usually desires a "flat" frequency response).

The phase response is also very important and Eq. (7.51) states that the phase shifts in $f(t)$ should be negated in such a way that all frequency components in $f(t)$ add *in phase* at exactly the time $t = t_m$. In contrast, the noise spectral components add with random phases so that the peak-signal-to-rms-noise ratio is maximized.

The signal representation $f(t)$ is assumed to have a finite duration $(0, T)$. The impulse response of the matched filter $f(t_m - t)$ can be obtained by folding $f(t)$ about the vertical axis and shifting it to the right by t_m seconds. Restricting consideration to the physically realizable case with minimum delay, we choose $t_m = T$. This is illustrated in Fig. 7.45.

At the point $t = t_m$, the signal output of the matched filter is given by substituting Eq. (7.51) in Eq. (7.43) with $k = 1$:

$$f_o(t_m) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega = E. \tag{7.53}$$

Thus the output of the matched filter at $t = t_m$ is independent of the particular waveform chosen and depends only on its energy! The mean-square noise output of the matched filter is [Eq. (7.53) in Eq. (7.50)]

$$\overline{n_o^2(t)} = E\frac{\eta}{2}. \tag{7.54}$$

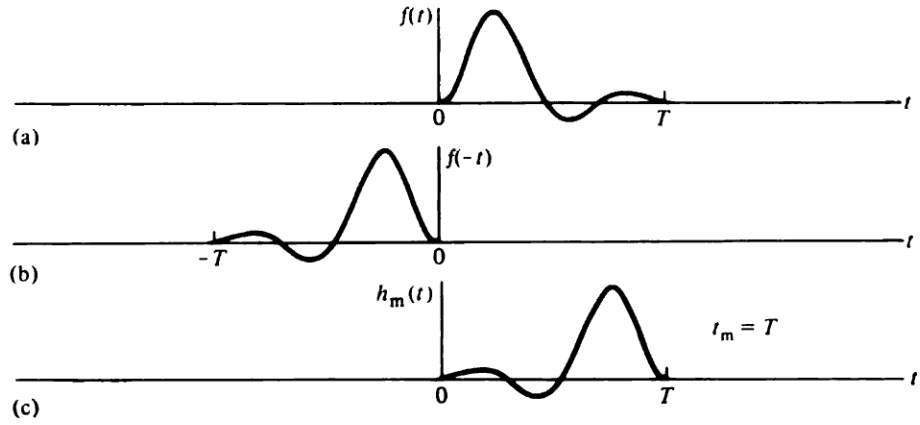


Figure 7.45 The matched filter for a physically realizable system, minimum delay.

The matched filter must be tailored to the signal waveform to achieve the maximum signal-to-noise ratio. One way in which this can be accomplished for the general case is by noting that the impulse response of a linear time-invariant filter can be approximated by a delay line that is tapped at various points and weighted by a set of fixed gains (cf. Section 3.10). Such a system is shown in Fig. 7.46, where the delay line has taps at the delays $k\Delta\tau$. For an N-tap delay line, the response $g(t)$ can be written as

$$g(t) = \sum_{k=0}^{N} f(t - k\Delta\tau)h(k\Delta\tau)\Delta\tau, \tag{7.55}$$

where the output of each tap is multiplied by the preset weight $h(k\Delta\tau)\Delta\tau$.

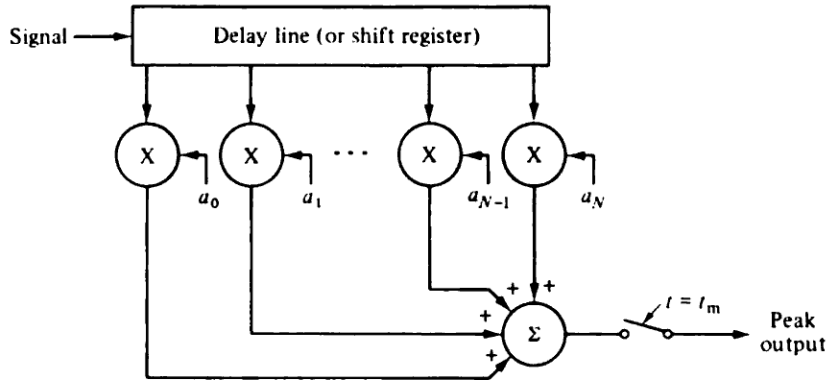


Figure 7.46 Delay-line realization of the matched filter.

For the specific case of a matched filter (real-valued case), we also have [cf. Eq. (7.52)]

$$h(t) = f(t_m - t),$$

so that the tap gains in Fig. 7.46 are given by

$$a_k = f(t_m - k\Delta\tau)\,\Delta\tau\,. \tag{7.56}$$

Thus once the signal $f(t)$ and the time t_m are known, the tap gains can be set to approximate the desired matched-filter characteristic.

An alternative realization of the matched filter is as follows. Let the input to the matched filter be $y(t) = f(t) + n(t)$ and the corresponding output be

$$\begin{aligned} g(t) &= y(t) \circledast f^*(t_m - t) = f^*(t_m - t) \circledast y(t) \\ &= \int_{-\infty}^{\infty} f^*(t_m - \xi)y(t - \xi)\,d\xi\,. \end{aligned} \tag{7.57}$$

We let $\zeta = t_m - \xi$, so that Eq. (7.57) becomes

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} f^*(\zeta)y(\zeta + t - t_m)\,d\zeta \\ &= r_{fy}(t - t_m)\,, \end{aligned} \tag{7.58}$$

where

$$r_{fy}(\tau) = \int_{-\infty}^{\infty} f^*(\zeta)y(\zeta + \tau)\,d\zeta$$

is the time cross-correlation function for energy signals (cf. Chapter 4). Noting that $y(t) = f(t) + n(t)$, we find

$$g(t) = r_f(t - t_m) + r_{fn}(t - t_m)\,, \tag{7.59}$$

where

$$r_f(\tau) = \int_{-\infty}^{\infty} f^*(\zeta)f(\zeta + \tau)\,d\zeta$$

is the time-autocorrelation function for energy signals. If the cross-correlation between $f(t)$ and $n(t)$ is zero, then the peak output $g(t_m)$ is given by $r_f(0)$.

A block diagram of the time cross-correlator is shown in Fig. 7.47. The incoming waveform $y(t)$ is multiplied by $f(t)$. This requires that the signal $f(t)$ be known *a priori* and either stored in memory or supplied from another source. The output of the multiplier is integrated to form $r_{fy}(\tau)$ in Eq. (7.58). To complete the operation, the switch is closed at $t = t_m$ to give the output $r_{fy}(0)$. Note that matched-filter detection is essentially a synchronous (i.e., coherent) detection.

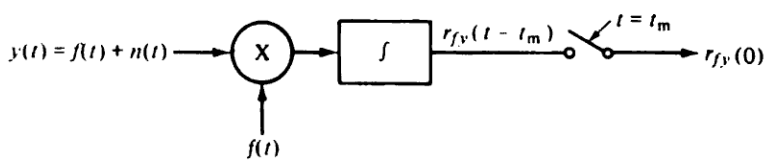


Figure 7.47 The time correlator.

In some cases the realization of a matched filter may be much easier to accomplish than the above methods would indicate. For example, consider the matched filter for the rectangular pulse waveform shown in Fig. 7.48(a). The impulse response of the matched filter is given by $h(t) = f(T - t) = f(t)$ so that Fig. 7.48(a) also describes $h(t)$. The convolution of these rectangular waveforms yields the triangular waveform shown in Fig. 7.48(b).

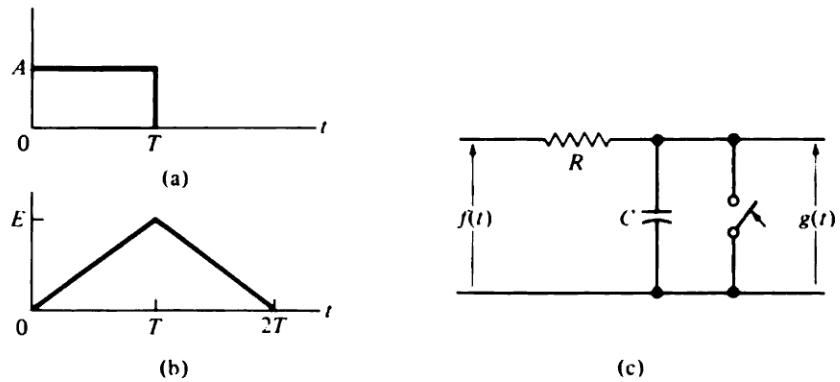


Figure 7.48 Matched filtering of a rectangular pulse waveform: (a) input; (b) output; (c) integrate-and-dump circuit approximation.

At the time $t = T$, the output of the matched filter is maximum. A sample is taken for the output and the trailing portion of the triangular waveform for $t > T$ is not used. The linearly increasing portion of the triangular output waveform for $t < T$ can be realized by integrating an incoming rectangular pulse. The output is taken at $t = T$ and then the integrator is reset (dumped) and the procedure is repeated. A relatively simple circuit to accomplish this *integrate-and-dump* filtering is shown in Fig. 7.48(c). For $RC \gg T$, the circuit acts like an integrator on the incoming pulse waveform. The output $g(t)$ is sampled at $t = T$ and then the capacitor is shorted momentarily to reset the integrator. The integrate-and-dump filter thus forms a relatively simple realization of a matched filter for rectangular pulses. Note that synchronization is required for the sampling and reset operations.

DRILL PROBLEM 7.11.1

Compute the matched-filter (physically realizable, minimum-delay) output over $(0, T)$ to the pulse waveform

$$f(t) = \begin{cases} \exp(-t) & 0 < t < T, \\ 0 & \text{elsewhere.} \end{cases}$$

ANSWER: $\exp(-T) \sinh t$.

7.12 MATCHED-FILTER CODEWORD DETECTION

In the previous section we discussed the matched-filter detection of pulse waveforms on a pulse-by-pulse basis. We can easily extend these concepts to the next level of detection—that of codewords. This has application to the problem of an optimum detector of PCM in the presence of white noise.

Our interest in the use of the matched filter for PCM centers on the design of a filter that will yield a maximum output only for a particular codeword. If we can design this filter, then we can use n such filters, each set for one of the n codewords available in the code. These filters can be followed by a decision to determine which output is maximum. The appropriate signal level associated with the codeword chosen is then generated and the receiver has performed its detection role at a particular sample point.

Both the tapped delay line and the time-correlator realizations of the matched filter can readily be extended to the case of codewords. We shall restrict our consideration here to the binary case. This makes the required circuitry easy to build using digital-logic elements. As a result of the similarities between the two matched-filter methods for the binary case, we shall refer to both of them as correlators. Both methods require bit and word synchronization.

The design of an optimum PCM receiver now becomes evident. A code with n possible codewords is chosen and then n correlators, each with a stored replica of one codeword, are used. The output of the n correlators is supplied to decision logic that determines which correlator output is largest at the sample time $t = t_m$. The appropriate signal amplitude is then assigned to that particular sample point and the cycle is repeated. Such a receiver configuration is shown in Fig. 7.49. The reader will find it interesting to compare this diagram with that of Fig. 2.5.

Often these operations are performed digitally and the procedures are illustrated here with an example. Suppose that we decide to use a 16-level binary PCM code ($n = 16$) and for simplicity we assume that there is no noise present. At the receiver, each binary code symbol is stored in memory and the required multiplication and summation is performed digitally on command of a clock. For convenience, we designate a one by "+" and a zero by "−". Let us consider the correlator output before the sample switch for the code symbol of level 13 (that is, 1101, represented by $++-+$).[†] Upon command of the clock (which of course must be synchronized to the incoming bit stream), the incoming data are multiplied by the stored code symbols in each of the correlators and supplied to accumulators (i.e., integrators). At the next clock command, the incoming data are shifted by one bit interval and the operation is repeated.

[†] The usual (0, 1) binary notation can be used if the appropriate equivalence relation is defined for the multiplier ("EXCLUSIVE NOR"), although it will be shown in Section 10.3 that best performance is obtained if the binary signal states have equal energy and opposite signs.

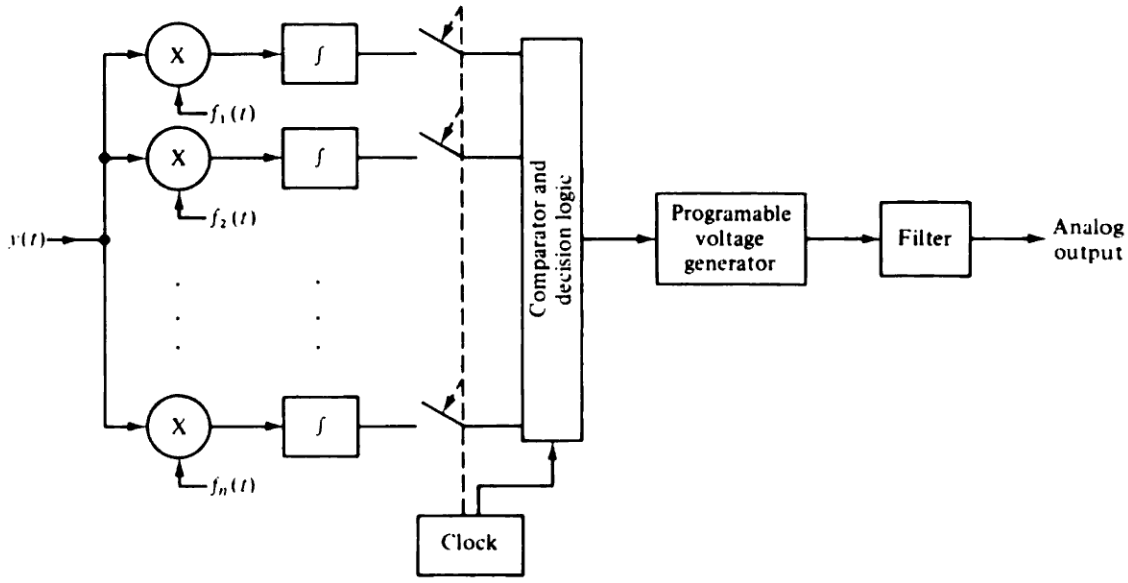


Figure 7.49 An optimum receiver for PCM using correlation.

Let us assume that the incoming data are in the sequence 6, 13, 10, and that the correlator is set for the code symbol for 13. A listing of the successive correlator outputs before the sampling switch is shown in Table 7.3 and the results are graphed in Fig. 7.50. Note that the output is a maximum for $\tau = t_m$ and is proportional, at that point, to the signal energy. Closing the switch at that point assures us of the maximum output. However, this also requires receiver synchronization and the receiver must therefore maintain word synchronization as well as bit synchronization.

TABLE 7.3 Output of a Digital Correlator for the Binary Representation of 13 in the Sequence 6, 13, 10

Relative Shift, $\tau - t_m$	Incoming Data/Stored Replica		Accumulator Output $= r_h(\tau - t_m)$
	10 13 6		
−3	+−+−++−+−++−	y	
	++−+	f	0
−2	+−+−++−+−++−		
	++−+		+2
−1	+−+−++−+−++−		
	++−+		−2
0	+−+−++−+−++−		
	++−+		+4
+1	+−+−++−+−++−		
	++−+		−2
+2	+−+−++−+−++−		
	++−+		0
+3	+−+−++−+−++−		
	++−+		+2

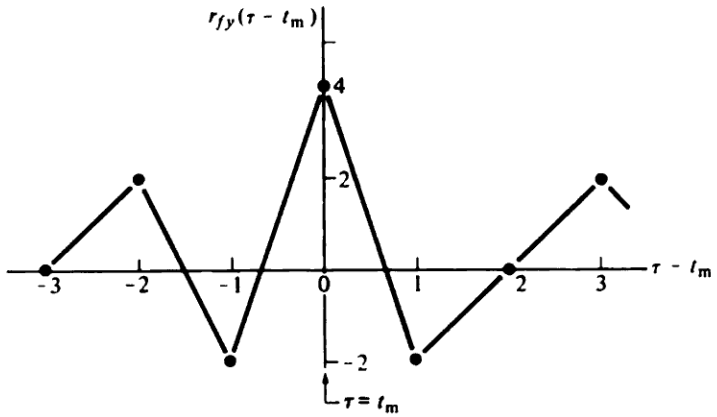


Figure 7.50 Correlator output for the PCM sequence of Table 7.3.

The necessity for word synchronization can be illustrated by considering the binary sequence for 3, 7 as the code enters the correlator shown in Table 7.3. The last two pulses of 3 and the first two pulses of 7 form the sequence $++-+$, matching the binary codeword for 13. Thus there will be a maximum correlator output midway between the processing for 3 and for 7. With word synchronization, however, this maximum will never occur at the correct time at which the output is taken unless the codeword for 13 is present.

Correlators can also be used to initially attain synchronization. In this case the receiver must search for the maximum correlator output. This search would be relatively easy if the correlator had an output only for $\tau - t_m = 0$. But this is not always the case, as can be seen from Fig. 7.50. One of the subjects of coding theory is to design codes which have minimum correlation for $\tau - t_m \neq 0$ (the maximum is proportional to the signal energy).

DRILL PROBLEM 7.12.1

A four-bit digital correlator similar to the one used for Table 7.3 is set for the binary code for 10 ($+-+-$). An infinite sequence of binary codes for a given number (0–15) is applied to the input. Determine the correlator output if the number is (a) divisible by 3; (b) divisible by 5 but not by 3.

ANSWER: (a) 0; (b) ± 4.

7.13 PSEUDONOISE (PN) SEQUENCES

To aid in synchronization, particularly in initial synchronization, it is desirable to use a codeword whose autocorrelation is as large as possible at zero shift and as small as possible at all other shifts. Upon reception, the receiver can determine, upon recognition, that transmission is about to begin and establish an accurate timing reference for

synchronization. This is particularly important in those transmission systems, such as satellites, that may not transmit data on a continuous basis.

Although we can envision codewords that, to accomplish this goal, are stored in memory both at the transmitter and at the receiver, it turns out that some relatively simple linear systems are available to generate codewords with surprisingly good correlations. These codewords are generated using shift-register sequence generators.

A shift register is a cascaded series of one-stage binary memories. The binary state of each memory is transferred to the next memory at the command of a clock. Thus the output of a three-stage shift register is the input delayed by three clock intervals; and at any one time, this shift register will contain the three successive most recent bits in the input bit stream. A shift-register sequence generator consists of an n-stage shift register with taps from certain stages and the output connected to logic combiners that feed back to the input. An example of a three-stage shift register sequence generator is shown in Fig. 7.51. In this example, the input to the shift register is the modulo-2 (i.e., binary add without carry) sum of the content of stages S_2 and S_3, and the output is read from the content of S_3.

Suppose we initialize the generator of Fig. 7.51 by loading all 1's into the shift registers. The two inputs to the modulo-2 ADD operation are 1 and 1; the result, 0, is held by the buffer. In the next clock interval, the content of S_2 is transferred to S_3, the content of S_1 is transferred to S_2, and the content of the buffer is transferred to S_1. The modulo-2 ADD operation again yields 0 and is held by the buffer for the next clock command. Continuing the sequence, we find that the three shift registers have the following contents:

$$111, 011, 001, 100, 010, 101, 110, 111, \ldots.$$

Because this generator uses information from only the three most recent bits, once the initial state (111) repeats, the entire remaining sequence will repeat. The generator output is taken from the content of the last shift register to give the output codeword:

$$1110010\ldots.$$

The output codeword is periodic, with a period of seven bits.

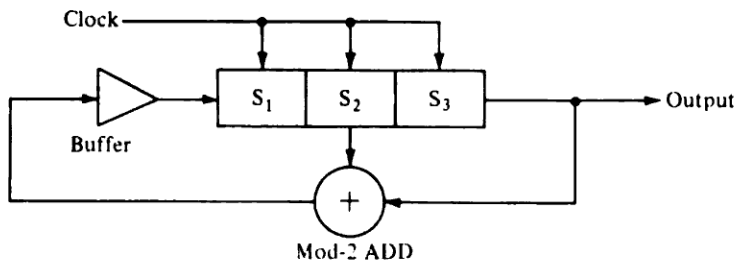


Figure 7.51 A three-stage shift register sequence generator for generation of a seven-bit PN sequence.†

† Often the buffer is included as part of the modulo-2 ADD operation. It is shown separately here for clarity.

In the progression of shift register contents above, every possible three-bit combination occurs with the sole exception of the all-zero word. To the informed observer, this progression is fixed by the number of shift register stages, and the choice of taps to apply to the modulo-2 ADD operation. To the uninformed observer, however, the progression (within a seven-bit period) appears to be entirely random. Therefore these sequences are sometimes called "pseudorandom" sequences.

The correlation properties of the above output sequence are of particular interest to us. Correlating the periodic sequence above with a seven-bit replica yields the correlation pattern shown in Fig. 7.52(a) for the on-off binary case (0, 1) using the EXCLUSIVE-NOR operation for multiplication, and the correlation pattern shown in Fig. 7.52(b) for the polar binary case $(-1, +1)$. In either case the maximum is given by the normalized energy (seven in this case) at zero shift. The correlation decreases to a plateau level for all other shifts up to one period away. As a result, the power spectral density approaches a white spectral density as the sequence length increases. Because these sequences are pseudorandom with white power spectral density, they are called *pseudonoise* (PN) sequences.

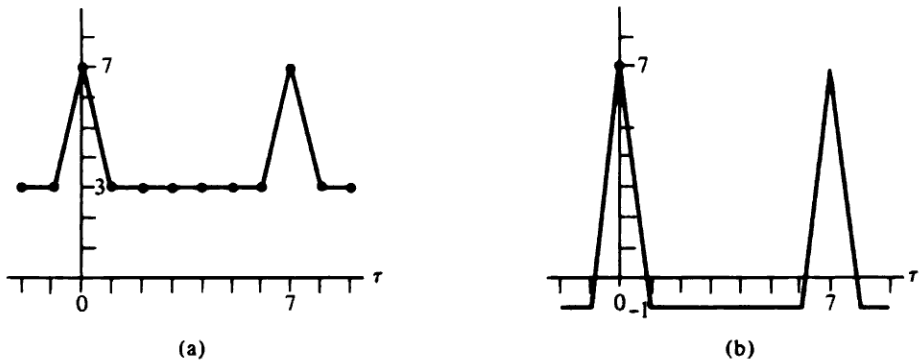


Figure 7.52 Correlation of a seven-bit PN sequence for (a) on-off binary, and (b) polar binary output waveform.

PN sequences are members of a class of codes known as maximal-length codes. By definition, maximal-length codes are the longest codes that can be generated by a shift register of given length. Some properties of maximal-length codes for the binary case are listed below.

1. The length of a maximal-length code that can be generated by n shift registers is $2^n - 1$.
2. With the exception of the all-zero word, every possible n-bit word exists in the n-stage shift register at some time during the generation of a complete code cycle, and only once during the cycle.
3. The number of 1's in an output sequence is one greater than the number of 0's.
4. The correlation of a code sequence with a one-period replica results in a peak equal to the sequence length $2^n - 1$ at zero shift, and at multiples of the sequence length. At all other shifts, the correlation is $2^{n-1} - 1$ for the on-off binary case (0, 1), and -1 for the polar binary case $(-1, +1)$.

Not all tap combinations of an n-stage shift register will yield a maximal-length code. Proper feedback connections for several values of n are given in Table 7.4. Tables for many other feedback connections are available in the literature.[†]

The difference between the correlation value of a code at zero shift and the maximum correlation value at any nonzero shift (up to one period) is called the *index of discrimination*. The larger the index of discrimination the better the code is suited for synchronization. For a maximal-length code, the index of discrimination is $(2^n - 1) - (-1) = 2^n$ for the polar binary case and $(2^n - 1) - (2^{n-1} - 1) = 2^{n-1}$ for the on-off binary case. Some degradation may be expected in the presence of noise, and design procedures usually allow for an error of one or two bits in a codeword.

TABLE 7.4 Feedback Connections for PN Code Generation

n	Sequence Length	Feedback Taps (for combination with output)	Output Sequence (initial state: all ones)
2	3	S_1	110
3	7	S_2	1110010
4	15	S_3	111100010011010
5	31	S_2	1111100110100100001010111011000

PN sequences are used widely in digital communication systems for synchronization. Often, if initial synchronization is the objective, several cycles of the sync codeword are sent and the receiver is required to identify two successive sync words. In a typical sync-search procedure, a bit (or group of bits) in the received bit stream is loaded into a receiver shift register and compared to the contents of a second register containing the sync codeword. If no match occurs, the process is repeated. If a match occurs, the receiver begins inputting additional bits and verifying. On verification of two successive sync words (within a given tolerance for error of, say, two bits), the receiver confirms that initial sync has been achieved. In conservatively designed systems, up to five sync words are sent to permit initial synchronization. A special codeword (meaning one that is distinct from the PN sequences) is placed after the PN sequences and becomes the signal that data transmission is about to begin.

PN sequences may be used also for frame synchronization. In this case, the sync word occurs periodically in the data transmission. A phase-locked loop (PLL) is used in the receiver to track this sync word and to provide accurate synchronization. The VCO in the PLL is a PN sequence generator.

PN sequences are used also in many other types of systems. Because of their good correlation properties, PN sequences find applications in ranging systems. In some systems, such as radar, the correlation of an aperiodic (one-shot) PN sequence is more important than that of a periodic sequence. An example of binary sequences with good aperiodic correlation properties are the Barker codes (Table 7.5). A Barker code of length N has an aperiodic autocorrelation with sidelobe levels no greater than

[†] See Reference 8.

TABLE 7.5 The Barker Codes

$N = 2$	++
	+−
$N = 3$	++−
$N = 4$	++−+
	+++−
$N = 5$	+++−+
$N = 7$	+++−−+−
$N = 11$	+++−−−+−−+−
$N = 13$	+++++−−++−+−+

$1/N$ times that of the main lobe. The longest known Barker code is of length $N = 13$.[†] (All of the known Barker codes are listed in Table 7.5.) PN sequences may also be used as discrete white-noise generators that test systems and scramble signal transmissions.

DRILL PROBLEM 7.13.1

The output of a six-stage PN sequence generator uses the polar binary representation −1V, +1V. (a) What is the (error-free) index of discrimination? (b) What is the average value of the sequence?

ANSWER: (a) 36.1 dB; (b) 15.9 mV.

7.14 SUMMARY

In pulse modulation, sample values are sent by varying a parameter in a train of pulses. The minimum sampling requirements are set by the sampling theorem.

In analog pulse modulation, the height, width, or position of the pulses is varied in direct proportion to the signal amplitude at the sampling instants. In pulse-code modulation, a code is changed in discrete steps in proportion to the quantized signal amplitude at the sampling instants.

Pulse-amplitude modulation (PAM) resembles conventional sampling except that the pulse tops are flat. Although this difference can be adjusted with equalizing filters, it is minimized by use of narrow pulse widths. This leaves space for interleaving pulses from other sampled signals. This method of combining several sampled signals in a definite time sequence is called time-division multiplexing (TDM). The minimum bandwidth to handle a TDM waveform must obey the requirements of the sampling theorem. PAM offers no signal-to-noise advantage.

[†] For added reading on coded radar signals, see, e.g., N. Levanon, *Radar Principles*, New York: John Wiley & Sons, 1988, Chapter 8.

Interference between pulses as a result of a finite bandwidth in a pulse modulation system is called intersymbol interference (ISI). Intersymbol interference can be controlled by widening the transmission bandwidth and by choosing the transmission spectral shaping. A good choice of the latter are the Nyquist waveforms, which have uniformly spaced zeros at multiples of a basic time interval governed by the bandwidth.

Pulse-timing modulation employs a variation in the timing of pulses to convey information. Examples of pulse-timing modulation are pulse-width modulation (PWM) and pulse-position modulation (PPM). Because pulse-edge response is important to the demodulation of pulse-timing modulation, it requires a much wider bandwidth than PAM. It also offers a signal-to-noise advantage similar to that for angle modulation.

Pulse-code modulation (PCM) uses a series of discrete code symbols to convey information. This type of modulation offers more versatility in signal design to combat effects of various types of noise. Different types of codes may be used, depending on system constraints on bandwidth, synchronization, etc.

The basic quantum unit for conveying information in a binary code is called the bit. Limiting bandwidth efficiency for a binary PCM system is two bits per second per Hz.

Fiber optic communication systems offer extremely wide bandwidths as a result of the very high carrier frequencies. Transmission through fibers is immune to electromagnetic interference and differences in electric potentials. Losses in fibers are relatively low, making such systems attractive for terrestial PCM systems.

One way to increase the reliability of communication is to add extra bits, at the expense of added transmission time or bandwidth, in order to detect and, possibly, to correct the errors. The use of parity-check bits and repetition codes are simple and effective methods to accomplish this objective.

Two basic types of digital data transmission are synchronous and asynchronous. In synchronous systems, all sampling and multiplexing is governed by a master clock. Asynchronous multiplexers are operated slightly faster than would be necessary for comparable synchronous systems. Either extra bits (called bit stuffing) or extra words (called word stuffing) are inserted to keep the inputs up to the desired rate. Intelligent multiplexers that can alter the data-handling and routing characteristics are called concentrators.

The Integrated Services Digital Network (ISDN) is an attempt to provide digital communication services to a wide range of users in the United States via a single interface. While centered on telephone services, it offers digital data services at a basic rate of 64 kbps. It takes advantage of packet-switching networks for the increased flexibility in combining various data rate requirements.

Detection of PCM demands the recognition of a given code symbol in the presence of noise. The optimum linear time-invariant filter to accomplish this task in the presence of white noise is called a matched filter. A matched filter is the optimum linear time-invariant filter that gives the highest peak-signal–to–rms-noise ratio in the presence of white noise. The matched filter can be realized using tapped delay lines or cross-correlation.

The principles of matched-filter detection can be extended to codewords. Codewords with good autocorrelation properties are desirable for synchronization in

matched-filter detection. A good autocorrelation function for synchronization is one that is maximum at zero shift and zero (or nearly zero) for all other shifts. Pseudonoise (PN) sequences provide good periodic autocorrelation properties, and they are relatively easy to generate using shift registers and digital logic.

Selected References for Further Reading

1. J. Martin. *Telecommunications and the Computer*. Second ed. Englewood Cliffs, N.J.: Prentice Hall, 1976.
   A qualitative description of data transmission methods using telephone networks; includes a glossary for definition of terms in telephony; easy reading.
2. J. A. Betts. *Signal Processing, Modulation and Noise*. New York: American Elsevier, 1970.
   Chapter 6 describes some effects of inadequate band-limiting in TDM methods.
3. H. S. Black. *Modulation Theory*. Princeton, N.J.: D. Van Nostrand, 1953.
   A good discussion of sampling theory and applications to pulse modulation.
4. M. Schwartz. *Information Transmission, Modulation and Noise*. Third ed. New York: McGraw-Hill, 1980.
   Chapter 3 contains a good discussion of sampling, time-division multiplexing, and waveshaping for PCM, and uses typical operating systems as examples.
5. J. C. McDonald. *Fundamentals of Digital Switching*. New York: Plenum Press, 1983.
   Covers the topics of switching systems and network synchronization.
6. A. P. Clark. *Principles of Digital Data Transmission*. Second ed. London: Pentech Press, 1983.
   Chapters 4, 7, 10 cover FDM, TDM, transmultiplexers, network timing, and synchronization methods for telephone networks.
7. J. R. Pierce and E. C. Posner. *Introduction to Communication Science Systems*. New York: Plenum Press, 1980.
   An interesting approach to matched filters and correlation is presented in Chapter 7.
8. R. C. Dixon. *Spread Spectrum Systems*. Second ed. New York: John Wiley & Sons, 1984.
   Chapter 3 contains a good discussion of maximal-length sequences and their generation. The remainder of this book makes interesting collateral reading.
9. P. Z. Peebles, Jr. *Communication Systems Principles*. Reading, Mass.: Addison-Wesley, 1976.
   Chapter 7 is an advanced discussion of sampling and S/N performance of PWM and PPM, including threshold effects.

PROBLEMS

7.1.1 The signals given below are not band-limited, so that when they are sampled some aliasing will result. Because the signals are low-pass, the primary aliasing will arise from those components at frequencies for which $\omega_s/2 \leq \omega < \omega_s$, where ω_s is the sampling frequency. Determine the minimum value of ω_s, in terms of the number of -3-dB bandwidths, for each of these signals so that the signal energy in the range $\omega_s/2 \leq \omega < \omega_s$ is down at least 10 dB from the energy in the frequency range $0 \leq \omega < \omega_s/2$.

a) $f(t) = e^{-at}u(t)$.

b) $f(t) = te^{-at}u(t)$.

7.1.2 A certain signal $f(t)$ is band-limited to f_m Hz. A given system samples this signal at the Nyquist rate. Each pulse out of the sampler is of the form $a_k \exp[-(t - t_k)/(RC)]u(t - t_k)$, where k designates the sample.

a) Make sketches similar to those in Fig. 7.3(c–f) for this case.
b) Can you devise an equalization filter that will yield samples which match those from an impulse sampler? Explain.
c) Can you devise an equalization filter that will yield PAM from this sampled signal? Explain.

7.1.3 The signal $f(t) = [W/(2\pi)]\mathrm{Sa}^2(Wt/2)$ is transmitted through a linear time-invariant system with a frequency transfer function $H(\omega) = F(\omega)$. The output $g(t)$ is sampled by an impulse sampler (i.e., a sequence of uniformly spaced impulses) at the Nyquist rate.

a) Sketch the spectral density of the sampled signal.
b) The sampled signal of part (a) is applied to a low-pass filter $H_1(\omega)$ with output $x(t)$. Determine and sketch $H_1(\omega)$ for the distortionless transmission of $f(t)$.

7.1.4 A sampling probe is sometimes used for observation of high-frequency periodic waveforms on an oscilloscope whose bandwidth is insufficient for a direct observation. The probe consists of a bandpass prefilter with bandwidth B_p, a high-speed sampler operating at rate f_s, followed by a low-pass filter with bandwidth B_x. The resultant lower-frequency periodic waveform is displayed on the oscilloscope.

a) Suppose that the high-frequency waveform is a sinusoid at $f = f_0$. Using the frequency domain and assuming an impulse sampler, determine the minimum and maximum bandwidth of the low-pass filter if the sampling probe operates at $f_s = (1 - \alpha)f_0$, $\alpha < 1$. Also, sketch to scale the resulting displayed waveform.
b) Repeat part (a) for the case in which n harmonics are necessary for a good representation of the periodic high-frequency waveform.
c) What is the purpose of the bandpass prefilter? What are the restrictions on the bandwidth B_p?

√ **7.1.5** Compute the trigonometric Fourier series coefficients for the periodic waveform:

$$x(t) = \begin{Bmatrix} \sin \omega_0 t & 0 \le t < T/4 \\ 0 & T/4 \le t < T \end{Bmatrix}.$$

Sample $x(t)$ at 10° increments over one period.

a) Tabulate the coefficient values for the first 40 harmonic terms ($0 \le n \le 40$).
b) Compute and plot the magnitude of the exponential Fourier series coefficients for the first 40 harmonic terms ($0 \le n \le 40$).
c) Using integration, the trigonometric Fourier series coefficients for $x(t)$ are

$$a_0 = 1/(2\pi);\ a_1 = 1/(2\pi);\ b_1 = 1/4;$$

$$a_n = [n \sin(n\pi/2) - 1]/[\pi(n^2 - 1)] \quad \text{for } n > 1;$$

$$b_n = [-n \cos(n\pi/2)]/[\pi(n^2 - 1] \quad \text{for } n > 1.$$

Using these results, plot the magnitude of the exponential Fourier series coefficients for the first 40 harmonic terms ($0 \le n \le 40$).

d) Do your results display any expected symmetry conditions? Explain.
e) Would you describe the equivalent sampling process in this case as an impulse sampler? Explain.

7.2.1 The first-order hold circuit shown in Fig. P–7.2.1 is an alternative to the sample-and-hold circuit for recovery of PAM signals. The input $f_s(t)$ is a sampled version of $f(t)$ and the delay is equal to the sampling period T.

a) In contrast to the sample-and-hold circuit, this is a linear time-invariant system; determine its impulse response and frequency transfer function.

b) Compare the magnitude frequency transfer function with that of an ideal low-pass filter.

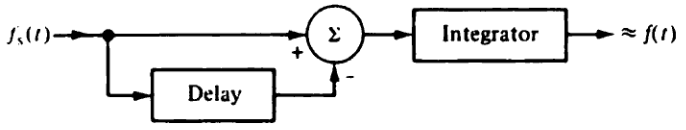


Figure P–7.2.1.

7.2.2 Suppose that the single-delay unit in Fig. P–7.2.1 were replaced with a unit having the impulse response $h(t) = 2\delta(t - T) - \delta(t - 2T)$; i.e., a unit having two parallel signal paths, one with a $\times 2$ amplifier and delay T and the second with a double delay time. Also, the integrator is changed to a double integrator.

a) Sketch a block diagram of this new system.

b) Determine the impulse response and the frequency transfer function of the system.

c) Compare the magnitude frequency transfer function with that of an ideal low-pass filter.

7.2.3 Twenty-four audio input signals, each band-limited to 3.3 kHz, are to be multiplexed and transmitted using FDM or TDM. The frequency-multiplexed systems allow 20% over minimum bandwidth for a guard band, and the time-multiplexed systems sample at 20% over the Nyquist rate. Determine the minimum bandwidth required for each method if the multiplexing and modulation used is a) FDM, SSB-SC; b) FDM, DSB-SC; c) TDM, PAM; d) PAM/AM; e) PAM/PM, $\beta = 9$.

7.2.4 A certain three-channel PAM system is shown in Fig. P–7.2.4; $f_1(t)$ is band-limited to 7 kHz and $f_2(t)$, $f_3(t)$ are each band-limited to 3 kHz.

a) Determine the minimum and maximum values of the commutator clock rate such that all three signals can be reconstructed at a receiver without distortion under ideal conditions.

b) Sketch a receiver block diagram for this system.

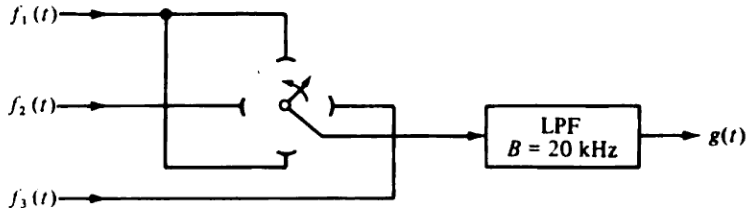


Figure P–7.2.4.

7.3.1 Twenty-four low-pass signals of equal bandwidth are time-multiplexed into a single channel using PAM, within a maximum bandwidth constraint of 150 kHz. Determine the maximum input-signal bandwidth if the output filter before transmission is a raised-cosine filter with a) $\alpha = 0$; b) $\alpha = 0.50$; c) $\alpha = 1$. d) Is there any advantage in using raised-cosine filters also for the input filters? Explain.

7.3.2 Two unit impulses are applied to the input of an ideal low-pass filter with bandwidth B. They are spaced so that the first zero crossings in the $(\sin x/x)$-type output waveforms coincide.

a) Determine the time difference between impulses (in terms of B).
b) Determine the largest sidelobe magnitude (express your answer in terms of dB, referenced to the peak value).
c) Repeat parts (a), (b) for the same pulse spacings if a raised cosine filter ($\alpha = 1$) were used whose -6-dB bandwidth is equal to B.

7.3.3 A PAM/AM communication system is to be designed for transmission of voice and data from a remotely located site. The requirements are for one voice channel, two medium-speed data channels, and four lower-speed data channels. The bandwidth allotment is 3.3 kHz for the voice channel, 800 Hz for each medium-speed data channel, and 400 Hz for each lower-speed data channel. To make allowance for adequate low-pass filtering in the receiver, all channels are oversampled by 20%. A raised-cosine filter ($\alpha = 0.50$) is planned for the filter following the multiplexer. Efforts should be made to conserve bandwidth, within the above constraints. The following two items are needed in a preliminary design study:

a) A timing diagram for the multiplexer; and
b) A plot of magnitude, in decibels, versus frequency for the PAM/AM transmission at a carrier frequency f_c.

7.3.4 The frequency transfer functions for each of three filters are given below for the positive frequency range. Assume that these frequency transfer functions are even in frequency, and equal to zero for $\omega > (1 + \alpha)W$. Determine whether each satisfies the Nyquist waveform criterion, within a constant, and find the corresponding impulse response of each.

a) $H_1(\omega) = 1 - u[\omega - (1 - \alpha)W][\omega - (1 - \alpha)W]/(2\alpha W)$
b) $H_2(\omega) = 1 - (\frac{1}{3})u[\omega - (1 - \alpha)W] - (\frac{1}{3})u[\omega - W] - (\frac{1}{3})u[\omega - (1 + \alpha)W]$
c) $H_3(\omega) = \begin{Bmatrix} 1 - 2[\omega/(2W)]^2 & \omega \le W \\ 2[(\omega - 2W)/(2W)]^2 & W < \omega \le 2W \end{Bmatrix}$

7.3.5 A number of channels are time-multiplexed using PAM. Each pulse is 5τ seconds in duration; neglect any guard time between channels. The resulting TDM pulse train is passed through a RC low-pass filter with $RC = \tau$. Assume that in the receiver the signal is gated with pulse width 5τ and then demultiplexed. We define a "percentage cross-talk" criterion as that signal energy from an adjacent time slot that "spills over" into the desired time slot, divided by the desired signal energy (in the correct time slot) when the pulses are of equal magnitude.

a) Using this criterion, calculate the percentage of cross-talk.
b) Investigate if a small delay in receiver timing will improve this situation.

7.4.1 A given PWM signal has fixed leading edges at the times $t = mT$. Sketch typical waveforms to show under what conditions the RC low-pass filter will approximately reconstruct the modulating waveform.

7.4.2 The fact that the pulse transitions are all positive at the clock times in Problem 7.4.1 is a clue that the pulse train is "self-clocking"—i.e., that there is a spectral line from which a phase-locked loop can extract the clock frequency. To see this, sketch a PWM waveform, differentiate once, and separate the result into the sum of two functions, one with positive transitions and one with negative.

a) Determine the spectral density of the function containing the positive transitions. (Analysis of the function containing the negative transitions is more complicated, and not attempted here.)
b) Is a PPM waveform "self-clocking?" Explain.

7.4.3 A possible receiving system for pulse timing systems in which it is known that there is zero average value in the modulating signal is proposed as follows. Timing points are determined and very narrow positive pulses are generated at the half-period ($T/2$) points. The trailing edges of the incoming pulse timing waveform are detected and these generate narrow negative pulses. The resulting pulse train is applied to a constant-amplitude pulse generator with the following logic operation. If the first pulse to arrive at the generator within a clock period is positive, then a positive level is generated until a negative pulse arrives to terminate it. If the first pulse to arrive at the generator within a clock period is negative, then a negative level is generated until a positive pulse arrives to terminate it. The resulting waveforms are then smoothed with a low-pass filter.

a) Will this system work satisfactorily for PPM? Explain.
b) Will this system work satisfactorily for PWM? Explain.

7.4.4 A given single-channel PPM system has the following specifications:

Input signal:	$\lvert f(t)\rvert < 10$ V
Desired accuracy:	± 0.1 V for $[S/N]_{pk} = 10$ dB
Average pulse rate:	8 kHz
Pulse width:	5 μsec

a) Estimate the minimum bandwidth necessary for transmission in the presence of noise using a guard time equal to the pulse width.
b) Let the desired accuracy be Δa. Plot B versus Δa for $[S/N]_{pk} = 10, 20, 30$ dB.

7.4.5 Two channels of a PWM system having equal bandwidths (f_m) are time multiplexed. The width of each pulse, w, is given at the sample times t_i by

$$w(t_i) = [5 + 0.5f(t_i)]\ \mu\text{sec},$$

where $f(t)$ is restricted to positive values and is measured in millivolts. The desired accuracy of the system is ± 0.5 mV and the guard time between channels is set at 5 μsec.

a) What is the minimum bandwidth required for transmission?
b) Plot the input amplitude range versus the input bandwidth f_m for this system.

7.4.6 Twenty-four audio input signals (each of bandwidth 0–3.3 kHz) are to be sampled and time-multiplexed using PPM. The output is filtered to the minimum bandwidth. Preliminary design specifications are as follows. The inputs are to be oversampled by 20%, the estimated $[S/N]_{pk}$ ratio is 10 dB, the absolute accuracy is 5%, and the pulse width and guard time per channel in the TDM output is three times the pulse rise time. Use a trapezoidal pulse approximation, as shown in Fig. 7.19, to make preliminary design estimates on the system, including

a) bandwidth required for the output filter;
b) pulse width and guard time allotments.

7.5.1 Let the input to a PPM system be a sinusoid, $a \cos(2\pi f_m t)$, where a and f_m are assumed to be the maximum values permitted by the system. Let τ_p be the pulse width and τ_g be the guard time. Develop an equation to replace Eq. (7.30) for these conditions; express your answer in terms of (B/f_m), $(f_m\tau_p)$, and $(f_m\tau_g)$.

7.5.2 A 3-kHz sinusoid with a peak amplitude of 2 V is sampled at an 8-kHz rate and transmitted using PPM. The pulse width τ_p and the guard time τ_g are each constrained to be equal to $5/B$. The S/N improvement factor desired is 30 dB.

a) Determine the minimum transmission bandwidth required within these constraints.
b) To obtain this S/N improvement using FM, how much bandwidth is required?

7.5.3 A single-channel PPM system is operated at a sample rate of 25 kHz with sinusoidal modulation at the maximum peak-to-peak amplitude. In the receiver, the signal pulses have a peak amplitude of 1 V. The half-amplitude width of the pulses is 4 μsec, and the rise and fall times are each 0.5 μsec. (Approximate the rise and fall times as linear line segments to yield a trapezoidal pulse shape.) The rms pulse edge deviation as a result of additive noise is 0.05 μsec. The guard time allowance is 4 μsec.

a) What is the minimum bandwidth required for transmission?
b) Find the peak and the average input signal-to-noise ratios.
c) Determine the signal-to-noise improvement factor of the system.

7.6.1 Demonstrate that the magnitude of the RB code representation is self-clocking (cf. Problem 7.4.2).

7.6.2 a) Show that $2^m - 1$ comparators are needed for a m-bit parallel encoder.
b) A $(m/2)$-bit parallel encoder feeds a D/A converter (DAC). The output of the DAC is subtracted from the input voltage, and the difference is converted by a second $(m/2)$-bit parallel encoder. Show that this arrangement can be a valid substitute for the m-bit parallel encoder in part (a). Compare the two methods.

7.6.3 A four-bit D/A converter (such as could be used in Fig. 7.28) is shown in Fig. P–7.6.3.

a) Explain the operation of the circuit.
b) A difficulty in accuracy of this D/A converter is that of accuracy of resistance values. Suppose that the feedback resistance R is taken as the reference, and let the switched resistances be $(R + \Delta R_1)$, $(R + \Delta R_2)$, etc. Determine the rms value over all measurement errors. Assume that all codes are equally likely and that the ΔR's are uncorrelated and the $\Delta R/R$'s are small and equal.

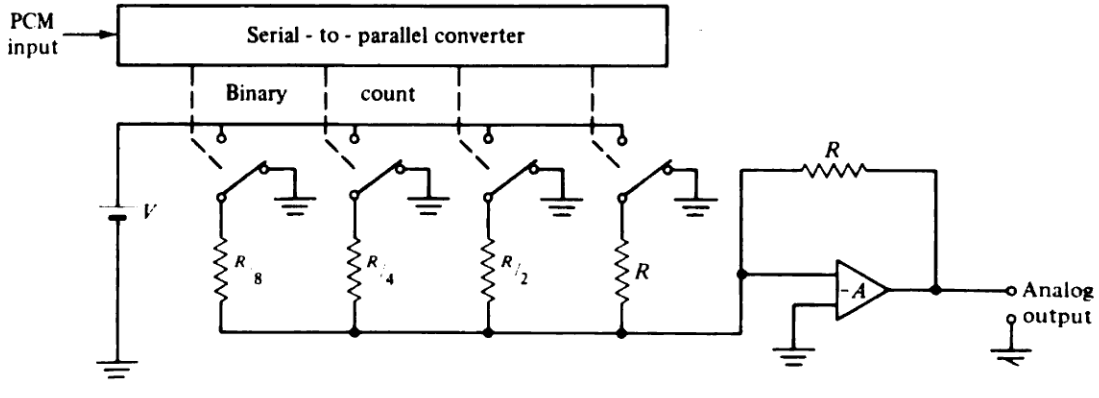


Figure P–7.6.3.

7.6.4 Consider the binary sequence 1011000101. Using vertical and horizontal lines as guides, sketch the binary waveforms for the following code representations. (Begin in the low state for those representations that depend on a prior state.)

a) Split phase (Manchester)
b) RZ
c) NRZ (M)
d) Delay modulation (Miller)

7.6.5 In this problem, we assume the return-to-zero (RZ) code representation but investigate the use of more than two levels. Suppose that we wish to send a binary input sequence at a clock rate of 64 kbps.

a) Determine the minimum bandwidth required for transmission using two-level RZ.
b) Instead of sending the pulse train in part (a), we generate one pulse having four possible levels for every group of two successive pulses in the input sequence. (We retain the RZ representation, and allow the zero level as one of the four choices.) Determine the minimum bandwidth required.
c) Repeat part (b), but assuming that three successive input digits are combined into one pulse with eight possible amplitudes.

7.6.6 Assume that the maximum usable bandwidth (based primarily on phase distortion) of a telephone channel is 2400 Hz, and that PCM/AM is used. Also assume that raised-cosine spectral shaping, $\alpha = 0.5$, is used.

a) Determine the maximum bit rate using a two-level transmission.
b) Repeat part (a), but assume that two successive input digits are combined into output pulses having four possible levels in the transmission.
c) Repeat part (a), but assume that three successive input digits are combined into output pulses having eight possible levels in the transmission.
d) What prevents us from continuing this progression?

7.6.7 Twenty-four input signals, each band-limited to 3.3 kHz, are each sampled at an 8-kHz rate and time multiplexed. Calculate the minimum bandwidth required for transmission using

a) five-bit PCM;
b) seven-bit PCM.
c) Repeat parts (a), (b), but assume AMI is used.

7.6.8 Data can be encoded on magnetic ("floppy") disks using the following methods. In the frequency modulation (FM) method, each pulse in the high state exists for one-fourth the clock interval. A one is represented by two cycles of an alternating high/low state pattern. A zero is represented by a high state for one-fourth the clock interval and a low state for three-fourths of the clock interval. In the second method, called modified frequency modulation (MFM), the high-state pulses are one-half the clock interval and the following encoding rule is used. A one is represented by a low state followed by a high state. A zero is represented by either of two patterns: A high state followed by a low state, if no high state occurred at the end of the previous clock interval, or by two consecutive low states if a high state occurred at the end of the previous clock interval. Thus the MFM method guarantees at least one low state between every two high states.

a) Following the directions of Problem 7.6.4, sketch the FM and MFM representations of the binary sequence given in Problem 7.6.4.
b) Estimate the bandwidth efficiency of each method. (Assume an equal mix of ones and zeros, and assume in MFM that one-half of the zeros are represented by each of the two possible representations.)
c) Why do you suppose the former method is called "FM?" Is it self-clocking? Explain.

7.6.9 A series of high-density bipolar representations, known as HDB1, HDB2, HDB3, . . . , can be used to eliminate the long successions of zero-amplitude pulses that may occur in BRZ. The HDBn representations avoid the occurrence of more than n successive pulses of zero amplitude. They use the AMI rules but substitute special sequences for each succession of $(n + 1)$ 0's, and also use AMI rule violations to convey the extra information needed to replace the 0's. The most important of the HDB codes is the HDB3.

If the string 0000 occurs in HDB3, either of the special sequences 000D or 100D may be sent. The "D" is replaced by level +1 or −1 in such a way as to violate the AMI rule at that point. Choice of 000D or 100D is such that pulses violating the AMI rule take on the levels +1 and −1 alternately.

Write out the HDB3 pulse train for the following input: 100001100000000, assuming that 000D is the first special sequence. Then write out the HDB3 pulse train, assuming that 100D is the first special sequence.

7.8.1 Consider the use of the codewords listed in Drill Problem 7.8.1. If only these codewords can be used,

a) How many data bits can be allocated per codeword?
b) How many errors can be detected?
c) How many errors can be corrected?
d) How many possible codewords in the set are unavailable in parts (a–c)?

7.8.2 a) Is it possible to make an undetected digit interchange error in Drill Problem 7.8.2? Explain.
b) Would the rule of tripling alternate digits work as well in this scheme? Explain.

★ **7.9.1** Design an M13 bit-interleaved multiplexer (see Table 7.1) that has the same frame format as that given in Fig. 7.43, with the exception that only two bits per input channel are taken before a control bit is added. In designing this multiplexer, calculate the four quantities requested in Example 7.9.1.

★ **7.9.2** Design an M12 bit-interleaved multiplexer (see Table 7.1) that has the same frame format as the multiplexer in Example 7.9.1, with the exception of the added seven-bit marker word 1110010. In this design, calculate the four quantities requested in Example 7.9.1. Compare these results with those of Section 7.9.1.

★ **7.9.3** A certain character-interleaved, time-division multiplexer with word stuffing uses an all-zero word as the stuff word. In the multiplexer, the m-bit input data word $\mathbf{A} = (a_1, a_2, \ldots, a_m)$ is encoded into the $(m + 1)$-bit output word $\mathbf{B} = (b_1, b_2, \ldots, b_m, b_{m+1})$ according to the following algorithm:[†]

$$b_{m+1} = 0 \quad \text{if } w(\mathbf{A}) \geq \left[\frac{m}{2}\right] + 1$$

$$b_{m+1} = 1 \quad \text{if } w(\mathbf{A}) \leq \left[\frac{m}{2}\right]$$

$$b_i = b_{m+1} \oplus a_i, \qquad 1 \leq i \leq m,$$

[†] M. M. Buchner, Jr., "An Asymmetric Encoding Scheme for Word Stuffing," *Bell System Technical Journal* (March 1970): 379–398; see also J. J. Spilker, Jr., *Digital Communications by Satellite*, Englewood Cliffs, N.J.: Prentice-Hall, 1977, pp. 124–128.

where $w(\cdot)$ is the Hamming distance measured from the all-zero word, the notation $[m/2]$ represents the largest integer $\leq m/2$, and $\oplus$ indicates modulo-2 addition. At the demultiplexer, the bit order is reversed so that b'_{m+1} is detected first and the decoding algorithm is

$$a'_i = b'_{m+1} \oplus b'_i, \qquad 1 \leq i \leq m,$$

where the primes indicate bits in the demultiplexer.

For such a multiplexer using $m = 8$-bit input data words,

a) Make a table of several arbitrary data words and their encoded versions.
b) Estimate the minimum distance of the output words from the all-zero stuff word.

7.11.1 If $f(t)$ is in volts, $H(\omega)$ is dimensionless and $S_n(\omega)$ is in watts per Hz across 1 ohm, what are the dimensions of the constant k in the matched filter?

7.11.2 Let $f(t) = \sqrt{S}\exp[-t^2/(2\sigma_s^2)]$ and let $S_n(\omega) = N_0 \exp(-\sigma_n^2\omega^2)$, where S is in watts and N_0 is in watts per Hertz. The ratio of the signal bandwidth to the noise bandwidth is σ_n/σ_s. Determine the maximum value of $|f_o(t_m)|^2/\overline{n_o^2(t)}$ versus the ratio of signal bandwidth to the noise bandwidth.

7.11.3 A physically realizable filter is matched to a signal approximated by the waveform shown in Fig. P–7.11.3.

a) Determine and sketch the impulse response of the matched filter for minimum delay.
b) Determine and sketch the matched-filter output, versus time, for the minimum-delay case.
c) Sketch the matched filter impulse response for an arbitrary delay time.
d) Sketch the matched filter output for the delay chosen in part (c).

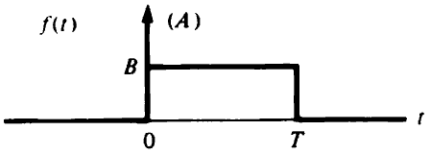


Figure P–7.11.3.

7.11.4 A given pulse signal waveform $f(t)$ is shown in Fig. P–7.11.4.

a) Determine the frequency transfer function of the matched filter for $f(t)$.
b) Develop a relation for maximum peak-signal–to–rms-noise ratio at the output of this filter if the input noise is white with a (two-sided) power spectral density $\eta/2$ watts/Hz across 1 ohm.
c) Repeat part (b) if the input noise is not white, but it is given that the output noise power is N_o if $S_n(\omega)$ is applied to the input of the filter whose frequency transfer function was determined in part (a).

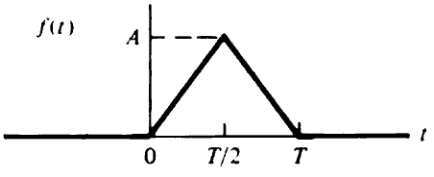


Figure P–7.11.4.

7.11.5 Let a received signal be $f(t) = e^{-at}$ over $(0, T)$, zero elsewhere, in the presence of additive white noise.

a) Determine the matched filter response versus time.

b) Find the time response of the output if $f(t)$ plus additive white noise is applied to a filter whose impulse response is $h(t) = A$ over $(0, T)$, zero elsewhere. Compare your answer with that in part (a); normalize your answers so that the rms noise outputs of the two filters are equal.

7.11.6 a) Determine and sketch the time response of a minimum-delay physically realizable matched filter to the signal $f(t) = \mathcal{R}e\{Ae^{j\omega_c t}\}$ over $(0, T)$, zero elsewhere.

b) Determine the maximum value of $|f_o(t_m)|^2/\overline{n_o^2(t)}$ if $S_n(\omega) = \eta/2$.

7.12.1 Codewords for synchronization can be sent efficiently by modulating the phase of a sinusoid by 0 and π radians and by turning the sinusoid off when not modulating (i.e., at end of codeword). Multiplication of this modulated sine wave by a sinusoid of the same frequency and coded phase in the receiver yields a three-level output in each bit interval: 0, $+1$, or -1. Adding the results of this multiplication in each bit interval over the number of bits in the stored codeword provides the correlation of the input codeword with the stored codeword. Consider the following two codewords: $+ + - +$ and $+ + + -$. Sketch the output of a four-bit correlator that is set for each of these codewords under the following conditions:

a) The codeword is aperiodic (assuming all 0's elsewhere); the correlator is set for the same code as the one that is sent (examine both cases).

b) The codeword is repeated continuously (all other conditions as above).

c) Repeat parts (a) and (b) if the correlator reference is accidently set for the other codeword.

7.12.2 Repeat Problem 7.12.1 using 0 to replace minus $(-)$, 1 to replace plus $(+)$, and an EXCLUSIVE-NOR operation to replace multiplication in the correlator.

√ **7.12.3** Use a digital computer and graphics monitor to compute and display an aperiodic eight-bit codeword and its autocorrelation function using 128 points. Allow the possible amplitudes $+1$, -1, 0 to be set within each bit interval. Display and print the input and output waveform for each of the following codewords:

a) $+ + + - - + -$

b) $+ + + - - + + -$

c) $+ + + - - - + -$

d) Which of the above would you prefer? Explain.

e) Can your program handle the case in which the input codeword is periodic? Explain.

7.12.4 It is suggested that a way to perform the correlation of two sequences of pulses with ± 1 amplitudes is to replace each pulse with an impulse function with area ± 1, and perform the correlation at the various shifts to obtain the weights of a new sequence of impulse functions. It is argued that the autocorrelation function is the correlation of this new sequence of impulse functions with the autocorrelation of one pair of pulse waveforms, normalized to a peak of one. Will this work? Explain. (It may be helpful here to consider correlation as convolution without reversal of one waveform.)

7.12.5 Draw a logic diagram of a correlator for an input binary codeword and a stored replica.

7.13.1 a) Why can't the all-zero condition be allowed in the shift registers of a maximal-length code generator?

b) Demonstrate the effect of an odd number of taps (including the last stage) in a shift register sequence generator. (Begin by examining a four-stage case.)

7.13.2 Consider a code sequence generated by feeding back the first and third stages of a three-stage shift register.

a) Assuming an initial state of 1 1 1, find all possible shift-register states in order of their occurrence.

b) Determine and plot the correlation of a periodic PN sequence output with a one-period stored replica. Use a bipolar representation.

c) Determine and plot the correlation of an aperiodic PN sequence output with a one-period stored replica. Use a bipolar representation.

d) Compare your results with those listed in Table 7.4 for the choice of feedback from the second and third stages.

7.13.3 a) What is the length of a maximal-length code generated by a 15-stage shift register sequence generator?

b) Express the discrimination index of the above PN code, in dB.

c) Repeat parts (a) and (b) for a 16-stage shift register sequence generator.

d) Estimate the worst case index of discrimination in the correlation of a periodic PN sequence with a one-period stored replica if a one-bit error occurs in the input sequence. Assume use of the bipolar representation.

7.13.4 A certain receiving system uses a periodic PN code sequence for accurate synchronization in the presence of additive noise. Synchronization timing in this system is determined by detection of the correlation pulse near zero shift with a preset threshold (e.g., at half-amplitude). Determine the minimum number of shift registers, n, required if synchronization must be held to within $\pm 1\%$ of a bit interval in the presence of additive noise of ± 1. Assume use of the bipolar representation.

7.13.5 The Barker codes (cf. Table 7.5) are bipolar binary code sequences of length N that are designed for autocorrelation peaks N and maximum sidelobe levels (at nonzero shifts) no greater than 1.

a) Sketch the autocorrelation of each of the two $N = 4$ Barker codes in Table 7.5. Assume "0" beyond the code sequence length, and define multiplication by 0 to be zero.

b) Repeat part (a), with the exception that the input code sequence consists of two repetitions (i.e., length $2N$) of the Barker code sequence in the codeword but the receiver remains matched for one (length N).

DESIGN PROBLEMS

Design Problem

7.1 Telemetry System for Model Control

A short-range, low-power, small PPM/AM telemetry system is to be designed for the remote control of models. The system is to have 8 channels and to meet a ±5% accuracy requirement for 20-dB peak S/N at the output of an envelope detector. (Allow a 5-dB margin for detection and postdetection filtering. The bandwidth is narrowed to B Hz after envelope detection.) Maximum frequency response per channel is 20 Hz, and the system is to oversample by 25%. Minimum pulse width and guard time requirements are each $5/B$. The RF section of the receiver has a noise bandwidth of 1 MHz and the total noise temperature (including that of the antenna) is estimated to be 10^5 K. Assume total propagation losses (including path and antennas) to be 90 dB. Make a preliminary design of the system, including: a) minimum bandwidth B of the filter after the envelope detector, b) timing diagrams, and c) an estimate of the minimum peak transmitter power requirement to meet these objectives. State any assumptions you make. d) If the above losses were for a distance of 1 mile, estimate the range of your system for a peak transmitter power of 1 W. (Assume that the path losses, in decibels, increase as 20 log d, where d is in miles.)

CHAPTER

8 Probability and Random Variables

Waveforms whose values can be specified exactly are said to be "deterministic." While deterministic signals cannot always be expressed explicitly, they can, within rather broad limitations, be expressed in terms of summations of explicitly known functions (e.g., the Fourier series). The use of deterministic signals has played a major role in our analysis of communication systems up to this point. Now we consider methods for describing nondeterministic, or "random," waveforms. To do this, we shall first introduce some basic concepts from probability theory.

Our approach to probability is heuristic. Because we are interested in some specific applications to communication systems, we cover only certain aspects of probability. For a more complete treatment of the subject of probability, the reader is advised to consult the references.

8.1 PROBABILITY

In signal analysis, it is common to use the concepts of signal energy and power as measures of the signal under consideration. In spatial coordinates, the use of distance and squared length are commonly employed. In a somewhat related manner, the concept of probability is used as a means by which we can measure (numerically) the favorable outcomes of a given experiment. To be a valid measure, we first need to discuss the basis for the measurement.

Suppose that some experiment is considered in which the outcome does not remain constant. One of the possible outcomes of the experiment is labeled A. Tossing a coin is such an experiment, the possible outcomes being heads and tails. If the experiment is repeated N times, suppose that the outcome A will occur N_A times. The *relative frequency of occurrence* of A is N_A/N. This ratio is not very predictable when N is small. If the experiment has statistical regularity, however, the relative frequency of a particular outcome may approach a limit as the number of repetitions N becomes very large. This limiting value of the relative frequency of occurrence is called the *probability* of the outcome A and is written as $P(A)$:

$$P(A) = \lim_{N \to \infty} \frac{N_A}{N}. \tag{8.1}$$

The limit in Eq. (8.1) is not a limit in the usual functional sense, but is used to indicate that if N is very large then the expected deviation of the ratio N_A/N from a constant becomes very small. This statement is sometimes called the "empirical law of large numbers" and agrees with our intuitive ideas of probability. For example, if an honest coin is tossed a large number of times, we expect that heads will show in about one-half the tosses and that this approximation becomes better as more tosses are attempted.

EXAMPLE 8.1.1

An honest coin is tossed ten times and the outcomes of this experiment are: $\{H,T,T,H,H,H,H,T,T,T\}$. Let A be the occurrence of a head and let $x = N_A/N$. (a) Calculate x as the experiment progresses. (b) Also, as a measure of possible convergence to $x = 0.50$, calculate

$$y = \left[\sum_{i=1}^{N} (x_i - 0.5)^2/N\right]^{1/2}.$$

Solution The results of the above computation are shown below. Note the (slow) convergence as the number of trials is increased even though the given sequence is not a particularly "favorable" one.

N	H	x	y
1	1	1.000	0.500
2	1	0.500	0.354
3	1	0.333	0.304
4	2	0.500	0.264
5	3	0.600	0.240

N	H	x	y
6	4	0.667	0.229
7	5	0.714	0.227
8	5	0.625	0.217
9	5	0.556	0.206
10	5	0.500	0.195

In many cases, the experiments needed to determine the probability of an event can be done conceptually. For example, suppose that a die is rolled and we wish to calculate the probability that a "1" will appear. Assuming an honest die (so that all six outcomes are equally probable) and its symmetry, we can conclude that the probability of a one is 1/6 without ever actually running a practical experiment. This type of reasoning leads to the following alternative definition of the probability of occurrence of an outcome A:

$$P(A) = \frac{\text{number of possible outcomes favorable to the event } A}{\text{total number of possible equally likely outcomes}}. \tag{8.2}$$

Care is needed in a correct interpretation of Eq. (8.2) because the assumption that all outcomes are equally likely may be wrong and will lead to erroneous conclusions.

DRILL PROBLEM 8.1.1

What is the probability of an even integer showing on (a) a roll of one honest die? (b) the sum of two honest dice? (c) What would you expect from the sum of three honest dice?

ANSWER: (a) 1/2; (b) 1/2; (c) 1/2.

A consideration of Eqs. (8.1) and (8.2) shows that our definition of the probability of occurrence of a given outcome is a positive real number between zero and one, or

$$0 \leq P(A) \leq 1 . \tag{8.3}$$

Let us denote the fact that an outcome is an absolute certainty by S and the fact that it is not possible at all by $\varnothing$. Again, using Eqs. (8.1) and (8.2), we conclude that

$$\begin{cases} P(\varnothing) = 0 , \\ P(S) = 1 . \end{cases} \tag{8.4}$$

Two possible events are defined as being *mutually exclusive* (or disjoint) if they cannot possibly occur simultaneously: i.e., if the occurrence of one outcome precludes the occurrence of the other. In this case we can write[†]

$$P(A \text{ or } B) = P(A + B) = \lim_{N \to \infty} \frac{N_A + N_B}{N} = \lim_{N \to \infty} \frac{N_A}{N} + \lim_{N \to \infty} \frac{N_B}{N} , \tag{8.5}$$

or

$$P(A + B) = P(A) + P(B) . \tag{8.6}$$

The additivity of probabilities as expressed by Eq. (8.6) permits us to assign probabilities to elementary events and then derive the probability of other events from these. In fact, in an advanced course in probability, Eqs. (8.3), (8.4), and (8.6) can be postulated and then the theory of probability is deduced from these three postulates.

It is convenient to represent A and B by closed plane figures, known as "Venn diagrams," which can be used as visual aids in the above operations. A Venn diagram for the two mutually exclusive events A and B is shown in Fig. 8.1. Because A and B are mutually exclusive, a given event cannot be represented as lying in both areas and therefore we conclude that the two areas do not overlap if the events are mutually exclusive. Because probability is a numerical measure of the outcomes of an experiment, Eq. (8.6) follows from this Venn diagram.

If all possible outcomes are included, an extension of Eq. (8.6) yields

$$P(A_1 + A_2 + A_3 + \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots , \tag{8.7}$$

[†] In mathematics, it is common to use the set notations of "union" ($A \cup B$) and "intersection" ($A \cap B$) instead of the sum ($A + B$) and product (AB) notations, respectively.

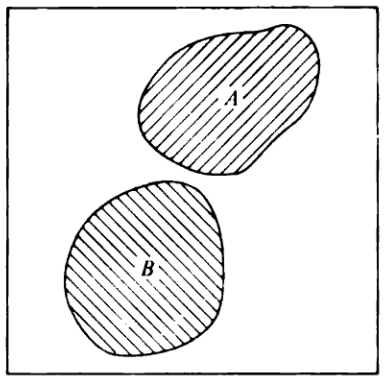


Figure 8.1 A Venn diagram of two mutually exclusive events.

or, using Eq. (8.4),

$$\sum_{i=1}^{N} P(A_i) = 1 . \tag{8.8}$$

If in addition all the outcomes have equal probability, then:

$$P(A_i) = 1/N . \tag{8.9}$$

In the latter case the outcomes are said to be *equally likely*.

EXAMPLE 8.1.2

Consider the experiment of tossing two honest coins. What is the probability of obtaining two heads or two tails?

Solution I There are four possible outcomes of the experiment: $\{HH\}, \{HT\}, \{TH\}, \{TT\}$. The probability of each outcome [cf. Eq. (8.2)] is 1/4. Let the outcome $\{HH\}$ be A and $\{TT\}$ be B. As these outcomes are mutually exclusive,

$$P(A \text{ or } B) = P(A + B) = P(A) + P(B) = 1/4 + 1/4 = 1/2 .$$

Solution II The number of favorable outcomes is 2. A direct application of Eq. (8.2) then yields a probability of $2/4 = 1/2$.

DRILL PROBLEM 8.1.2

Consider the experiment of rolling two honest dice. The events of the experiment are the sums of the numbers showing up. What is the probability that this sum is less than 5?

ANSWER: 1/6.

8.2 CONDITIONAL PROBABILITY AND STATISTICAL INDEPENDENCE

Consider the possible events A and B. These events may or may not occur together. (If they are mutually exclusive they cannot occur together.) The probability of the joint occurrence of A and B is $P(A \text{ and } B)$, or, more simply, $P(AB)$.† Suppose that we repeat an experiment N times and let N_{AB} be the number of times that A and B occur together. Using the relative-frequency approach, we find that the joint probability of their occurrence is (the limit as $N \to \infty$ is implied)

$$P(AB) = \frac{N_{AB}}{N}. \tag{8.10}$$

Recall that N_A is the number of times that A occurs and N_B is the number of times that B occurs. Because A and B may not always occur together, we have $N_{AB} \leq N_A$ and $N_{AB} \leq N_B$.

It may be that the occurrence of the event B depends in some way on the occurrence of A. The probability of B given that A is known to have occurred is called the *conditional probability* of B given A and is written as $P(B \mid A)$. Using the relative-frequency approach for the N_A trials in which A occurs, we can write

$$P(B \mid A) = \frac{N_{AB}}{N_A} = \frac{N_{AB}/N}{N_A/N} = \frac{P(AB)}{P(A)}. \tag{8.11}$$

In a similar manner,

$$P(A \mid B) = \frac{N_{AB}}{N_B} = \frac{N_{AB}/N}{N_B/N} = \frac{P(AB)}{P(B)}. \tag{8.12}$$

Combining Eqs. (8.11) and (8.12), we can express the joint probability $P(AB)$ as

$$P(AB) = P(B \mid A)P(A) = P(A \mid B)P(B), \tag{8.13}$$

or

$$P(B \mid A) = \frac{P(B)P(A \mid B)}{P(A)}. \tag{8.14}$$

Equation (8.14) is known as *Bayes' theorem* in statistical decision theory.

EXAMPLE 8.2.1

A standard deck of 52 playing cards is shuffled and two cards are drawn at random. After looking at the first card, what is the probability that the second is a heart if (a) the first card is a heart; (b) the first card is not a heart? (c) Calculate $P(AB)$ for part (a).

† This is equivalent to the set notation $P(A \cap B)$.

Solution Let A = event of a heart on the first card; B = event of a heart on the second card; C = event of no heart on the first card.

a) Because there are 13 hearts in the original deck and one is known to have been removed,

$$P(B|A) = 12/51 .$$

b) In this case, all 13 hearts are still distributed among $(52 - 1) = 51$ cards so that

$$P(B|C) = 13/51 .$$

c) $P(A) = 13/52$; using Eq. (8.11) and the result of part (a) gives us

$$P(AB) = \frac{12}{51}\frac{13}{52} = \frac{3}{51} .$$

DRILL PROBLEM 8.2.1

Two honest dice are rolled. Event A is the sum that is divisible by 2. Event B is the sum that is divisible by 4. Calculate (a) $P(AB)$; (b) $P(B|A)$.

ANSWER: (a) 1/4; (b) 1/2.

Now suppose that B is independent of A so that the occurrence of A does not in any way influence the occurrence of B. If this is true, the conditional probability $P(B|A)$ is simply the probability $P(B)$; that is,

$$P(B|A) = P(B) \quad \text{if independent.} \tag{8.15}$$

Using Eq. (8.15) in Eq. (8.13), we conclude that the two events A and B are *statistically independent* if

$$P(AB) = P(A)P(B) . \tag{8.16}$$

Therefore when the possible outcomes of an experiment are statistically independent, the joint probability of the particular outcomes is equal to the product of the probabilities of the individual outcomes. This result may be extended to any arbitrary number of outcomes.

Note that the property of statistical independence is quite different from that of mutual exclusiveness. In fact, if A and B are mutually exclusive, the probability of their joint occurrence is zero by definition—that is, $P(AB) = 0$.

EXAMPLE 8.2.2

Equation (8.6) gives the probability of two mutually exclusive events in terms of the sum of their probabilities of occurrence: $P(A + B) = P(A) + P(B)$. Derive a comparable result if the events are not mutually exclusive.

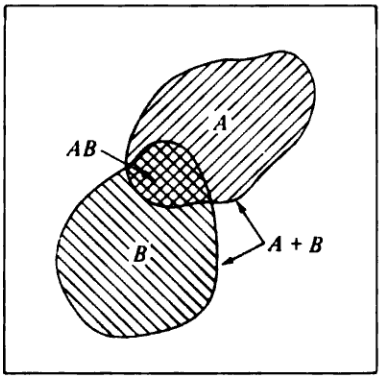


Figure 8.2 A Venn diagram of two events not mutually exclusive.

Solution Let the two outcomes be A and B; a Venn diagram is shown in Fig. 8.2. Because A and B are not mutually exclusive, the joint probability $P(AB)$ is nonzero, as represented by an overlap in the Venn diagram. (If it were zero, then A and B would be nonoverlapping, or disjoint.)

From a consideration of the Venn diagram, it is quite obvious that simply adding A and B gives the correct result (A or B) only when A and B are disjoint (i.e., mutually exclusive). In the more general case, an addition of A and B includes (A and B) twice and to correct for this we write

$$P(A + B) = P(A) + P(B) - P(AB).$$

This is our desired result and is a more general version of Eq. (8.6). Note that when A and B are mutually exclusive, $P(AB) = 0$ and this result simplifies to that given in Eq. (8.6).

DRILL PROBLEM 8.2.2

A number of honest dice are rolled. Calculate the probability of every die rolled showing a "1" if the number of dice rolled is (a) 2; (b) 3; (c) N.

ANSWER: (a) 1/36; (b) 1/216; (c) $(1/6)^N$.

8.3 THE RANDOM VARIABLE AND CUMULATIVE DISTRIBUTION FUNCTION

The concept of probability was introduced in the previous sections to measure numerically the possible outcomes of an experiment. The possible outcomes themselves are not necessarily numerical, however (e.g., heads and tails of a coin), and it becomes cumbersome trying to describe them. The concept of a random variable is

introduced to signify a rule (or mapping) by means of which a real number is assigned to each of the possible outcomes of an experiment. A combination of these two concepts will enable us to make numerical graphs of probability versus events for a given experiment.

Let the events of an experiment having a finite number of possible outcomes be identified by the symbols λ_i. These symbols need not represent numbers but only the possible outcomes of the experiment. We establish some rule or functional relationship (mapping), $X(\lambda_i)$, which assigns a real number to each possible outcome. This rule is called a *random variable*.

As an example, we could assign the number -1 to "tails" and the number $+1$ to "heads" in the toss-of-a-coin experiment. Alternatively, we could just as well have associated $\sqrt{2}$ with "tails" and π with "heads." In either case, we have accomplished our purpose in assigning a number to each possible outcome. Note that in defining the random variable we have not assigned a measure to each outcome (i.e., whether a head occurs on the average in one-half of the tosses). The roles of probability and random variable are distinctly different.

Sometimes the identifying symbols λ_i turn out to be numerical (e.g., the six faces of a die—they are already assigned numbers in an arbitrary sequence). If we wish, we could assign new numbers to these events, e.g., squaring each number. Thus a random variable itself is not really random but is an arbitrarily established rule to assign numbers to events.

Although we have introduced the concept for a finite number of possible outcomes, a random variable may be discrete or continuous depending on the nature of the possible outcomes of the experiment. For example, the numbers assigned to the tossing of a coin (for a finite number of tosses) are discrete because the possible outcomes form a finite countable set. In contrast, the angular position of a pointer on a rotating wheel is a random variable whose values form a continuum over the range $0 < \theta \leq 360°$. To allow for both the discrete and continuous cases, we shall drop the subscript i and simply write $X(\lambda)$, realizing that we must supply the subscript when we are referring to the discrete case.

It is convenient and, in fact, quite advantageous to plot the outcome probabilities for a given experiment versus the random variable assigned. In order to make such a plot meaningful, we shall first define what we mean by the cumulative distribution function.

The *cumulative distribution function*, $F_X(x)$, associated with a random variable X is defined as[†]

$$F_X(x) \triangleq P\{X(\lambda) \leq x\}. \tag{8.17}$$

The cumulative distribution function is dependent on both the random variable X and the value of the argument x. It is for this reason that the subscript X is included in the notation.

[†] It is conventional to use a capital letter to denote a random variable and the corresponding lower-case letter to denote the value within the range.

Because the cumulative distribution function is based directly on the concept of probability, it has the following properties:

$$0 \le F_X(x) \le 1\,, \tag{8.18}$$

$$F_X(x_1) \le F_X(x_2), \quad \text{if } x_1 < x_2\,, \tag{8.19}$$

$$\begin{cases} F_X(-\infty) = 0\,, \\ F_X(+\infty) = 1\,. \end{cases} \tag{8.20}$$

The property in Eq. (8.18) follows from the fact that $F_X(x)$ is a probability and therefore must obey Eq. (8.3). The property in Eq. (8.19) holds also as a result of Eq. (8.3) because probability is always nonnegative; for $x_1 < x_2$, $F_X(x_2)$ includes as many or more of the possible outcomes than are included in $F_X(x_1)$. The property in Eq. (8.20) follows from Eq. (8.4) because $F_X(-\infty)$ includes no events possible while $F_X(+\infty)$ includes all events possible.

For a discrete random variable $X(\lambda_i)$ with associated probabilities P_i, the cumulative distribution function can be expressed as

$$F_X(x) = \sum_i P_i u(x - x_i)\,. \tag{8.21}$$

Therefore the cumulative distribution function of a discrete random variable has a series of finite ("stepwise") discontinuities occurring at the points $x = x_i$. The height of each step is given by: $P_i = P(X = x_i)$ and between the steps $F_X(x)$ is constant. At a point of finite discontinuity, $F_X(x)$ is taken to be continuous to the right.

EXAMPLE 8.3.1

Consider the experiment that consists in the tossing of three honest coins. The random variable chosen is defined by assigning 0 to a "tail" and 1 to a "head," and then adding the numbers. Determine and plot the cumulative distribution function.

Solution The eight possible outcomes are

000	100
001	101
010	110
011	111

If we assume equally likely outcomes, the probability of each is 1/8. The outcomes are mutually exclusive so the probabilities are additive. A table of the random variable $X(\lambda_i)$ and the associated probabilities is shown below.

x	P
0	1/8
1	3/8
2	3/8
3	1/8

The cumulative distribution function is found by adding the probabilities for all outcomes below a given number and is plotted in Fig. 8.3(a). For example, the cumulative distribution function evaluated at $x = 2$ is

$$\begin{aligned} F_X(2) &= P\{X \le 2\} \\ &= P(0) + P(1) + P(2) \\ &= 1/8 + 3/8 + 3/8 = 7/8 . \end{aligned}$$

In equation form, we can write our answer as [cf. Eq. (8.21)]

$$F_X(x) = \tfrac{1}{8}u(x) + \tfrac{3}{8}u(x-1) + \tfrac{3}{8}u(x-2) + \tfrac{1}{8}u(x-3) .$$

Note that $F_X(x)$ satisfies all of the properties given by Eqs. (8.18)–(8.20).

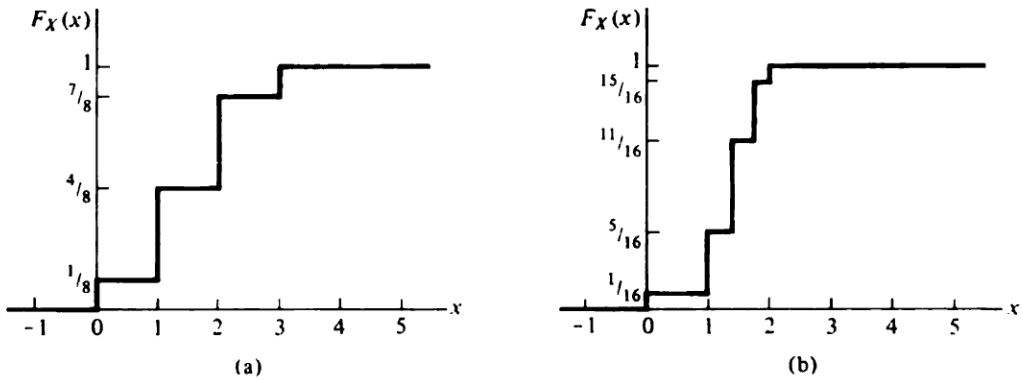


Figure 8.3 Cumulative distribution function for the discrete random variable of (a) Example 8.3.1 and (b) Drill Problem 8.3.1.

DRILL PROBLEM 8.3.1

Four honest coins are tossed and X is designated as the square root of the number of heads that turn up. Determine and plot the cumulative distribution function, $F_X(x)$.

ANSWER: The result is plotted in Fig. 8.3(b); the equation is

$$F_X(x) = \tfrac{1}{16}u(x) + \tfrac{4}{16}u(x-1) + \tfrac{6}{16}u(x-\sqrt{2}) + \tfrac{4}{16}u(x-\sqrt{3}) + \tfrac{1}{16}u(x-2) .$$

8.4 THE PROBABILITY DENSITY FUNCTION

Instead of tossing coins or rolling dice, suppose we spin a wheel with an attached pointer. As a result of frictional losses, the wheel gradually slows and comes to rest at some angular position θ. Suppose that we mark off 360 equal divisions around the

circumference. The stopping point of the wheel is the desired outcome of the experiment. There is a uncountably infinite number of possible outcomes and so we assign a continuous random variable θ to the experiment.

Consider the probability of stopping at points around the rim of the wheel. Because the wheel must stop somewhere within the range, the total probability of $0 < \theta \leq 360°$ is one. Because we have divided this range uniformly, the probability of $0 < \theta \leq 1°$ is $1/360$. What, then, is the probability that the wheel comes to rest at a given point? There is an uncountably infinite number of points within each division (no matter how small we make the divisions), and the only reasonable answer is that this probability must be 0; otherwise we cannot possibly end up with a total probability of 1.

Even though $P\{\theta = \theta_1\} = 0$, it makes good sense to speak of $P\{\theta \leq \theta_1\}$, or $P\{\theta_1 < \theta \leq \theta_2\}$ for $\theta_1 < \theta_2$. The concept of the cumulative distribution function is still valid, but it is more convenient in such cases to define a function whose *area* is the probability of occurrence within a specified range. Because area is being equated to probability, such a function is termed a *probability density function* (pdf).

Therefore we define the probability density function, $p_X(x)$, of a random variable X by

$$\int_{-\infty}^{x} p_X(x)\,dx = P\{X \leq x\}. \tag{8.22}$$

However, we can also write [cf. Eq. (8.17)]

$$P\{X \leq x\} = F_X(x), \tag{8.23}$$

so that

$$\int_{-\infty}^{x} p_X(x)\,dx = F_X(x). \tag{8.24}$$

Taking the derivative of each side of Eq. (8.24), we find that an equivalent definition of the pdf is

$$p_X(x) = \frac{d}{dx}F_X(x). \tag{8.25}$$

Thus the pdf of a random variable X is the derivative of the cumulative distribution function, $F_X(x)$.

It may happen that the derivative of the cumulative distribution function will not exist at all points. In this case we must resort to a limiting operation. The delta function notation used in Chapters 2 and 3 becomes useful in many cases.

The random variable subscript notation begins to get cumbersome at this point. When there is only a single random variable under discussion and no ambiguity can result, we shall omit the subscript in the probability density function for simplicity of notation.

From Eq. (8.24) and the properties of the cumulative distribution function [cf. Eqs. (8.18)–(8.20)], some properties of the probability density function are

$$\int_{-\infty}^{\infty} p(x)\,dx = F_X(\infty) = 1\,, \tag{8.26}$$

$$\int_{x_1}^{x_2} p(x)\,dx = F_X(x_2) - F_X(x_1) = P\{x_1 < X \le x_2\}\,. \tag{8.27}$$

Because Eq. (8.27) must always be nonnegative for any (x_1, x_2), we can also write

$$p(x) \ge 0 \quad \text{for all } x\,. \tag{8.28}$$

EXAMPLE 8.4.1

A wheel with a pointer is rotated and allowed to slow down and stop. When the wheel comes to rest, the angular position of the pointer is θ. A scale, $0 < \theta \le 360°$, is marked with uniform increments around the rim of the wheel. Determine the pdf of θ.

Solution Because all values in the range $0 < \theta \le 360°$ are equally probable, the pdf is a constant over this range so that

$$p(\theta) = \begin{cases} K & 0 < \theta \le 360°, \\ 0 & \text{elsewhere.} \end{cases}$$

The constant K can be evaluated because the total probability of the wheel stopping within the range is known to be 1:

$$\int_{-\infty}^{\infty} p(\theta)\,d\theta = \int_0^{360} K\,d\theta = 360K = 1\,,$$

$$K = 1/360\,.$$

Note that the probability of the pointer stopping exactly at the point $\theta = 1°$ is, formally,

$$P(1) = \lim_{\epsilon \to 0} \int_{1-\epsilon}^{1+\epsilon} K\,d\theta = 0\,,$$

but the probability of the pointer stopping within the interval $0 < \theta \le 1°$ is

$$\int_0^1 \frac{1}{360}\,d\theta = \frac{1}{360}\,.$$

A graph of the pdf is shown in Fig. 8.4.

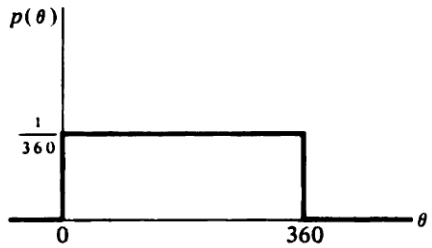


Figure 8.4 The probability density function of Example 8.4.1.

DRILL PROBLEM 8.4.1

In the reliability testing of small electronic component parts, the probability of the time-to-failure is often approximated using the exponential pdf, $p(y) = K \exp(-y/\alpha)u(y)$, where y is the equipment operating time in hours and the constant α is experimentally determined (and is called the *mean-time-before-failure*). (a) Compute the probability that a given part described by this pdf will fail within the first α hours of operation. (b) If $\alpha = 10{,}000$ hours, what is the probability that the part will fail during the first hour of operation?

ANSWER: (a) 0.632; (b) 0.0001.

It is convenient to extend our definition of probability density function to include the case of discrete random variables. This follows readily if we recall that the cumulative distribution function is composed of a series of step functions and that the derivative of a step function, at least formally, is an impulse function. Using Eq. (8.21), we can write

$$p(x) = \frac{d}{dx}\left[\sum_i P_i u(x - x_i)\right],$$

or

$$p(x) = \sum_i P_i \delta(x - x_i). \tag{8.29}$$

Although our examples tend to fall into either of the two categories, there is no theoretical reason for a pdf to be purely discrete or purely continuous.

EXAMPLE 8.4.2

Consider the experiment that consists in the rolling of two honest dice. The random variable X is assigned to the sum of the numbers showing up on the two dice. Determine and plot the cumulative distribution function and the pdf of X.

Solution There are 36 possible outcomes because each die may show any of the numbers 1–6; the range of the sum then goes from 2–12. Each possible outcome is assumed equally likely and so its probability is equal to 1/36. The events are mutually exclusive so additivity applies. We find that $P(2) = P(12) = 1/36$, $P(3) = P(11) = 2/36$, $P(4) = P(10) = 3/36$, $P(5) = P(9) = 4/36$, $P(6) = P(8) = 5/36$, $P(7) = 6/36$. The resulting cumulative distribution function and pdf are shown in Fig. 8.5.

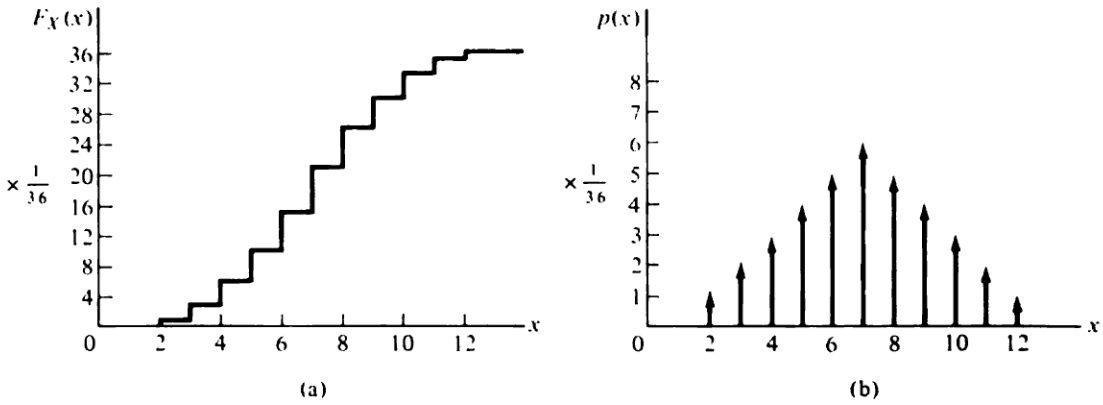


Figure 8.5 (a) The cumulative distribution function and (b) the probability density function for Example 8.4.2.

DRILL PROBLEM 8.4.2

Four honest coins are tossed and X is designated as the number of heads that turn up. Determine the pdf of X.

ANSWER:† $p(x) = \frac{1}{16}\delta(x) + \frac{4}{16}\delta(x-1) + \frac{6}{16}\delta(x-2) + \frac{4}{16}\delta(x-3) + \frac{1}{16}\delta(x-4)$.

A comparison with Chapter 4 (cf. Section 4.2) reveals that our development of the cumulative power spectrum and power spectral density is practically identical to the above development of the cumulative distribution function and the probability density function. This analogy can be extended to the mathematical role of the periodic signal in power spectral density and the discrete random variable in the probability density function, since both give rise to series of impulse functions. Both density functions are nonnegative; both cumulative functions are therefore monotone nondecreasing with increasing argument. Each point in the power spectral density has zero power associated with it while the area under the density curve gives the power within a specified frequency range. Each point in the probability density has zero probability associated with it, while the area under the density curve gives the probability within a specified range of the random variable.

† A more compact notation for problems of this type is discussed in Section 8.6.2.

8.5 STATISTICAL AVERAGES

The probability density function is useful in evaluating the most probable values of desired quantities. A way to find the most probable value of X is to locate that point in the pdf of X which has the maximum value. This point is called the *mode*. A necessary condition for the determination of the mode (in the continuous case) is that $(d/dx)p(x) = 0$. The location of this point is extremely sensitive to minor changes in the pdf, however. Another measurement is to locate the point that divides the area into two equal parts. This is called the *median*. The determination of the median is improved over that of the mode but is still quite sensitive to minor changes in the pdf. Alternatively, we can use statistical averaging.

The statistical average of a random variable X (or a function of a random variable) is the numerical average of the values which X (or a function of X) can assume, weighted by their probabilities. If $X = x_1$ is observed n_1 times, $X = x_2$ is observed n_2 times, etc., until, finally, $X = x_k$ is observed n_k times, then $n_1 + n_2 + \ldots + n_k = N$ and the observed value is

$$n_1x_1 + n_2x_2 + n_3x_3 + \cdots = \sum_{i=1}^{K} n_ix_i. \tag{8.30}$$

Dividing by N and then letting $N \to \infty$ to employ our relative-frequency approach to probability, we have

$$M_1 = \lim_{N\to\infty} \sum_{i=1}^{K} x_i\frac{n_i}{N} = \sum_{i=1}^{K} x_iP_i = m_X, \tag{8.31}$$

where M_1 is called the first moment of X and m_X is the *average value or mean value of* X. Equation (8.31) is also called the *expected value of* X. Note that for the special case in which all events are equally likely [cf. Eq. (8.9)], Eq. (8.31) simplifies to the arithmetic average of X:

$$m_X = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{8.32}$$

In a similar manner the second moment M_2 is equal to the expected value of X^2:

$$M_2 = \sum_{i=1}^{K} x_i^2P_i, \tag{8.33}$$

and is called the *mean-square* value of X. The square root of M_2 is called the *root-mean-square* (rms) value of X. Higher-ordered moments can be expressed in a similar way.

For the case of a continuous random variable, the summation in Eqs. (8.31) and (8.33) approaches an integration over the entire range of X so that the expected value of X, designated by $E\{X\}$, is

$$M_1 = E\{X\} = \int_{-\infty}^{\infty} xp(x)\,dx = m_X, \tag{8.34}$$

and the expected value of X^2, designated by $E\{X^2\}$, is

$$M_2 = E\{X^2\} = \int_{-\infty}^{\infty} x^2 p(x)\,dx. \tag{8.35}$$

By extension, the expected value of X^n is

$$E\{X^n\} = \int_{-\infty}^{\infty} x^n p(x)\,dx. \tag{8.36}$$

The expected value $E\{X^n\}$ is referred to as the "nth moment" of the random variable X. This is analogous to the concept of a moment in mechanics. If $p(x)$ were the mass density (normalized to 1), the first moment [cf. Eq. (8.34)] corresponds to the center of gravity, the second [cf. Eq. (8.35)] to the moment of inertia about the origin, etc.

A knowledge of all the moments of a random variable, provided they are finite, specifies that random variable as completely as a knowledge of the pdf. Fortunately, we usually can get by with working only with the first two moments.

Often it is convenient to subtract the first moment $M_1 = m_X$ before computing the other moments. This is analogous to mechanics, where moments are referred to the center of gravity rather than the origin of the coordinate system. Such moments are referred to as "central moments." The second central moment is of particular importance and is called the *variance* of the random variable X. Thus the variance, σ_X^2, of the random variable X is defined as

$$\sigma_X^2 = E\{(X - m_X)^2\} = \int_{-\infty}^{\infty} (x - m_X)^2 p(x)\,dx. \tag{8.37}$$

The square root of the variance is called the *standard deviation*, σ_X, of the random variable X. The standard deviation has the same units as X and is a convenient indicator of the effective width of the probability density function. For example, if two random variables, X, Y, both have average values of 1 but $\sigma_X = 1$ and $\sigma_Y = 2$, we can conclude that the effective width of the pdf for Y is more distributed (spread out) than that for X. Stated differently, the probability of X being close to 1 is greater than the probability that Y is close to 1.

Finally, the expected value of a function of a random variable $g(X)$ is defined as

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x)p(x)\,dx. \tag{8.38}$$

When $g(X) = X^n$, the expected value is the nth moment of X, as described above. If $g(X)$ is equal to a constant, then Eq. (8.38) gives that constant because $\int_{-\infty}^{\infty} p(x)\,dx = 1$. The expected value operator $E\{\cdot\}$ is linear so that

$$E\{a_1 X + a_2\} = a_1 m_X + a_2.$$

EXAMPLE 8.5.1

Show that an equivalent expression to Eq. (8.37) is

$$\sigma_X^2 = E\{X^2\} - [E\{X\}]^2 = M_2 - M_1^2.$$

Solution Expanding the squared term and integrating term by term, we can rewrite Eq. (8.37) as

$$\begin{aligned}\sigma_X^2 &= E\{X^2 - 2M_1X + M_1^2\}\\ &= E\{X^2\} - 2M_1E\{X\} + M_1^2\\ &= M_2 - 2M_1^2 + M_1^2\\ &= M_2 - M_1^2.\end{aligned}$$

Note that if $M_1 = 0$, then $\sigma_X^2 = M_2 = E\{X^2\}$.

DRILL PROBLEM 8.5.1

Compute the mean m_X and the standard deviation σ_X of the random variable X described by the following probability density functions: (a) $p(x) = \delta(x - 1)$; (b) $p(x) = (8/9)x[u(x) - u(x - 3/2)]$; (c) $p(x) = \exp(-x)u(x)$.

ANSWER: (a) 1, 0; (b) 1, $1/(2\sqrt{2})$; (c) 1, 1.

★ **EXAMPLE 8.5.2**

The foregoing remarks about the variance of a random variable as a relative indicator of the spread of the probability density function can be supported by a very general result from statistics known as *Chebyshev's inequality:*

$$P(|X - m_X| > k\sigma_X) < \frac{1}{k^2} \qquad \text{for } k > 0.$$

This inequality states that the probability of X being outside $\pm k$ standard deviations of the mean is less than $1/k^2$ *regardless of* $p(x)$. Therefore the probability of finding X appreciably different from the mean is greater for a large standard deviation than for a small one. Because the total probability is unity, Chebyshev's inequality may also be written as

$$P(|X - m_X| \le k\sigma_X) \ge 1 - \frac{1}{k^2}.$$

A proof of Chebyshev's inequality may be found in texts on probability and statistics.[†]

As an example, compute $P(X > 5)$ for the pdf of Drill Problem 8.5.1(c), and compare with the bound given by the Chebyshev inequality.

Solution Using a knowledge of the pdf, we have

$$P(X > 5) = \int_5^\infty e^{-x}\,dx = e^{-5} = 0.0067.$$

[†] Chebyshev is spelled Tchebycheff in some texts.

Using Chebyshev's inequality with $m_X = 1$, $\sigma_X = 1$, we have

$$P(X > 5) = P(|X - 1| > 4) < \frac{1}{4^2} = 0.0625 .$$

Although this bound is almost an order of magnitude larger than the exact value in this case, note that the computation of this bound requires no knowledge of the pdf (only the first two moments) and hence is a very general result.

Although the equations for expected values have been written in terms of continuous random variables, they also hold for the case of discrete random variables when the impulse functions arising in the probability densities are handled correctly. Recall that the probability density for a discrete random variable can be written as [cf. Eq. (8.29)] $p(x) = \Sigma_i P_i \delta(x - x_i)$ so that the expected value of X for the discrete case is, from Eq. (8.34),

$$\begin{aligned} m_X = E\{X\} &= \int_{-\infty}^{\infty} x \sum_{i=1}^{K} P_i \delta(x - x_i)\, dx \\ &= \sum_{i=1}^{K} x_i P_i , \end{aligned} \tag{8.39}$$

and is in agreement with Eq. (8.31). In a similar manner,

$$E\{X^2\} = \int_{-\infty}^{\infty} x^2 \sum_{i=1}^{K} P_i \delta(x - x_i)\, dx = \sum_{i=1}^{K} x_i^2 P_i , \tag{8.40}$$

and the variance of X is

$$\sigma_X^2 = \sum_{i=1}^{K} (x_i - m_X)^2 P_i . \tag{8.41}$$

For the equally likely case, Eq. (8.40) simplifies to

$$E\{X^2\} = \frac{1}{N} \sum_{i=1}^{N} x_i^2 , \tag{8.42}$$

and Eq. (8.41) becomes

$$\sigma_X^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - m_X)^2 . \tag{8.43}$$

EXAMPLE 8.5.3

An exam is taken by nine students. After grading, it is found that the percentage scores are as listed below. Compute the median and average grade (treat all students equally!). Also compute the standard deviation.

N	Z	N	Z
1	33	6	82
2	47	7	88
3	58	8	94
4	67	9	100
5	75		

Solution The student number that divides the distribution equally is No. 5 so the median grade is 75. Because all students are to be treated equally, $P_i = 1/9$ for all i. The average grade is [Eq. (8.32)]

$$E\{Z\} = \sum_{i=1}^{9} z_i P_i = 644/9 = 71.6 .$$

Using Eq. (8.43), we find the standard deviation is given by

$$\sigma_Z = \sqrt{\sum_{i=1}^{9} (z_i - 71.6)^2/9} = 21.1 .$$

DRILL PROBLEM 8.5.2

Compute the mean and the standard deviation of the random variable X in Example 8.4.2.

ANSWER: 7, 2.42.

The number of possible trials in an experiment with a discrete number of states is called a *population*. A *sample* is a finite number of trials that is less than the population. In statistical averaging, we make a distinction between a population average and a sample average. In the former, all possible trials are available to us for use in the statistical averaging, whereas in the latter we must estimate the statistical averages of the population based only on a finite number of samples.

Suppose that we wish to estimate the population average θ based on the N samples from the population which are available to us. The estimate of θ will be denoted by $\hat{\theta}$. We look for certain desirable qualities in our estimate, two of which we mention here.

First, we would like to have the estimate correct if the sample size were expanded to the entire population. Thus an estimator $\hat{\theta}$ is called an *unbiased estimator* of θ if

$E\{\hat{\theta}\} = \theta$. A second desirable quality is that the estimate converge to the correct value as the number of samples becomes large. An estimator $\hat{\theta}$ is called a *consistent estimator* of θ (in a mean-square sense) if

$$\lim_{N\to\infty} E\{(\hat{\theta} - \theta)^2\} = 0 .$$

It can be shown (cf. Example 8.5.4) that the sample mean [Eq. (8.32)] is an unbiased and consistent estimator of the population mean. The sample variance [Eq. (8.43)] is a consistent estimator; however,†

$$E\{\hat{\sigma}^2\} = \frac{N-1}{N}\sigma_X^2 , \tag{8.44}$$

so that it is not an unbiased estimator. For this reason, the unbiased sample-variance estimator,

$$\sigma_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - m_X)^2 , \tag{8.45}$$

is usually used. For large N, the difference between the two becomes negligible.

The concepts of estimation are not confined to discrete random variables. They can readily be extended to the case of continuous random variables.

EXAMPLE 8.5.4

Show that the sample mean [Eq. (8.32)] is (a) an unbiased and (b) a consistent estimator of the population mean.

Solution

a)
$$E\{\hat{m}_X\} = E\left\{\frac{1}{N}\sum_{i=1}^{N} x_i\right\} = \frac{1}{N}\sum_{i=1}^{N} E\{x_i\} = m_X .$$

b)
$$E\{(\hat{m}_X - m_X)^2\} = E\left\{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - m_X)\right]^2\right\} = \frac{1}{N^2}E\left\{\sum_i\sum_k (x_i - m_X)(x_k - m_X)\right\} = \frac{1}{N^2}\sum_i\sum_k E\{x_i x_k - m_X^2\} .$$

† See, for example, J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, Second ed. New York: Wiley-Interscience, 1986, Chapter 4.

If we assume that the samples x_i, x_k are statistically independent, only those terms for $i = k$ are nonzero, and we have (using the result of Example 8.5.1),

$$E\{(\hat{m}_X - m_X)^2\} = \frac{1}{N^2}\sum_{i=1}^{N}[E\{x_i^2\} - m_X^2] = \frac{1}{N}\sigma_X^2.$$

Therefore the variance of the sample mean is equal to the population variance divided by the sample size. Because this approaches zero as $N \to \infty$, the sample mean is a consistent estimator of the population mean.

DRILL PROBLEM 8.5.3

Five resistors are selected at random from a box of 1-kΩ resistors. The measured values (in ohms) are 980, 1009, 1026, 963, and 997. (a) Estimate the mean and the standard deviation of the resistance values of the resistors in the box based on this sample. (b) Calculate the mean and standard deviation if these five were the only resistors in the box.

ANSWER: (a) 995.0 Ω, 24.55 Ω; (b) 995.0 Ω, 21.95 Ω.

8.6 SOME PROBABILITY DISTRIBUTIONS

A few of the numerous examples of probability distributions, both discrete and continuous, which arise in physical problems and are the most useful to us in this book are described here.

8.6.1 The Uniform Distribution

A random variable that is equally likely to take on any value within a given range is said to be *uniformly distributed*. The random variable may be discrete or continuous. In the former case, the pdf is a series of equally weighted impulse functions. In the latter case, the pdf is a rectangular function, as illustrated in Fig. 8.6(a). Note that

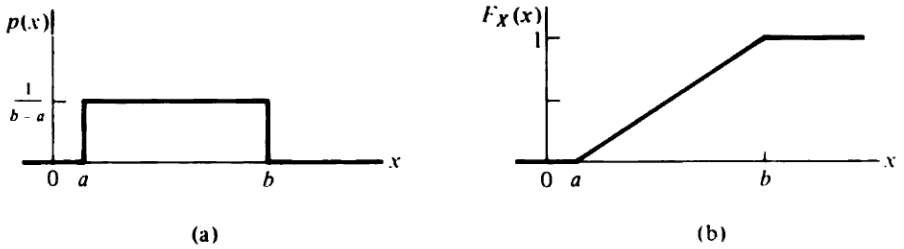


Figure 8.6 The uniform probability density function and corresponding cumulative distribution function.

the height of the pdf must be chosen to give unit area. The corresponding cumulative distribution function is shown in Fig. 8.6(b).

EXAMPLE 8.6.1

Consider the random variable $X = \cos(\omega_0 t_1 + \theta)$, where θ is a random variable with a uniform probability density over $-\pi < \theta \leq \pi$. Compute the mean and rms value.

Solution

$$E\{X\} = \int_{-\pi}^{\pi} \cos(\omega_0 t_1 + \theta)(1/2\pi)\,d\theta = 0,$$

$$E\{X^2\} = \int_{-\pi}^{\pi} \cos^2(\omega_0 t_1 + \theta)(1/2\pi)\,d\theta$$

$$= \int_{-\pi}^{\pi} \frac{1}{4\pi}\,d\theta + \int_{-\pi}^{\pi} \frac{1}{4\pi}\cos(2\omega_0 t_1 + \theta)\,d\theta = \frac{1}{2} + 0,$$

$$E\{X^2\}^{1/2} = 1/\sqrt{2} = 0.707.$$

That these results are exactly the same results as we obtain using time averages is not accidental. We shall return to the topic of statistical averages and time averages in a later section.

DRILL PROBLEM 8.6.1

Determine the mean and standard deviation of the uniform pdf shown in Fig. 8.6(a).

ANSWER: $(b + a)/2$; $(b - a)/2\sqrt{3}$.

8.6.2 The Binomial Distribution

Consider an experiment having only two possible outcomes, A and B, which are mutually exclusive. Let the probabilities be $P(A) = p$ and $P(B) = 1 - p = q$. The experiment is repeated n times and the probability of A occurring i times is

$$P_i = \binom{n}{i} p^i q^{n-i}, \tag{8.46}$$

where $\binom{n}{i}$ is the binomial coefficient,

$$\binom{n}{i} = \frac{n!}{i!\,(n-i)!}. \tag{8.47}$$

The binomial coefficient enters here to account for the total number of possible ways to combine n items taken i at a time, with all possible permutations permitted. We let $0! = 1$ by definition. The derivation of Eq. (8.46) can be found in a text on probability.

When n, i are not too large, the coefficients can also be read from Pascal's triangle. Pascal's triangle contains the coefficients of the expansion of $(p + q)^n$ and is shown in Fig. 8.7 for $1 \leq n \leq 5$. The triangle is constructed by noting that each row begins and ends with the number 1. The other numbers in a given row are determined by adding those two numbers in the row immediately above the desired number. For example, the numbers in the fourth row (that is, $n = 4$) are $1, 4 = 1 + 3, 6 = 3 + 3, 4 = 3 + 1, 1$; these numbers are the coefficients in the expansion of $(p + q)^4$. Note that each coefficient can also be expressed as $\binom{n}{i}$ where i is the power of one of the variables in the expansion of $(p + q)^n$. The binomial coefficients are useful because they enable us to find the number of possible ways to arrange n items taken i at a time.

n											
0						1					
1					1		1				
2				1		2		1			
3			1		3		3		1		
4		1		4		6		4		1	
5	1		5		10		10		5		1

Figure 8.7 Pascal's triangle.

The binomial probability density function can be obtained by substituting Eq. (8.46) into Eq. (8.29):

$$p(x) = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \delta(x - i), \tag{8.48}$$

where i is the integer number of times that event A occurs. The mean value of the binomial distribution is (np) and the variance is (npq).[†]

EXAMPLE 8.6.2

A standard deck of 52 playing cards is shuffled and two cards are drawn at random. What is the joint probability of drawing two hearts? (Cf. Example 8.2.1.)

Solution We shall use the relative-frequency concept of probability [see Eq. (8.2)]. First we find the number of possible favorable outcomes and then divide by the total number of possible equally likely outcomes.

[†] The random variable x_i may be substituted for the integers i in Eq. (8.48) if the mean and variance are appropriately modified.

The favorable outcomes are the drawing of two hearts. There are 13 hearts in the deck; taken two at a time, the number of favorable outcomes is, using Eq. (8.47),

$$\binom{13}{2} = \frac{13!}{2!\,11!} = \frac{12 \cdot 13}{2} = 78 .$$

The total number of possible equally likely outcomes is found by taking any two cards at a time out of a deck of 52 cards, or

$$\binom{52}{2} = \frac{52!}{2!\,50!} = \frac{51 \cdot 52}{2} = 1326 .$$

The joint probability of drawing two hearts out of the deck is then

$$P(AB) = \frac{78}{1326} = \frac{3}{51} .$$

This agrees with the answer obtained in Example 8.2.1(c).

EXAMPLE 8.6.3

Three honest dice are rolled. Write the pdf for the number of ones showing up.

Solution Let X be the number of ones; then $p = \frac{1}{6}$, $q = \frac{5}{6}$, $n = 3$, and Eq. (8.48) gives

$$p(x) = \sum_{i=0}^{3} \binom{3}{i} (\tfrac{1}{6})^i (\tfrac{5}{6})^{3-i} \delta(x - i) ,$$

$$p(x) = \tfrac{125}{216}\delta(x) + \tfrac{75}{216}\delta(x - 1) + \tfrac{15}{216}\delta(x - 2) + \tfrac{1}{216}\delta(x - 3) .$$

DRILL PROBLEM 8.6.2

A number is selected at random between 0 and 1. (Include 0 but not 1.) What is the probability that in the decimal expansion of the selected number the sum of the first two digits is 5 or less?

ANSWER: 6!/100.

DRILL PROBLEM 8.6.3

Three honest coins are tossed and X is designated as the number of heads that turn up. Determine the pdf of X, using Eq. (8.48); repeat for the case of six coins.

ANSWER: See Fig. 8.8.

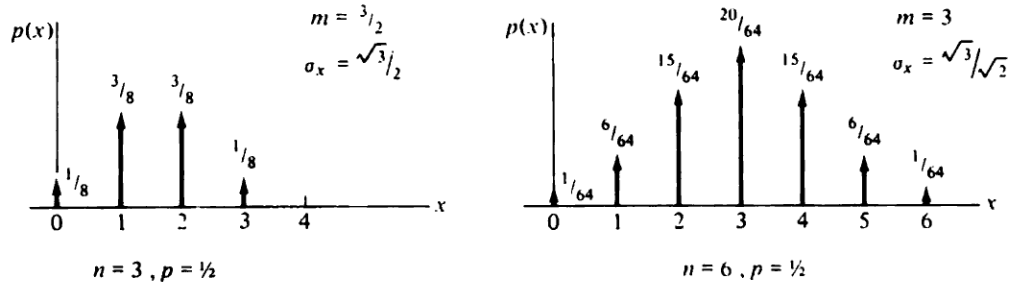


Figure 8.8 The binomial probability density function for two different values of n.

8.6.3 The Poisson Distribution

In certain types of problems we may wish to count the number of occurrences of some event during various intervals of time. A special case of such a counting process, if the average rate of occurrence is known, is that described by the Poisson distribution,

$$P_i = e^{-\nu}\frac{\nu^i}{i!} \tag{8.49}$$

where ν is the average value of the rate of occurrence of the desired event. The mean value of the Poisson distribution is ν and the variance is ν.

The Poisson distribution arises in some physical problems involving counting, e.g., counting the particles emitted over a time duration in radioactive decay, the random number of photoelectrons generated by a photodetector, and monitoring the number of telephone calls arriving at a switching center. Our interest here arises from the fact that use of the Poisson distribution is pertinent to the problem of the transmission of many data bits when the error rates are low. The binomial distribution becomes awkward to handle in such cases. However, if the mean value of the error rate remains finite (and equal to ν) we can approximate the binomial distribution by the Poisson distribution. This is illustrated in the following examples.

EXAMPLE 8.6.4

Two honest dice are rolled and the desired outcome is the appearance of all 1's. (a) Calculate the probability that this will occur twice in forty rolls, first using the binomial distribution and then approximating with the Poisson distribution. (b) Repeat for the case of four dice and 1000 rolls.

Solution a) Using the binomial distribution, we have

$$P_2 = \binom{40}{2}\left[\frac{1}{36}\right]^2\left[\frac{35}{36}\right]^{38} = \frac{40!}{2!\,38!}\frac{35^{38}}{36^{40}} = 0.2063\,,$$

and using the Poisson distribution, we get

$$\nu = np = 40/36\,,$$

$$P_2 \approx e^{-\nu}(\nu)^2/2! = 0.2032\,.$$

The approximation becomes better with larger n for a given ν. Note the comparatively simple calculation when contrasted with the binomial distribution; this is emphasized in part (b).

b) Using the binomial distribution, we have

$$P_2 = \binom{1000}{2}\left[\frac{1}{1296}\right]^2\left[\frac{1295}{1296}\right]^{998} = 0.13765\,,$$

and using the Poisson distribution, we get

$$\nu = np = 1000/1296\,,$$

$$P_2 \approx e^{-\nu}(\nu)^2/2! = 0.13761\,.$$

DRILL PROBLEM 8.6.4

A certain communication system transmits digital data with an average error rate of one digit in 10^6 digits. What is the probability that three errors will occur in a given transmission of 10^6 digits?

ANSWER: 6.13%.

8.6.4 The Gaussian Distribution

A widely known probability density function is the gaussian or "normal" density function. It arises when a large number of independent factors (within some fairly broad restrictions) contribute additively to an end result. This result, known as a "central-limit theorem," states that the sum of N independent random variables approaches a gaussian distribution as N becomes large. This result is not dependent upon the distribution of each random variable (within broad restrictions) as long as the contribution of each is small compared to the sum. This theorem is not easy to prove and interpret correctly, however, and we shall not pursue the topic further except to observe that electrical noise often results from summations of a large number of random effects and therefore tends to be gaussian-distributed.

The *gaussian pdf* is continuous and is defined by

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma}e^{-(x-m)^2/2\sigma^2}, \tag{8.50}$$

where m, σ^2 are the mean and the variance respectively. The factor: $1/(\sqrt{2\pi}\,\sigma)$ is needed so that there is unit area. A graph of the gaussian pdf is shown in Fig. 8.9. Note that the gaussian pdf is uniquely specified by its first two moments. For convenience in tabulating numerical values, we define a *normalized gaussian pdf* having zero mean and unit variance,

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}. \tag{8.51}$$

The cumulative distribution function corresponding to the gaussian pdf is

$$P\{X \le x\} = F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(y-m)^2/2\sigma^2} dy. \tag{8.52}$$

This integral cannot be evaluated in closed form and requires numerical evaluation. The procedure is to first normalize and then refer to a table of tabulated values.

As an example, suppose we wish to calculate $P\{X \le (m + k\sigma)\}$ for a gaussian random variable X. Then we can write

$$P\{X \le (m + k\sigma)\} = \int_{-\infty}^{m+k\sigma} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-m)^2/2\sigma^2} dx. \tag{8.53}$$

To normalize, let $z = (x - m)/\sigma$ so that

$$P\{X \le (m + k\sigma)\} = \int_{-\infty}^{k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz. \tag{8.54}$$

Note that this normalization results in the following rule in changing the limits of integration: Subtract the mean value and then divide by the standard deviation to determine the new limits for integration of the normalized gaussian.

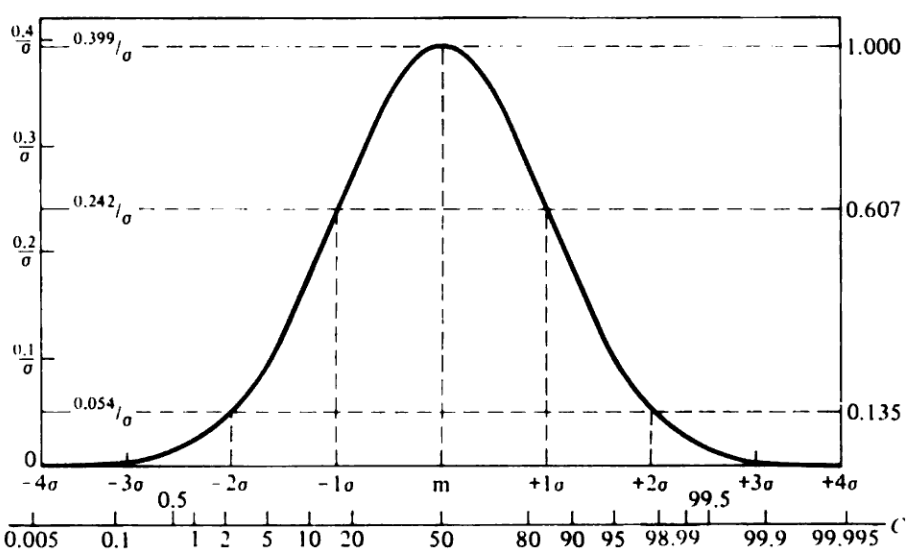


Figure 8.9 Gaussian (normal) probability density function. The standard deviation is σ, and scale C is the cumulative distribution function in percent.

As a result of the even symmetry in the integrand of Eq. (8.54), it is helpful to note that

$$\int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \frac{1}{2}. \tag{8.55}$$

For convenience in referring to tabulated values, we define the function $Q(x)$:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-z^2/2}\, dz. \tag{8.56}$$

As we observe in Fig. 8.10, $Q(x)$ represents the area beneath a plot of the normalized gaussian pdf over (x, ∞); the remainder of the area beneath this pdf is $[1 - Q(x)]$. Therefore our answer to the above example can be expressed as

$$P\{X \le (m + k\sigma)\} = 1 - Q(x). \tag{8.57}$$

Tabulated values of $Q(x)$ versus x are included in Appendix I.

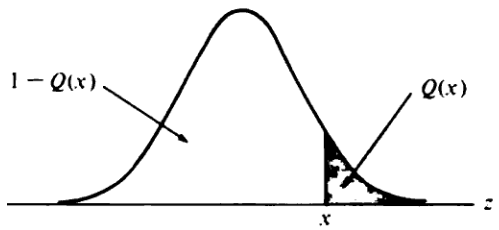


Figure 8.10 Area of the gaussian pdf, as described by Q(x).

EXAMPLE 8.6.5

The amplitude of a signal is gaussian-distributed with zero mean and a mean-square value σ^2. Find the probability of observing the signal amplitude above 3σ.

Solution

$$\begin{aligned} P\{X > 3\sigma\} &= \int_{3\sigma}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}\, dx \\ &= \int_{3}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \\ &= Q(3) = 0.0013 \ (\approx 10^{-3}) \end{aligned}$$

DRILL PROBLEM 8.6.5

A random variable X has a gaussian pdf. The mean value of X is +2 V and the variance is 4 V^2. Compute the following probabilities: (a) $P\{X < 6\}$; (b) $P\{X > 3\}$; (c) $P\{X < -2\}$; (d) $P\{2 < X < 3\}$.

ANSWER: (a) 0.9772; (b) 0.3085; (c) 0.0228; (d) 0.1915.

8.7 ★ THE HISTOGRAM

The probability density function provides much useful information about a random variable and allows us to determine statistical averages. But what if the pdf is not known? If experimental sample values of the random variable are available, we can estimate the pdf using a histogram.

A *histogram* is a tabulation of the frequency-of-occurrence of a random variable within preset ranges. The ranges for the tabulation are chosen to be adjacent and the total of the set of numbers gives the total number of events in the experiment. The histogram may be displayed by plotting the number of times that the random variable has been found in each range versus the value of the random variable. If the histogram is normalized by dividing by the total number of events, the height represents the relative frequency of occurrence within each range.

If the ranges are known or preset, the computation of a histogram is a straightforward number-tabulating routine and is easily implemented on a digital computer. This is illustrated in the following example.

✓ EXAMPLE 8.7.1

Compute the histogram of the sum of the first two digits in the fractional part of the square root of the integers from 1 to 1000.

Solution The total range of possible outcomes is from 0 (0 + 0) to 18 (9 + 9). Event counters are set in 19 memory registers such that register No. 0 increments its count by one if $X_i = 0$, register No. 1 increments its count by one if $X_i = 1$, etc. The results are listed below and are plotted in Fig. 8.11(a). (*Note:* A two-place round-off was used in determining the first two integers.)

Register Number	Event Count	Register Number	Event Count
0	31	10	88
1	12	11	82
2	31	12	71
3	41	13	64
4	46	14	50
5	57	15	43
6	65	16	24
7	82	17	20
8	91	18	0
9	102	Total	1000

The relative frequency can be found by dividing each count by the total (1000). (Note that the relatively large contribution at zero is caused by the 31 perfect squares which occur within that range.) On the basis of this histogram data, the mean and standard

deviations are 8.81 and 4.08, respectively. Also note that the histogram says nothing about the statistical independence (or lack of it) of the sample values.

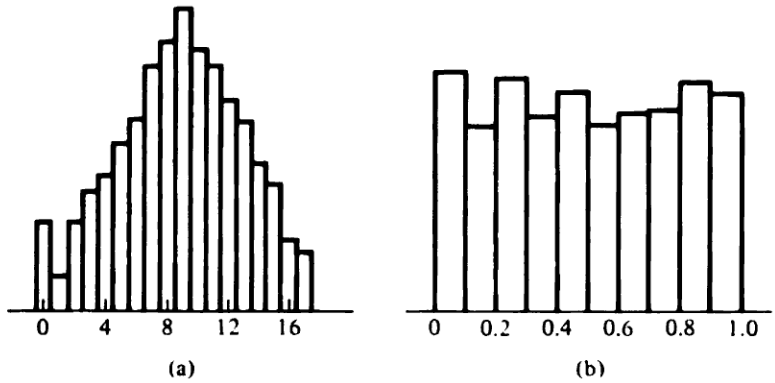


Figure 8.11 Histograms of (a) Example 8.7.1 and (b) Drill Problem 8.7.1.

✓ DRILL PROBLEM 8.7.1

A possible method of generating a "pseudorandom" number sequence within the range $0 \le X < 1$ in small calculators is to use the recursive formula:

$$x_i = \text{fractional part of } \{(\pi + x_{i-1})^5\}.$$

Compute the histogram of the first 1000 sample values of X, using $x_0 = 0$, when each number is truncated to one decimal place.

ANSWER: Histogram results for $0 \le X < 0.1$, $0.1 \le X < 0.2$, etc., are, respectively, 115, 88, 111, 93, 104, 88, 94, 96, 108, 103 (see Fig. 8.11b).

Computation of the histograms in Fig. 8.11 is straightforward because the ranges are preset. Each range interval for the tabulation in the histogram is often called a probability "cell." For instance, there are 19 cells in the Example 8.7.1 and 10 in Drill Problem 8.7.1. How many sample values should one take for each cell? This question is a typical one in applying a relative-frequency definition of probability and is not easy to answer concretely. "As many as possible" is desirable, of course, but a practical rule of about 100 per cell is reasonable.

If the cell widths are not known or preset, the generation of a meaningful histogram becomes more complicated and involves some compromises in setting the cell widths. On the one hand, narrow cell widths are desirable to give as accurate an approximation to the pdf as possible. On the other hand, wider cell widths allow more averaging and the variance in the estimate of the cell height is decreased. If the sample values can be stored (e.g., on magnetic tape), the results of a histogram using one set of cell widths can be used to suggest different widths for a second histogram

from the same data. The sample mean and variance can also be computed to aid in readjusting the cell widths and locations. The practical rule of tabulating at least 100 events per cell is applicable to retain some reliability in the estimate of the cell height. Obviously, the best histograms result from narrow cell widths and the tabulation of large amounts of data. It also helps to have some clues as to the nature of the pdf of the random variable before the tabulation is begun. For a more thorough discussion of these problems, the interested reader should consult the references.[†]

8.8 ★ TRANSFORMATIONS OF RANDOM VARIABLES

We are concerned with the transmission of signals through systems and are interested in how we can describe the effects of a system on a given input random variable. The approach here is to use the concept of a mapping of one random variable into another and to describe this in terms of probability density functions.

Let a random variable Y be a function $f(X)$ of a random variable X. We assume that $f(X)$ is known and has a continuous first derivative and that the pdf of X is known. Because probability is conserved in the mapping from X to Y, the probability of finding Y in a differential range dy is equal to the probability that X is in the corresponding range dx. To insure nonnegative pdf's, we use absolute magnitude signs and write:

$$p_Y(y)\,|dy| = p_X(x)\,|dx|\,, \tag{8.58}$$

or

$$p_Y(y) = p_X(x)\left|\frac{dx}{dy}\right|. \tag{8.59}$$

To obtain the transformation expressed in the new coordinate y, we assume that the inverse of $y = f(x)$ exists, so that our final result is

$$p_Y(y) = \left[p_X(x)\left|\frac{dx}{dy}\right|\right]_{x=f^{-1}(y)}. \tag{8.60}$$

Some intuitive insight into this procedure based upon a mapping of histograms is illustrated in Fig. 8.12. Here a histogram approximation to the output pdf $p_Y(y)$ is obtained from a histogram approximation of the input pdf $p_X(x)$ by a mapping using the output-input gain characteristic of the system, $y = f(x)$. Note that $f(x)$ is not linear (however, it is linear with a slope of 2 for $x < 2$ and linear with a slope of 1 for $x > 6$).

In the transformation, each probability cell in the histogram of the input is mapped into the corresponding cell in the histogram of the output. Each cell width in the out-

[†] See, for example, M. Schwartz and L. Shaw, *Signal Processing: Discrete Spectral Analysis, Detection, and Estimation*, New York: McGraw-Hill, 1975.

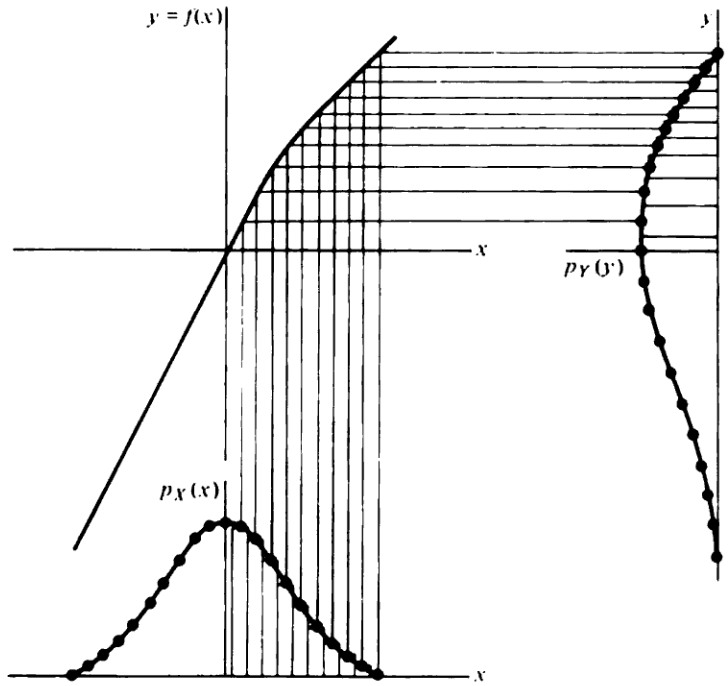


Figure 8.12 A histogram interpretation of the transformation of random variables.

put histogram is proportional to the width of the corresponding cell in the input histogram times the incremental slope (i.e., gain) of the output-input system gain characteristic. To conserve cell probability, each input histogram value must then be multiplied by the inverse of the magnitude of the incremental slope of the output-input gain characteristic. As the number of cells in the input histogram becomes very large, the transformation approaches the differential form expressed in Eq. (8.59). The final step taken in Eq. (8.60) is needed to express the result in the new coordinate y.

EXAMPLE 8.8.1

A gaussian-distributed random variable X with zero mean and unit variance is applied to a "full-wave" rectifier whose output-input gain characteristic is $y = |x|/a$, $a > 0$. Determine the pdf of the output random variable Y.

Solution The mapping is one-to-one for $x < 0$ and one-to-one for $x > 0$ (a sketch may help); in both cases, $y > 0$. Using Eq. (8.60) we have, for $x > 0$,

$$p_Y(y) = a\left[\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right]_{x=ay} = \frac{a}{\sqrt{2\pi}}e^{-a^2y^2/2}, \qquad y > 0,$$

and for $x < 0$

$$p_Y(y) = a\left[\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right]_{x=-ay} = \frac{a}{\sqrt{2\pi}}e^{-a^2y^2/2}, \qquad y > 0.$$

The solution is then

$$p_Y(y) = a\sqrt{\frac{2}{\pi}}e^{-a^2y^2/2}u(y).$$

It can be shown (see Appendix A) that the mean of Y is $(\sqrt{2/\pi})/a$ and the standard deviation is $\sqrt{1 - (2/\pi)}/a$.

DRILL PROBLEM 8.8.1

A random variable X is uniformly distributed over $(0, 2)$ and is applied to the input of a system whose output-input gain characteristic is $y = 2x + 1$. Determine the pdf of the output random variable Y.

ANSWER: $p_Y(y) = [u(y - 1) - u(y - 5)]/4$.

In cases in which the derivative of $f(X)$ has finite discontinuities, the above procedure can be altered somewhat to account for those probability cells that are all mapped into one. This can be handled conveniently by use of impulse functions and is demonstrated in the following example.

EXAMPLE 8.8.2

The random variable X of Example 8.8.1 is applied to the half-wave rectifier whose output-input gain characteristic is $y = (x/a)u(x)$. Determine the pdf of the output.

Solution For $x > 0$, the solution proceeds in the same way as for Example 8.8.1. For $x < 0$, however, all points of the input pdf are mapped into zero in the output. To conserve probability we must add a contribution of $\int_{-\infty}^{0} p_X(x)\,dx = 1/2$ at the point $y = 0$ so that

$$p_Y(y) = \frac{a}{\sqrt{2\pi}}e^{-a^2y^2/2}u(y) + (\tfrac{1}{2})\,\delta(y).$$

The mean of Y is $(\sqrt{2\pi}\,a)^{-1}$, and the standard deviation is $\sqrt{\pi - 1} \times (\sqrt{2\pi}\,a)^{-1}$.

DRILL PROBLEM 8.8.2

The random variable X of Drill Problem 8.8.1 is applied to the input of a system whose output-input gain characteristic is (a) $y = x - (x - 1)u(x - 1)$; (b) $y = xu(x - 1)$. Determine the pdf of the output.

ANSWER: (a) $\frac{1}{2}[u(y) - u(y - 1)] + (\frac{1}{2})\delta(y - 1)$; (b) $\frac{1}{2}[u(y - 1) - u(y - 2)] + (\frac{1}{2})\delta(y)$.

8.9 JOINT AND CONDITIONAL DENSITY FUNCTIONS

In some cases we require two or more random variables for a description of a physical process. The same approach that was used for probability density functions of one dimension can be extended to more than one dimension. For simplicity we consider only the two-dimensional case.

The probability of finding X in the range $x_1 < X \le x_2$ while also finding Y in the range $y_1 < Y \le y_2$ is [cf. Eq. (8.27)]

$$P(x_1 < X \le x_2, y_1 < Y \le y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} p_{XY}(x, y)\, dx\, dy, \tag{8.61}$$

where $p_{XY}(x, y)$ is the *joint probability density function* of the random variables X and Y. For the two-dimensional case, probability equals the volume under the surface described by $p_{XY}(x, y)$, as depicted in Fig. 8.13.

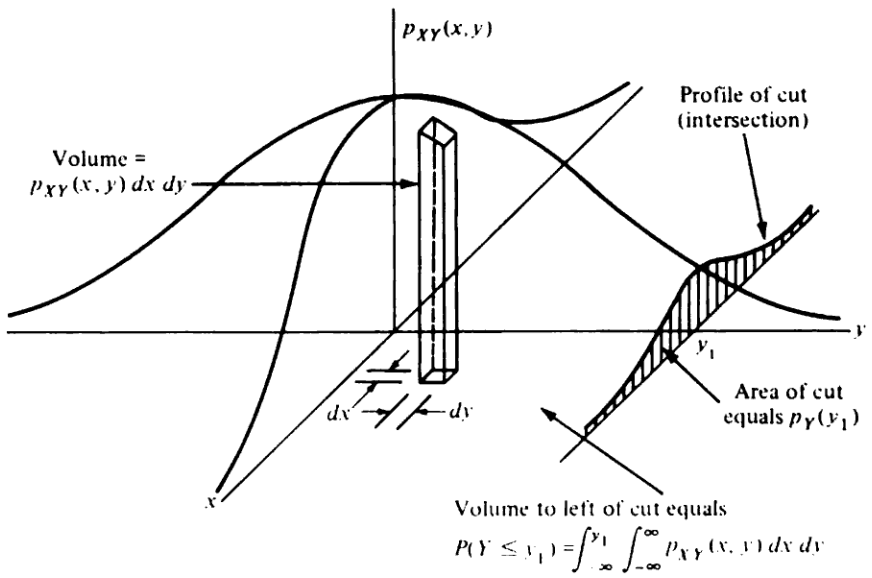


Figure 8.13 A geometric interpretation of the joint probability density function.

Given the joint probability density function $p_{XY}(x, y)$, the probability density function $p_Y(y)$ can be found by obtaining the cumulative distribution function $F_Y(y)$ in the following manner. If we are interested only in Y, the value of X does not matter, so that

$$P(Y \leq y) = P(X \leq \infty, Y \leq y), \tag{8.62}$$

or

$$F_Y(y) = F_{XY}(\infty, y), \tag{8.63}$$

yielding

$$F_Y(y) = \int_{-\infty}^{y} \int_{-\infty}^{\infty} p_{XY}(x, y)\, dx\, dy. \tag{8.64}$$

Differentiating both sides of Eq. (8.64) with respect to y, we obtain

$$p_Y(y) = \int_{-\infty}^{\infty} p_{XY}(x, y)\, dx. \tag{8.65}$$

Thus the probability density function of Y can be found by integrating the joint probability density function $p_{XY}(x, y)$ over all values of X. Similar arguments in the other dimension yield $p_X(x)$.

A geometric interpretation of Eqs. (8.64) and (8.65) is shown in Fig. 8.13. Intersecting the surface $p_{XY}(x, y)$ by the plane $y = y_1$, we obtain the cross-sectional profile shown. From Eq. (8.65), we see that the cross-sectional area gives $p_Y(y_1)$, and $p_Y(y)$ can be determined by finding the dependence of the cross-sectional area on y. Note that the volume to the left of the intersection in Fig. 8.13 is equal to the integral of $p_Y(y)$ over all $y \leq y_1$; i.e.,

$$\int_{-\infty}^{y_1} p_Y(y)\, dy = \int_{-\infty}^{y_1} \int_{-\infty}^{\infty} p_{XY}(x, y)\, dx\, dy,$$

verifying Eq. (8.64). Similar arguments in the other dimension yield $p_X(x)$. The probability density functions $p_X(x)$ and $p_Y(y)$ are called *marginal densities*.

It is intuitively appealing to attempt to identify the profile formed by the intersection of the plane $y = y_1$ and the surface $p_{XY}(x, y)$ as the conditional probability density for X given $Y = y_1$. An investigation shows, however, that this profile function would not generally qualify as a probability density function because it does not necessarily have unit area. To achieve this requirement, we can normalize by dividing by the area. Thus we can write

$$p_X(x \mid y_1) = \frac{p_{XY}(x, y_1)}{\text{area shaded}},$$

$$= \frac{p_{XY}(x, y_1)}{p_Y(y_1)}.$$

Similar arguments hold for $p_Y(y|x)$ so that, in general, this result can be rewritten as[†]

$$p_{XY}(x,y) = p_X(x)p_Y(y|x) = p_Y(y)p_X(x|y). \tag{8.66}$$

In the particular case in which the two random variables X and Y are statistically independent, $p_X(x|y) = p_X(x)$, $p_Y(y|x) = p_Y(y)$, and Eq. (8.66) reduces to

$$p_{XY}(x,y) = p_X(x)p_Y(y). \tag{8.67}$$

In the case of two or more random variables, the concept of expected values can be extended by using multiple integration over the joint probability density function in the general form,

$$E\{g(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p_{XY}(x,y)\,dx\,dy. \tag{8.68}$$

EXAMPLE 8.9.1

Let X and Y be two random variables with variances σ_X^2 and σ_Y^2, respectively. Let the sum be $Z = X + Y$, with variance σ_Z^2. Determine σ_Z^2 in terms of σ_X^2, σ_Y^2 if X and Y are statistically independent.

Solution For convenience, let the means of X and Y be zero so that

$$\sigma_Z^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)^2 p_{XY}(x,y)\,dx\,dy.$$

If X, Y are statistically independent, then

$$\begin{aligned}\sigma_Z^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)^2 p_X(x)p_Y(y)\,dx\,dy \\ &= \int_{-\infty}^{\infty} x^2 p_X(x)\,dx + \int_{-\infty}^{\infty} y^2 p_Y(y)\,dy \\ &\quad + 2\int_{-\infty}^{\infty} x p_X(x)\,dx \int_{-\infty}^{\infty} y p_Y(y)\,dy.\end{aligned}$$

Because both factors of the last term of this equation are zero by definition, we have our desired result:

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$$

if X, Y are statistically independent. It is left to the student to show that this result also holds for nonzero means. Therefore, we conclude that the variance of a sum is the sum of the variances if the random variables are statistically independent.

[†] The random variable subscript notation becomes cumbersome and is often deleted. Doing so streamlines the notation, but the random variable subscripts should be replaced if possible confusion could result in the notation.

EXAMPLE 8.9.2

A random variable Y is composed of the mean of N statistically independent random variables $X_1, X_2, \ldots, X_N$, each with a mean of m_X and variance σ_X^2. (a) Determine m_Y in terms of m_X. (b) Determine σ_Y^2 in terms of σ_X^2.

Solution We can write

$$Y = \frac{1}{N}\sum_{i=1}^{N} X_i = \sum_{i=1}^{N} X_i/N.$$

a) $$m_Y = E\left\{\frac{1}{N}\sum_{i=1}^{N} X_i\right\} = \frac{1}{N}\sum_{i=1}^{N} E\{X_i\},$$

$$m_Y = m_X.$$

b) Treating X_i/N as new random variables and using the result of Example 8.9.1, we have

$$\sigma_Y^2 = \sigma_1^2/N^2 + \sigma_2^2/N^2 + \cdots + \sigma_N^2/N^2 = N\sigma_X^2/N^2,$$

$$\sigma_Y^2 = \sigma_X^2/N.$$

Combining these two results, we conclude that if an experiment having statistical regularity is repeated, then our confidence that the averaged results are closer to the true results is improved. Stated more precisely, the width of the pdf of the random variable describing the experiment narrows proportional to $\sqrt{N}$.

DRILL PROBLEM 8.9.1

Let X and Y be two random variables with means m_X and m_Y respectively. Let the sum be $Z = X + Y$, with mean m_Z. Determine m_Z in terms of m_X and m_Y.

ANSWER: $m_Z = m_X + m_Y$ (i.e., "the mean of a sum is the sum of the means.") Note that this result holds regardless of whether or not the random variables are statistically independent and is a special case of the more general statement that the expected value operation is linear.

8.10 ★ CORRELATION BETWEEN RANDOM VARIABLES

In an analogy to the variance of one random variable, the *covariance* of two random variables X and Y is defined as

$$\mu \triangleq E\{(X - m_X)(Y - m_Y)\}. \tag{8.69}$$

If X and Y are statistically independent, use of Eq. (8.67) in Eq. (8.69) gives

$$\begin{aligned}\mu &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - m_X)(y - m_Y)p_{XY}(x, y)\,dx\,dy \\ &= \int_{-\infty}^{\infty}(x - m_X)p_X(x)\,dx\int_{-\infty}^{\infty}(y - m_Y)p_Y(y)\,dy \\ &= (m_X - m_X)(m_Y - m_Y) = 0 .\end{aligned} \tag{8.70}$$

Thus the covariance is zero if X and Y are statistically independent. Next suppose that X and Y are completely dependent so that either $X = Y$ or $X = -Y$. In these cases, we find that Eq. (8.69) gives

$$\mu = \sigma_X^2 = \sigma_Y^2 = \sigma_X\sigma_Y , \tag{8.71}$$

and

$$\mu = -\sigma_X^2 = -\sigma_Y^2 = -\sigma_X\sigma_Y . \tag{8.72}$$

At this point it is convenient to define a normalized covariance that is called the *correlation coefficient:*

$$\rho \triangleq \frac{\mu}{\sigma_X\sigma_Y} . \tag{8.73}$$

Because Eqs. (8.71) and (8.72) are the extreme cases, we conclude that

$$-1 \le \rho \le 1 . \tag{8.74}$$

The correlation coefficient ρ serves as a measure of the extent to which X and Y are dependent. When $\rho = 0$, the random variables X and Y are said to be *uncorrelated.*

From Eqs. (8.70) and (8.73) we conclude that *if two random variables are statistically independent then they are uncorrelated.* However, *the converse is not necessarily true.*

In statistical testing, usually the joint probability function is unknown and a uniform distribution is assumed. Thus Eqs. (8.69) and (8.73), for the discrete case, become

$$\rho = \frac{\Sigma_i(x_i - m_X)(y_i - m_Y)}{[\Sigma_i(x_i - m_X)^2\Sigma_i(y_i - m_Y)^2]^{1/2}} . \tag{8.75}$$

Equation (8.75) is readily calculated from numerical data.

The correlation coefficient, as defined here, refers to a linear correlation between the random variables X, Y. Thus the correlation coefficient is not so much a measure of how dependent the random variables X, Y are in a general sense but is a measure of the linearity of that relationship. In fact, the square of the correlation coefficient can be used as a relative measure, on a mean-square basis, of how well a straight line can be fit to the sample values of X, Y. For example, the two random variables X, Y may have perfect dependence such as in the relation $X^2 + Y^2 = 1$, yet may have a zero correlation coefficient [as they do if uniformly distributed over $(-1, 1)$]. This behavior is also illustrated in Example 8.10.1.

Computation of the correlation coefficient is a convenient indicator of how dependent two random variables are on each other within the above restriction. For example, the correlation coefficient determined between adjacent sample values (that is, x_i, x_{i-1}) in Example 8.7.1 is $\rho = 0.391$, while for Drill Problem 8.7.1, $\rho = 5.96 \times 10^{-4}$. Thus the successive sample values of the latter number generator are much more uncorrelated than the former.

An expansion of Eq. (8.69) gives

$$\mu = E\{XY\} - m_X m_Y . \tag{8.76}$$

The *correlation* of the two random variables X and Y is defined as the first term in the right-hand side of Eq. (8.76):

$$E\{XY\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy p_{XY}(x, y)\, dx\, dy . \tag{8.77}$$

For zero mean values of either X or Y, the correlation is equal to the covariance of X and Y.

EXAMPLE 8.10.1

Let the joint probability density function of the random variables X, Y be uniform over the unit circle defined by $X^2 + Y^2 \leq 1$, and suppose that $Y = X^2$. Show that even though X and Y are not statistically independent, they are uncorrelated.

Solution Using Eq. (8.65), we have

$$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x, y)\, dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \left(\frac{1}{\pi}\right) dy = \frac{2}{\pi}\sqrt{1 - x^2}, \qquad -1 \leq x < 1 ,$$

so that

$$E\{X\} = \frac{2}{\pi} \int_{-1}^{1} x\sqrt{1 - x^2}\, dx = 0 .$$

Equation (8.77) gives

$$E\{XY\} = \int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x^3 \left(\frac{1}{\pi}\right) dx\, dy = 0 .$$

Therefore $\rho = 0$ [cf. Eqs. (8.73) and (8.76)] and X, Y are uncorrelated.

DRILL PROBLEM 8.10.1

Compute the correlation coefficient of the first six successive digits in the fractional part of π.

ANSWER: -0.194.

EXAMPLE 8.10.2

(Linear Regression) Suppose that a group of data points described by the random variables X, Y are to be approximated by the straight line $Y = aX + b$. The best straight-line fit is defined here by minimizing the mean-square error $E\{[Y - (aX + b)]^2\}$.

a) Find a and b in terms of the statistical averages of X, Y.

b) Show that ρ^2 is a relative measure of how well the straight line fits the data.

Solution a)

$$\begin{aligned} E\{[Y - (aX + b)]^2\} &= E\{Y^2 - 2aXY - 2bY + a^2X^2 + 2abX + b^2\} \\ &= E\{Y^2\} - 2aE\{XY\} - 2bE\{Y\} + a^2E\{X^2\} \\ &\quad + 2abE\{X\} + b^2. \end{aligned}$$

Taking a partial derivative with respect to b and setting the result equal to zero, we obtain

$$-2E\{Y\} + 2aE\{X\} + 2b = 0$$

or

$$b = m_Y - am_X .$$

Using this value of b in the equation above, we get

$$\begin{aligned} E\{[Y - (aX + b)]^2\} &= E\{[Y - (aX + m_Y - am_X)]^2\} \\ &= E\{[Y - m_Y) - a(X - m_X)]^2\} \\ &= E\{(Y - m_Y)^2 - 2a(X - m_X)(Y - m_Y) + a^2(X - m_X)^2\} \\ &= \sigma_Y^2 - 2a\mu + a^2\sigma_X^2 . \end{aligned}$$

Taking a partial derivative with respect to a and setting the result equal to zero, we obtain

$$-2\mu + 2a\sigma_X^2 = 0$$

or

$$a = \mu/\sigma_X^2 \quad \text{and} \quad b = m_Y - \frac{\mu}{\sigma_X^2} m_X .$$

b)

$$\begin{aligned} E\{[Y - (aX + b)]^2\} &= \sigma_Y^2 - 2a\mu + a^2\sigma_X^2 \\ &= \sigma_Y^2 - 2\frac{\mu^2}{\sigma_X^2} + \frac{\mu^2}{\sigma_X^2} \\ &= \sigma_Y^2 - \frac{\mu^2}{\sigma_X^2\sigma_Y^2}\sigma_Y^2 \\ &= \sigma_Y^2(1 - \rho^2) . \end{aligned}$$

The value of ρ^2 is called the *coefficient of determination*. Note that for $\rho = 1$, there is a perfect fit of the data to the straight line calculated from the results above. Many modern electronic calculators are preprogrammed to perform these calculations.

DRILL PROBLEM 8.10.2

Find the best straight-line fit to the following six (X, Y) data points and calculate the coefficient of determination: (1.0, 1.1), (2.0, 1.4), (3.0, 3.1), (4.0, 2.5), (5.0, 5.9), (6.0, 3.2).

ANSWER: $y = 0.67x + 0.53$; $\rho^2 = 0.53$.

8.11 ★ THE BIVARIATE GAUSSIAN DISTRIBUTION

A random variable or "variate" representing the outcome of a given experiment may require more than one dimension for its description of all possible outcomes. In this case we can speak of a "multivariate" probability distribution. Such a multivariate probability distribution may assume any form, as long as the integration or summation over all possible outcomes is normalized to 1. However, we only discuss one very important case here—that of the bivariate (two-dimensional) gaussian distribution.

Two random variables X and Y are said to be jointly gaussian if their joint probability density is

$$p_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{\frac{-\left[\dfrac{(x-m_X)^2}{\sigma_X^2} + \dfrac{(y-m_Y)^2}{\sigma_Y^2} - \dfrac{2\rho(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y}\right]}{2(1-\rho^2)}\right\}. \tag{8.78}$$

We can make the following conclusions for the bivariate gaussian distribution based on Eq. (8.78) and some integrations (which we do not include here).

1. *The bivariate gaussian distribution depends only on the first- and second-order moments of the random variables.* Thus Eq. (8.78) requires only the means, variances, and covariances of the random variables.
2. *All individual (one-dimensional) pdf's are gaussian.* For example, an application of Eq. (8.65) to Eq. (8.78) gives

$$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x, y)\,dy = \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-(x-m_X)^2/2\sigma_X^2}. \tag{8.79}$$

3. *All conditional pdf's are gaussian.* This follows from a use of Eq. (8.66):

$$p(x \mid y) = \frac{p_{XY}(x, y)}{p_X(x)},$$

and Eqs. (8.78) and (8.79).

4. *Uncorrelated gaussian variates are also statistically independent.* (The converse is true for any distribution.) This is easily shown by setting $\rho = 0$ in Eq. (8.78) so that [cf. Eq. (8.79)]

$$p_{XY}(x, y) = p_X(x)p_Y(y).$$

5. *A linear combination of jointly gaussian variates has a gaussian distribution.* For example, if

$$Z = \alpha X + \beta Y,$$

where X and Y are jointly gaussian, then Z is gaussian with

$$m_Z = \alpha m_X + \beta m_Y,$$
$$\sigma_Z^2 = \alpha^2\sigma_X^2 + \beta^2\sigma_Y^2 + 2\rho\alpha\beta\sigma_X\sigma_Y.$$

As a result of these properties and the central limit theorem, the gaussian random variable is a popular choice for a probability model in engineering problems. The above concepts (and conclusions) can be extended to more than two variates. The resulting expressions become unwieldy, however, unless matrix notation is adopted.

8.12 RANDOM PROCESSES

To determine the probabilities of the various possible outcomes of an experiment, we repeat the experiment many times. For example, suppose we are interested in establishing the statistics associated with the tossing of a die. We might proceed by either of the two following methods. One the one hand, we might toss a large number of dice simultaneously. Alternatively, we could toss a single die many times. Intuitively, we would expect that both methods would yield the same results if all dice were identical and remained identical throughout the experiment. Assuming honest dice, we would expect, for example, that a "1" would show on one-sixth of the dice in the first method and would show one-sixth of the time in the second method.

Up to this point, we have conducted our experiments using the first method. However, our interests in the application of probabilistic concepts to communication systems actually are closer to the second method. Under what conditions can we expect the same results? We shall investigate this question in this section. But first we must specify more precisely just what we mean by the two methods above.

Conceptually, we could consider the second method as a random number generator which puts out a number between 1 and 6 at the command of a clock. Now suppose we have many of these random number generators and that they are all operated

from commands from a common clock. Such a collection of waveform sources is called an *ensemble*. The individual random number generator waveforms (as a function of time) are called *sample functions*. The sample functions can be written as $X(t, \lambda_i)$ when there are i discrete outcomes possible. The ensemble of possible sample functions is called a *random process*. This is illustrated in Fig. 8.14.

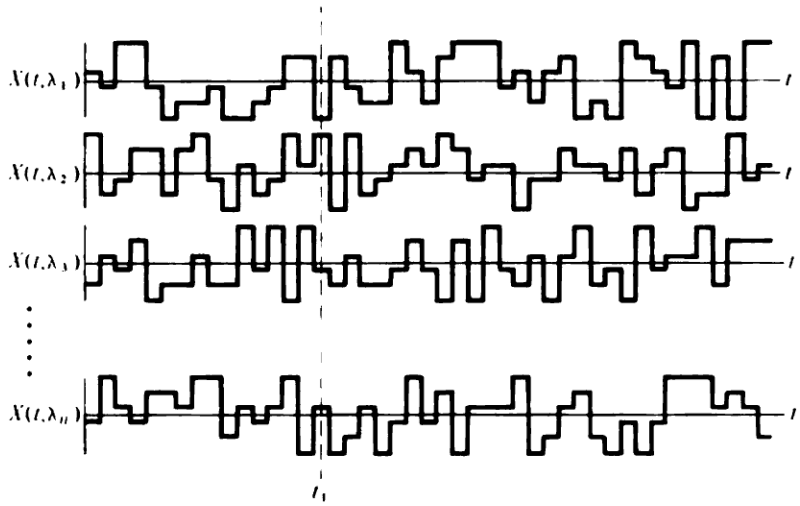


Figure 8.14 An ensemble of random waveforms.

The statistics of the waveforms in our ensemble may be determined from measurements made at some fixed time $t = t_1$ on all the sample functions. Thus we can write $X(t_1, \lambda_i)$, which is just another way of writing $X(\lambda_i)|_{t_1}$. Therefore a random process $X(t, \lambda_i)$ at a fixed time is simply the random variable X. A statistical average can be determined from the measurements made on $X(\lambda_i)|_{t_1}$ and the average so determined is called an *ensemble average*. For example, we could toss 100 coins simultaneously at $t = t_1$ and determine the statistics on the basis of this experiment. Note that the statistics and averages obtained in this method are those which we have been discussing in previous sections.

On the other hand, we could choose to observe just one sample function for a period of time and determine its statistics. (In the above example, this would correspond to taking one coin and tossing it 100 times in succession.) From these observations, we can compute averages that we call *time averages*.

We can illustrate some of these ideas by plotting the output of a number of identical random waveform generators as shown in Fig. 8.14. Such a collection of waveforms is called an ensemble and is described by the random process $X(t, \lambda_i)$. Each generator output as a function of time is one sample function of the ensemble. Values of all waveforms in the ensemble observed at a particular time $t = t_1$ are described by the ensemble statistics and designated by the random variable X.

An ensemble average is found by using the ensemble statistics at a particular time. A time average is found by using time averaging on a particular sample function. The question then arises: Are time averages equal to ensemble averages?

In general, ensemble averages and time averages are not equal. Suppose, for example, that the statistical characteristics of the sample functions in the ensemble varied with time. Then the ensemble averages would be different at different times but the time averages would not reflect this variation. When the statistical characteristics of the sample functions do not change with time, a random process is said to be *stationary*. However, even the property of being stationary does not ensure that the ensemble and the time averages are the same. For it may happen that while each sample function is stationary, the individual sample functions may differ statistically from one another. In this case the time averages will depend on the particular sample function that is used to form the average.

When the nature of a random process is such that ensemble averages and time averages are equal, the process is referred to as being *ergodic*. Hence if $X(t, \lambda_i)$ is ergodic, then all its statistics can be determined from a single sample function of the random process. *An ergodic random process is stationary but a stationary random process is not necessarily ergodic*. For example, the random process $X(t) = C$, where C is a random constant, is stationary but not ergodic.

The foregoing reasoning can be applied to the continuous as well as to the discrete case. Suppose, for example, that the ensemble is composed of many random noise generators. Each generates a sample function and at a particular instant of time the outputs of all noise generators may be observed to find an ensemble average. If the random process describing the outputs of the noise generators is ergodic, then observations made on one sample function over a period of time are sufficient to obtain the statistics of the random process.

We shall consider only those random processes which are ergodic. Thus we can equate ensemble averages [cf. Eqs. (8.34) and (8.35)] with the time averages [cf. Eqs. (4.23) and (4.24)]:

$$E\{X\} = \int_{-\infty}^{\infty} xp(x)\,dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\,dt, \tag{8.80}$$

$$E\{X^2\} = \int_{-\infty}^{\infty} x^2 p(x)\,dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)\,dt. \tag{8.81}$$

Also, the cumulative distribution function and the probability density function can now be determined from observations of one sample function just as well as from observations of an ensemble of sample functions at one particular instant in time.

A consideration of Eqs. (8.80) and (8.81) for ergodic random processes allows us to make the following statements:

1. The average or mean value, $E\{X\}$, is equal to the dc level of the signal.
2. The square of the mean, $[E\{X\}]^2$, is equal to the power in the dc component across 1 ohm.
3. The mean-square value, $E\{X^2\}$, is equal to the total average power of the signal $x(t)$ across 1 ohm.
4. The square root of the mean-square value is the rms value.

5. The variance, σ_X^2, is equal to the average power in the time-varying or ac component of the signal across 1 ohm.
6. The standard deviation, σ_X, is the rms value of the time-varying (ac) component of the signal.
7. If $E\{X\} = 0$, then σ_X is the rms value of the signal.

These relations tend to make us feel much more comfortable in the use of random signals. However, it must be remembered that these equalities hold only for ergodic random processes. As a result of the close tie with signal analysis, as indicated above, we shall designate a random process by the same type of notation that we use for deterministic waveforms so that $X(t, \lambda)$ will be written simply as $x(t)$.

It is instructive to review the relationship of the cumulative distribution function and the probability density function of an ergodic random process $x(t)$. Consider, for example, the sample function of an ergodic random process as shown in Fig. 8.15(a). This waveform is applied to the input of a comparator, which in turn activates a counter connected to a clock. The proportionate amount of time that the waveform spends below the preset threshold of the comparator, x_0, is then recorded versus the setting x_0 and is shown in Fig. 8.15(b). Because we have assumed ergodicity, we now recognize that this describes the cumulative distribution function of the random variable X associated with the random process $x(t)$. Taking the derivative of the cumulative distribution function yields the pdf as shown in Fig. 8.15(c).

Another possibility in making a measurement of the pdf of an ergodic random process directly is to use a series of threshold devices each of which detects whenever the input waveform is within a specified amplitude range. When the amplitude of the input waveform is within the specified range for a given threshold detector, a constant-level gate pulse is produced and integrated. At the end of the desired measurement interval, the outputs of the integrators are sampled and displayed, as shown in Fig. 8.16. As the time interval used becomes large, the resulting samples approximate the

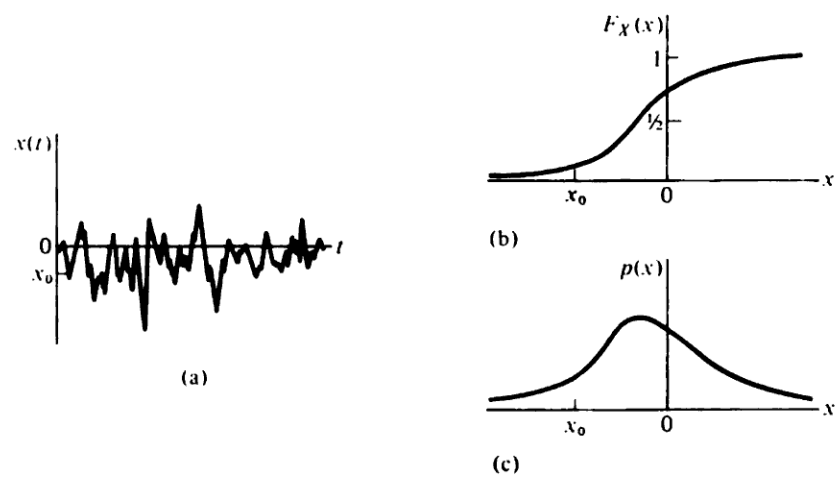


Figure 8.15 A sample function of an ergodic random process and its cumulative distribution function and pdf.

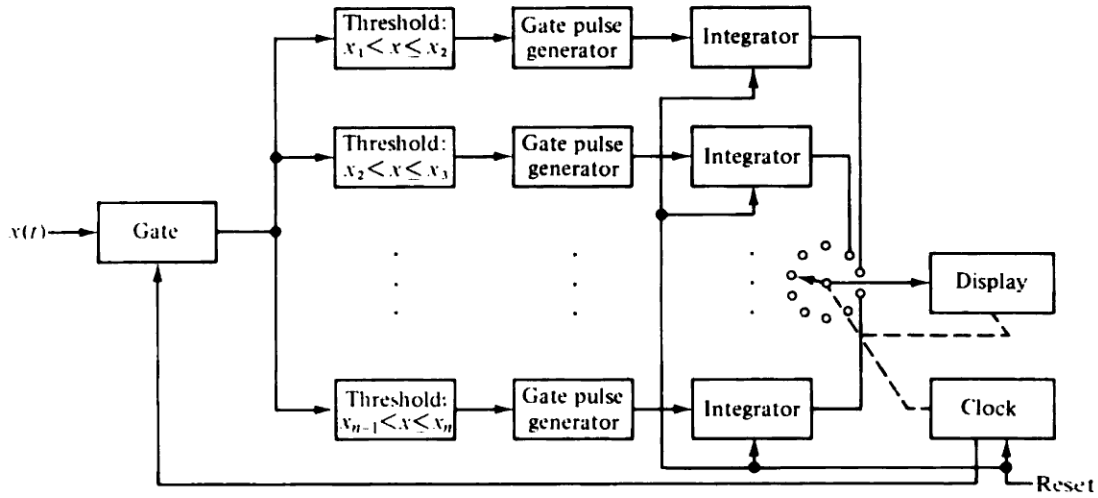


Figure 8.16 Relative measurement of the probability density function of an ergodic random process.

probability density function if they are normalized by the length of the time interval. A device available in the laboratory for performing this measurement is called a "multichannel analyzer." Note the similarities with the multichannel spectral analyzer shown in Fig. 4.2.

EXAMPLE 8.12.1

Suppose we wish to trigger an alarm system whenever a given signal voltage is present. There is also additive noise present; this noise voltage is gaussian-distributed with zero mean value. The alarm has a preset threshold such that whenever the total input voltage exceeds the threshold the alarm triggers and rings a bell.

a) It is desired to set the threshold at the lowest voltage level possible and yet have the alarm ring the bell no more than 1% of the time, on the average, when there is no signal present. Determine the required threshold voltage if the rms value of the noise is 0.5 V.

b) Assume that when the signal is present, this signal is a positive 2 V. With the threshold set as in part (a), what is the probability that the bell will not ring even when the signal is present?

Solution a) When no signal is present, $m_x = 0$ and $\sigma_x = \frac{1}{2}$. We determine the threshold x_0 such that (see Fig. 8.17a)

$$\int_{x_0}^{\infty} \frac{2}{\sqrt{2\pi}} e^{-2x^2}\, dx = 0.01 .$$

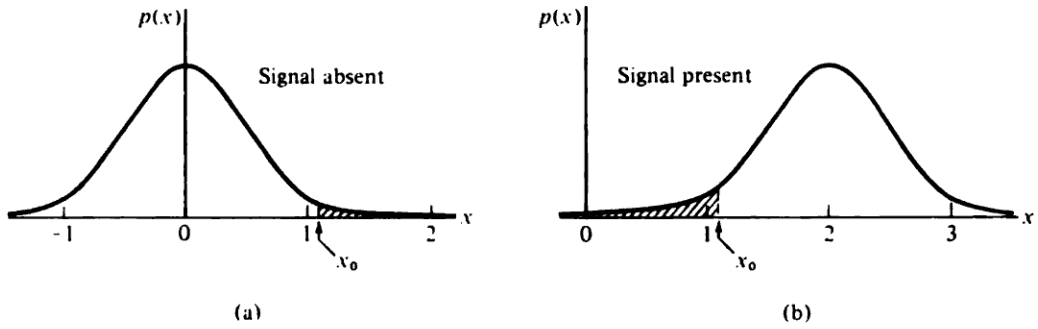


Figure 8.17 The probability density functions of Example 8.12.1.

Changing the variable, let $z = 2x$ so that this becomes

$$\int_{2x_0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = 0.01\,,$$

or

$$Q(2x_0) = 0.01\,.$$

Using the tables in Appendix I, we find that the required threshold value is $x_0 = 1.163$ V.

b) When the signal is present, $m_x = 2$ and $\sigma_x = \frac{1}{2}$. The probability that the bell will not ring even if the signal is present is (see Fig. 8.17)

$$\begin{aligned} P(x \le x_0) &= \int_{-\infty}^{1.163} \frac{2}{\sqrt{2\pi}} e^{-2(x-2)^2}\, dx \\ &= \int_{-\infty}^{-1.674} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \\ &= \int_{1.674}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \\ &= Q(1.674) = 0.0471\,. \end{aligned}$$

DRILL PROBLEM 8.12.1

(a) Determine a new alarm threshold setting for Example 8.12.1 such that the probability of the bell ringing with no signal present (i.e., a false alarm) is exactly twice that of the bell not ringing when a signal is present (i.e., a miss). (b) What is the threshold that will make these equal?

ANSWER: (a) 0.927 V; (b) 1.000 V.

8.13 AUTOCORRELATION AND POWER SPECTRA

Suppose that we have a random process $x(t)$, such as the one for which several sample functions are exhibited in Fig. 8.18. At a time $t = t_1$, the random process $x(t)$ is characterized by the random variable $X_1 = x(t_1)$. Similarly, at a time $t = t_2$, the random process $x(t)$ is characterized by the random variable $X_2 = x(t_2)$. The possibility now arises that we characterize the random process $x(t)$ by use of a joint probability density function of these random variables. If we were to completely describe the random process in this way we would, of course, have to use an infinite number of random variables and an infinite-dimensional joint density. This is not practical and therefore we limit ourselves to a two-dimensional joint density.

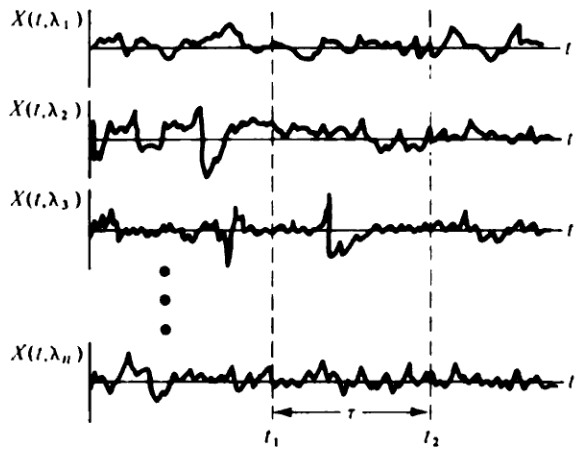


Figure 8.18 An ensemble of continuous random waveforms.

The first moment of the random process under consideration is

$$E\{x(t)\} = m_x(t) . \tag{8.82}$$

If the random process $x(t)$ is stationary, then its statistics do not change as a function of time and the first moment becomes a constant:

$$E\{x(t)\} = m_x \qquad x(t) \text{ stationary.} \tag{8.83}$$

We characterize the random process $x(t)$ in terms of a two-dimensional joint density $p_{X_1X_2}(x_1, x_2)$, where $X_1 = x(t_1)$ and $X_2 = x(t_2)$ (such as depicted in Fig. 8.18). There are three different second (joint) moments possible: $E\{X_1^2\}$, $E\{X_2^2\}$, and $E\{X_1X_2\}$. This last joint moment is of particular importance and is given the special notation,†

$$E\{x(t_1)x(t_2)\} \triangleq R_{xx}(t_1, t_2) , \tag{8.84}$$

† This definition is for a real-valued random process; for the complex-valued case, the second term in the expectation has a complex conjugate operation.

or, writing this out,

$$R_{xx}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p_{X_1 X_2}(x_1, x_2)\, dx_1\, dx_2 .$$

The function $R_{xx}(t_1, t_2)$ is called the *autocorrelation function* of the random process $x(t)$ evaluated at $t = t_1, t_2$. The term *autocorrelation* has been used previously (cf. Chapter 4) and we shall soon see why the same word is used again here.

If the random process $x(t)$ is stationary, the second joint moment is independent of the absolute time reference and is only dependent on the time difference between t_1 and t_2 (see Fig. 8.18) so that we can write

$$R_{xx}(t_1, t_2) = R_{xx}(t_2 - t_1) \qquad x(t) \text{ stationary.} \tag{8.85}$$

In this case, we can let $\tau = t_2 - t_1$, or $t_2 = t_1 + \tau$, and then delete the subscript on t_1 to obtain the simpler notation

$$R_{xx}(\tau) = E\{x(t)x(t + \tau)\} \qquad x(t) \text{ stationary.} \tag{8.86}$$

If the random process $x(t)$ is stationary, then Eq. (8.86) holds. The converse statement is not necessarily true but is defined by the weaker condition of a "wide-sense stationary" random process. In a wide-sense stationary random process the joint probability density function is dependent only on time differences between t_1 and t_2, but not on the actual values of t_1 and t_2.

The autocorrelation function is the second joint moment and therefore does not completely describe or define the random process $x(t)$ (i.e., it is not unique), but it does give us much information about $x(t)$. In particular, it gives us a measure of how dependent a particular value of a sample function is on another value that is removed τ units of time. For example, if t_1 and t_2 are sufficiently far apart such that $x(t_1)$ and $x(t_2)$ are statistically independent, the autocorrelation function reduces to

$$\begin{aligned} R_{xx}(\tau) &= E\{x(t)x(t + \tau)\} \\ &= E\{x(t)\}E\{x(t + \tau)\} \\ &= m_x^2 . \end{aligned} \tag{8.87}$$

If the mean value of $x(t)$ is zero, that value of τ over which $R_{xx}(\tau)$ first goes to zero represents the time separation required for the random process to become "uncorrelated." In other words, the shape of $R_{xx}(\tau)$ gives us some idea as to how past values affect present values of $x(t)$ and hence how fast a particular sample function can vary with time.

Thus for $m_x = 0$ we say that $x(t)$ and $x(t + \tau)$ are *uncorrelated* if $R_{xx}(\tau) = 0$. If $x(t)$ and $x(t + \tau)$ are statistically independent, they are uncorrelated. (However, if $x(t)$ and $x(t + \tau)$ are uncorrelated, they are not necessarily independent.)

In addition, if the random process $x(t)$ is ergodic (recall that ergodicity implies stationarity but not vice versa), the autocorrelation function can be found from any one sample function by use of time-averaging, or

$$R_{xx}(\tau) = E\{x(t)x(t+\tau)\}, \tag{8.88}$$

$$R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t+\tau)\,dt.$$

At this point we can see why the functional notation $R(\tau)$ and the name "autocorrelation" were used in Chapter 4. In fact, we can now tie our results together so that the relations for deterministic signals become special cases of the relations developed for random processes. For example, the power spectral density of a stationary random process $x(t)$ is given by

$$S_x(\omega) = \mathcal{F}\{R_{xx}(\tau)\} = \int_{-\infty}^{\infty} R_{xx}(\tau)e^{-j\omega\tau}\,d\tau. \tag{8.89}$$

If $x(t)$ is ergodic, then $R_{xx}(\tau)$ is given by Eq. (8.88), agreeing with the earlier result in Chapter 4. Conversely, the autocorrelation function is given by the inverse Fourier transform of the power spectral density:

$$R_{xx}(\tau) = \mathcal{F}^{-1}\{S_x(\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)e^{j\omega\tau}\,d\omega. \tag{8.90}$$

The average power in $x(t)$ is (for a 1-ohm load)

$$P = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x^2(t)\,dt = E\{x^2(t)\} = R_{xx}(0),$$

and from Eq. (8.90),

$$R_{xx}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)\,d\omega,$$

agreeing with the concept of a power spectral density developed in Chapter 4.

In Chapter 4 the concept of a power spectral density was introduced as a limiting case of an energy spectral density measured over an interval and divided by that interval [see Eq. (4.17)]. This approach served our purposes for applying deterministic functions, and it can be applied with care to some other situations. In general, however, the approach is not always correct. To be valid, the expected value of the quantity $|F_T(\omega)|^2/T$ should tend toward the true value $S_f(\omega)$ over all sample functions, and its variance should tend toward zero with $T \to \infty$. To remedy this possible defect, we use the following definition for the power spectral density of random waveforms:

$$S_f(\omega) = \lim_{T\to\infty}\left[E\left\{\frac{|F_T(\omega)|^2}{T}\right\}\right]. \tag{8.91}$$

Some implications of using Eq. (8.91) in the computation of power spectral density are discussed in Section 8.14.

All of the relations developed in Chapter 4 for the transmission of power spectra through linear systems can now be applied here. The methods for using random inputs form a very powerful analytical tool in general systems theory.

EXAMPLE 8.13.1

The stationary random process $x(t)$ has a power spectral density $S_x(\omega)$. Find the power spectral density of $y(t) = x(t) - x(t - T)$.

Solution I Using the autocorrelation function, we have

$$\begin{aligned} R_{yy}(\tau) &= E\{[x(t) - x(t - T)][x(t + \tau) - x(t + \tau - T)]\} \\ &= E\{x(t)x(t + \tau)\} - E\{x(t)x(t + \tau - T)\} \\ &\quad - E\{x(t - T)x(t + \tau)\} + E\{x(t - T)x(t + \tau - T)\} \\ &= R_{xx}(\tau) - R_{xx}(\tau - T) - R_{xx}(\tau + T) + R_{xx}(\tau), \\ S_y(\omega) &= 2S_x(\omega) - e^{-j\omega T}S_x(\omega) - e^{j\omega T}S_x(\omega) \\ &= 2(1 - \cos \omega T)S_x(\omega). \end{aligned}$$

Solution II Using the concept of a system transfer function, we can write

$$y(t) = x(t) \circledast h(t),$$

where

$$\begin{aligned} H(\omega) &= \mathcal{F}\{h(t)\} = 1 - e^{-j\omega T}, \\ S_y(\omega) &= S_x(\omega)\,|H(\omega)|^2 \\ &= |1 - e^{-j\omega T}|^2 S_x(\omega) \\ &= (2 - e^{-j\omega T} - e^{j\omega T})S_x(\omega) \\ &= 2(1 - \cos \omega T)S_x(\omega). \end{aligned}$$

The results must, of course, agree. In most cases the former approach is the more powerful, though not always the most direct.

DRILL PROBLEM 8.13.1

White noise with a (two-sided) power spectral density of $\eta/2$ W/Hz is applied to the input of an ideal low-pass filter with a cut-off frequency of B Hz.

a) Find the autocorrelation function of the output of this filter.

b) What is the minimum time interval over which the output becomes uncorrelated?

ANSWER: (a) $\eta B\mathrm{Sa}(2\pi B\tau)$; (b) $\tau = 1/(2B)$ sec.

The autocorrelation function and power spectral density, together with the probability density function, are widely used in both theoretical and applied signal and system analysis. These functions are sketched in Fig. 8.19 for some of the more commonly used waveforms.

EXAMPLE 8.13.2

(Random Binary Waveform) Consider a sample function $x(t)$ shown in Fig. 8.20 of the *random binary waveform* that has the following properties: (a) each pulse is of duration T_b; (b) the two possible pulse levels $-A$ and $+A$ are equally likely, and the presence of $-A$ or $+A$ in any one pulse interval is statistically independent of that in all other intervals; (c) the starting time T of $x(t)$ is uniformly distributed over $(0, T_b)$. Determine the autocorrelation function and the power spectral density of this random waveform.

Solution Let t_1, t_2 be two arbitrary values of time and assume that $0 < t_1 < t_2 < T_b$ and that $|t_1 - t_2| < T_b$. Then $X(t_1)$ and $X(t_2)$ are in the same pulse interval if (see Fig. 8.20)

$$t_2 \le T \le t_1 + T_b.$$

The probability that T is within this range is

$$P(t_2 \le T \le t_1 + T_b) = \int_{t_2}^{T_b + t_1} p_T(\xi)\, d\xi = \frac{T_b + t_1 - t_2}{T_b}.$$

Letting $\tau = t_2 - t_1$, we can rewrite this result as

$$P(t_2 \le T \le t_1 + T_b) = 1 - \frac{\tau}{T_b}.$$

Using absolute magnitude signs to permit the case for negative arguments, we have

$$E\{x(t)x(t+\tau)\} = A^2\left(1 - \frac{|\tau|}{T_b}\right) \qquad |\tau| < T_b.$$

For $|\tau| > T_b$ the random variables $X_t X_{t+\tau}$ are statistically independent (that is, the pulses are in different intervals), and we can write [cf. Eq. (8.87)]

$$E\{x(t)x(t+\tau)\} = m_x^2 = 0 \qquad |\tau| > T_b.$$

Combining the results above, we can write the autocorrelation function as [cf. Eq. (3.49)]

$$R_{xx}(\tau) = A^2\Lambda\left(\frac{\tau}{T_b}\right).$$

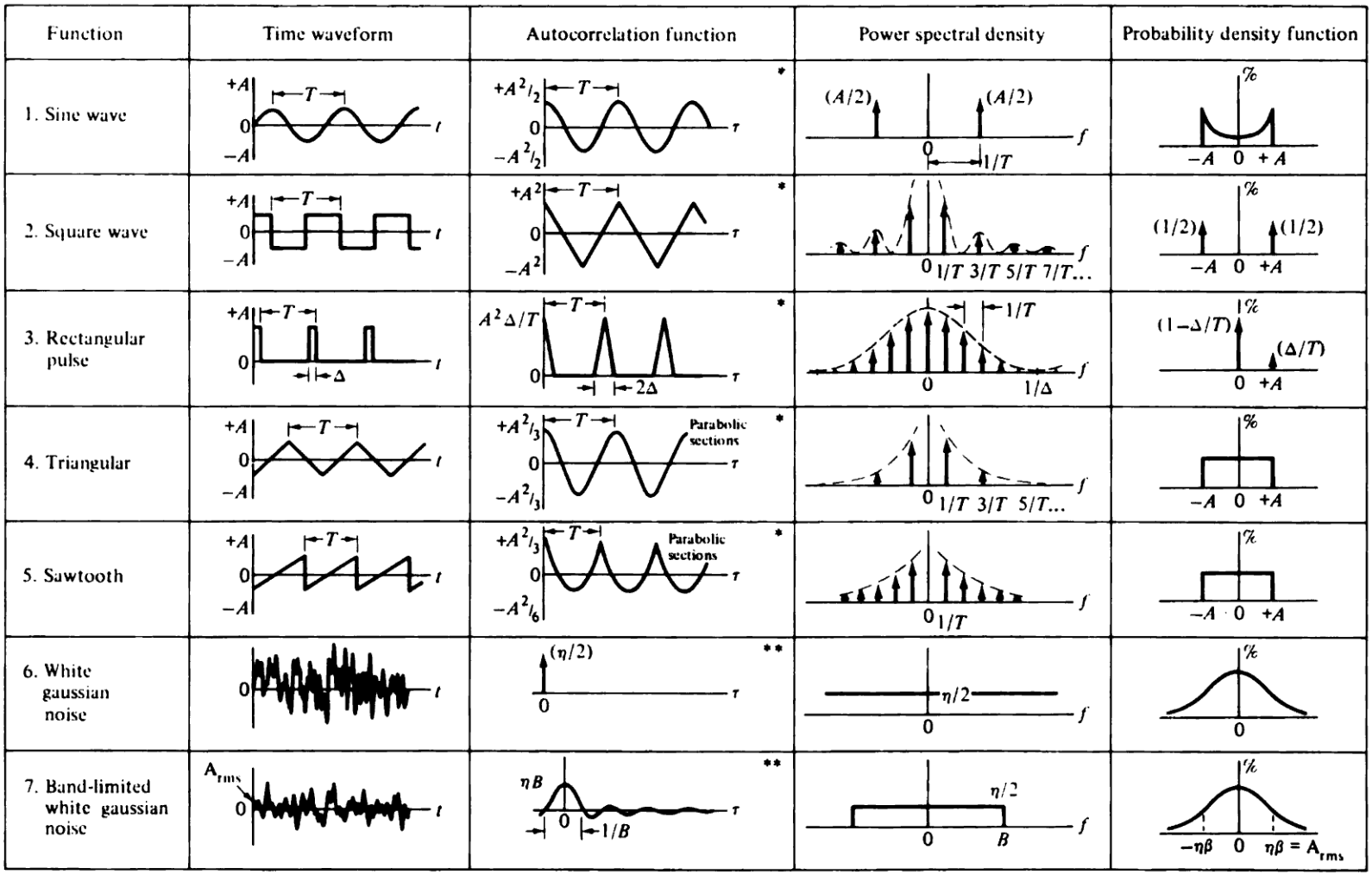


Figure 8.19 Autocorrelation, spectra, and probability densities for some commonly used waveforms (* = time, ** = statistical).

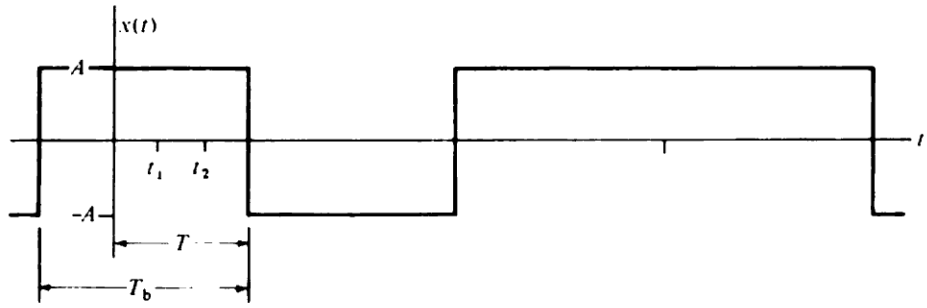


Figure 8.20 A sample function of the random binary waveform in Example 8.13.2.

Using Fourier transform pair #16 in Table 3.1, we find the corresponding power spectral density is

$$S_x(\omega) = A^2T_b[\mathrm{Sa}^2(\omega T_b/2)],$$

which can be written as

$$S_x(\omega) = \frac{|X(\omega)|^2}{T_b}.$$

★ EXAMPLE 8.13.3

A given random process $x(t)$ is applied to the input of an N-stage transversal filter (Fig. 3.14); the output of the filter is designated here by $y(t)$. Assume that $x(t)$ can be written in terms of a pulse waveform $f(t)$ of duration T_b and a stationary sequence of discrete random amplitudes $\{b_k\}$ to form the random pulse train†

$$x(t) = \sum_{k=-\infty}^{\infty} b_k f(t - kT_b - T),$$

where T is a random time delay with a uniform pdf over $(0, T_b)$. The random variable b_k governs the choice of amplitude over the interval $(0, T_b)$. The choice in each interval is statistically independent of the choices in all other intervals. For convenience, we assume that the random variable b_k has zero mean and unit variance. For example, $x(t)$ can be envisioned as a binary random waveform, with b_k equaling $+1$ with probability $\frac{1}{2}$, and -1 with probability $\frac{1}{2}$.

The transversal filter introduces a controlled amount of dependence in $y(t)$ between intervals. In general, this dependence will extend over N adjacent intervals; therefore N is called the *constraint length*. Also, the maximum number of discrete output levels in $y(t)$ is increased by N over the number of levels in $x(t)$. Therefore if

† Note that although the $\{b_k\}$ can be considered to be random variables here, they also represent a random process as a function of the discrete parameter k and, therefore, we use the lowercase notation.

$x(t)$ is a binary random waveform, $y(t)$ is a ternary random waveform for $N = 1$. Dependence between adjacent intervals introduced by the transversal filter alters the power spectral density of $x(t)$, and such filters can be used to alter the spectral density of a random waveform.

Designating the filter output $y(t)$ as

$$y(t) = \sum_{k=-\infty}^{\infty} c_k f(t - kT_b - T),$$

determine

a) the autocorrelation function $R_{cc}(k)$, and

b) the power spectral density of $y(t)$ in terms of the power spectral density of $x(t)$.

Solution a) The c_k's are statistically independent beyond the constraint length of the transversal filter so that

$$R_{cc}(k) = 0 \quad \text{for } k > N.$$

Writing out a few terms for $k \leq N$, we find that

$$\begin{aligned} c_k &= a_0 b_k + a_1 b_{k-1} + a_2 b_{k-2} + \cdots + a_N b_{k-N} \\ &= \sum_{i=0}^{N} a_i b_{k-i}. \end{aligned}$$

Using this result and adopting k for the shift parameter, we find the autocorrelation function is

$$\begin{aligned} R_{cc}(k) &= E\{c_\ell c_{\ell+k}\} = \sum_{i=0}^{N} a_i \sum_{j=0}^{N} a_j E\{b_{\ell-i} b_{\ell+k-j}\} \\ &= \sum_{i=0}^{N-k} a_i a_{i+k}. \end{aligned}$$

Combining the functions above results in our desired result,

$$R_{cc}(k) = \begin{cases} \sum_{i=0}^{N-k} a_i a_{i+k} & k \leq N \\ 0 & k > N \end{cases}.$$

b) Development of the desired result is facilitated by first choosing $N = 2$ and then generalizing for larger N. Noting that what is requested is the magnitude squared of the frequency transfer function of the filter [cf. Eq. (4.21)], we find the most straightforward way to proceed is to determine the impulse response, take a Fourier transform, and then take the magnitude squared. However, we follow a statistical approach here to illustrate the methods developed in this chapter. It is left to the reader to verify that both approaches yield the same result.

For the case of $N = 2$,

$$\begin{aligned} y(t) &= a_0x(t) + a_1x(t - T_b) + a_2x(t - 2T_b) \\ R_{yy}(\tau) &= E\{[a_0x(t) + a_1x(t - T_b) + a_2x(t - 2T_b)] \\ &\quad \times [a_0x(t + \tau) + a_1x(t + \tau - T_b) + a_2x(t + \tau - 2T_b)]\} \\ &= (a_0^2 + a_1^2 + a_2^2)R_{xx}(\tau) + (a_0a_1 + a_1a_2) \\ &\quad \times [R_{xx}(\tau - T_b) + R_{xx}(\tau + T_b)] + (a_0a_2)[R_{xx}(\tau - 2T_b) + R_{xx}(\tau + 2T_b)]. \end{aligned}$$

From part (a) above and $N = 2$,

$$\begin{aligned} R_{cc}(0) &= a_0^2 + a_1^2 + a_2^2, \\ R_{cc}(1) &= a_0a_1 + a_1a_2, \\ R_{cc}(2) &= a_0a_2, \\ R_{cc}(k) &= 0 \quad \text{for } k > 2, \end{aligned}$$

so that

$$\begin{aligned} R_{yy}(\tau) = R_{cc}(0)R_{xx}(\tau) &+ R_{cc}(1)[R_{xx}(\tau - T_b) + R_{xx}(\tau + T_b)] \\ &+ R_{cc}(2)[R_{xx}(\tau - 2T_b) + R_{xx}(\tau + 2T_b)]. \end{aligned}$$

Taking the Fourier transform of both sides and using the delay theorem yield

$$\begin{aligned} S_y(\omega) = R_{cc}(0)S_x(\omega) &+ R_{cc}(1)S_x(\omega)[e^{-j\omega T_b} + e^{j\omega T_b}] \\ &+ R_{cc}(2)S_x(\omega)[e^{-j2\omega T_b} + e^{j2\omega T_b}] \end{aligned}$$

or

$$S_y(\omega) = S_x(\omega)\left[R_{cc}(0) + 2\sum_{k=1}^{2} R_{cc}(k)\cos(k\omega T_b)\right].$$

Generalizing this result, we have

$$S_y(\omega) = S_x(\omega)\left[R_{cc}(0) + 2\sum_{k=1}^{N} R_{cc}(k)\cos(k\omega T_b)\right].$$

Note that the terms within the brackets represent the spectral shaping (filtering) that results from the dependency between N adjacent pulses.

★ **EXAMPLE 8.13.4**

Consider a random binary waveform $x(t)$ that has the following properties: (a) each pulse is of duration T_b; (b) the two possible states in each interval are represented by the waveforms $f_1(t)$, $f_2(t)$ with corresponding Fourier transforms $F_1(\omega)$, $F_2(\omega)$; (c) the probability that $f_1(t)$ is selected in any interval is p and the probability that $f_2(t)$ is selected is $q = (1 - p)$; (d) the choice in any interval is statistically independent of

that in all other intervals. Determine the power spectral density of this random binary waveform.[†]

Solution As a first step, we exhibit the possible periodic content in $x(t)$ as $v(t)$:

$$v(t) = \sum_{n=-\infty}^{\infty} [pf_1(t - nT_b) + qf_2(t - nT_b)].$$

Checking, we observe that $v(t + T_b) = v(t)$ and therefore $v(t)$ is periodic with period T_b. Therefore it can be expressed in terms of a Fourier series whose coefficients are the V_n. The V_n can be expressed in terms of the Fourier transforms $F_1(\omega), F_2(\omega)$ as [cf. Eq. (3.15)]

$$V_n = \frac{1}{T_b}[pF_1(n\omega_b) + qF_2(n\omega_b)],$$

where $\omega_b = 2\pi/T_b$. The resulting power spectral density is [cf. Eq. (4.20)]

$$S_v(\omega) = \frac{1}{T_b^2} \sum_{n=-\infty}^{\infty} |pF_1(n\omega_b) + qF_2(n\omega_b)|^2 2\pi\,\delta(\omega - n\omega_b).$$

To consider the remaining (nonperiodic) content, we truncate the waveform to an interval T and then proceed to use Eq. (8.91). Let the truncated waveform $x_T(t)$ be represented by

$$x_T(t) = \sum_{n=-N}^{N} x_n(t),$$

where

$$x_n(t) = \begin{cases} f_1(t - nT_b) & \text{with prob } p \\ f_2(t - nT_b) & \text{with prob } q \end{cases}.$$

Subtracting out the periodic content, let $u_T(t) = x_T(t) - v_T(t)$, or

$$u_T(t) = \sum_{n=-N}^{N} u_n(t),$$

where

$$u_n(t) = \begin{cases} f_1(t - nT_b) - pf_1(t - nT_b) - qf_2(t - nT_b) & \text{with prob } p \\ f_2(t - nT_b) - pf_1(t - nT_b) - qf_2(t - nT_b) & \text{with prob } q \end{cases}.$$

This can be rewritten in the more compact notation

$$u_n(t) = a_n[f_1(t - nT_b) - f_2(t - nT_b)],$$

[†] W. R. Bennett and J. R. Davey, *Data Transmission*, New York: McGraw-Hill, 1965, pp. 317–319; see also W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Englewood Cliffs: Prentice Hall, Inc., 1973, pp. 17–19. The derivation in this example follows the approach used in Bennett and Davey.

where

$$a_n = \begin{cases} q & \text{with prob } p \\ -p & \text{with prob } q \end{cases}.$$

The $\{a_n\}$ form a sequence of random amplitudes with properties

$$E\{a_n\} = 0; \qquad E\{a_m a_n\} = \begin{cases} 0 & \text{for } m \neq n \\ pq & \text{for } m = n \end{cases}.$$

Allowing an interchange in integration and summation, and using the delay property, we find the Fourier transform of $u_T(t)$ to be

$$U_T(\omega) = \sum_{n=-N}^{N} a_n e^{-jn\omega_b}[F_1(\omega) - F_2(\omega)].$$

Proceeding, we write

$$E\{|U_T(\omega)|^2\} = E\left\{\sum_{m=-N}^{N}\sum_{n=-N}^{N} a_m a_n e^{j(n-m)\omega_b}[F_1(\omega) - F_2(\omega)][F_1^*(\omega) - F_2^*(\omega)]\right\}.$$

To evaluate the expectation, we note that the only variables are the a's, and this expectation is nonzero only for $n = m$, so that

$$\begin{aligned} E\{|U_T(\omega)|^2\} &= \sum_{n=-N}^{N} pq|F_1(\omega) - F_2(\omega)|^2 \\ &= (2N + 1)pq|F_1(\omega) - F_2(\omega)|^2. \end{aligned}$$

Using Eq. (8.91) and noting that $T = (2N + 1)T_b$, we find the power spectral density of $u(t)$ is

$$\begin{aligned} S_u(\omega) &= \lim_{N\to\infty} \frac{(2N + 1)pq|F_1(\omega) - F_2(\omega)|^2}{(2N + 1)T_b} \\ &= pq\frac{|F_1(\omega) - F_2(\omega)|^2}{T_b}. \end{aligned}$$

Combining the results for both nonperiodic and periodic content, we have

$$\begin{aligned} S_v(\omega) &= p(1 - p)\frac{1}{T_b}|F_1(\omega) - F_2(\omega)|^2 \\ &\quad + \frac{2\pi}{T_b^2}\sum_{n=-\infty}^{\infty} |pF_1(n\omega_b) + (1 - p)F_2(n\omega_b)|^2\,\delta(\omega - n\omega_b). \end{aligned}$$

Note that if the Fourier transforms of both $f_1(t)$ and $f_2(t)$ vanish at a harmonic of ω_b, or if the weighted contributions from the Fourier transforms are equal but opposite signs, then no discrete component will appear in the power spectral density at the fre-

quency of that harmonic. For the special case in which $p = \frac{1}{2}$ and $f_1(t) = -f_2(t)$, this result simplifies to

$$S_x(\omega) = \frac{|F(\omega)|^2}{T_b}.$$

8.14 ★ NUMERICAL COMPUTATION OF POWER SPECTRA

In the preceding section the concept of a power spectral density was introduced as the limiting case of the ensemble average of an energy spectral density [Eq. (8.91)]

$$S_f(\omega) = \lim_{T\to\infty}\left(E\left\{\frac{|F_T(\omega)|^2}{T}\right\}\right).$$

In this equation, $S_f(\omega)$ is the power spectral density of the signal $f(t)$ having a Fourier transform $F_T(\omega)$ over $(-T/2, T/2)$. It is not practical to take the limit as $T \to \infty$, however, and we wish to examine some of the consequences of using a finite observation interval. Our emphasis here is on the numerical methods to compute power spectral density.

We use the notation $S_T(\omega)$ for the power spectral density estimate of $f(t)$ based on an observation interval of T units. Thus, for a given sample function, Eq. (8.91) becomes

$$S_T(\omega) = \frac{|F_T(\omega)|^2}{T}. \tag{8.92a}$$

The discrete case is often of interest in which the discrete Fourier transform (DFT) is used to approximate the Fourier transform based on N discrete data samples. Thus N replaces T and $F_N(\omega_n)$ replaces $F_T(\omega)$, where $F_N(\omega_n)$ is the DFT of $f(kT/N)$, $k = 0, 1, 2, \ldots, (N-1)$. With these changes, we can rewrite Eq. (8.92a) as

$$S_N(\omega_n) = \frac{|F_N(\omega_n)|^2}{N}. \tag{8.92b}$$

For convenience we shall treat both the continuous and the discrete cases and it will be obvious from the notation which one is being considered.

The spectral estimate given by Eq. (8.92) is often called a *periodogram* because it was first used to search for possible periodicities in data records. Its use in power spectral estimation from numerical data is becoming more popular as a result of the efficiency and convenience of the FFT algorithm for finding $F_N(\omega_n)$.

Because the data are assumed to arise at least partially from random sources (e.g., additive noise), different sample functions $f(k)$, $k = 0, 1, 2, \ldots, (N-1)$ will result in different values of $S_N(\omega_n)$ for each ω_n and the collection of these values forms an

ensemble. A desirable characteristic in the estimator is that the mean of this ensemble should be the true value at that frequency, i.e., that $E\{S_N(\omega_n)\} = S_f(\omega_n)$. If this is true, then we say that the estimator is *unbiased*.

Using Eq. (8.92) and taking the expected value, we have

$$\begin{aligned} E\{S_T(\omega)\} &= E\left\{\mathcal{F}\left\{\frac{1}{T}\int_{-\infty}^{\infty} f(t)\ \mathrm{rect}\ (t/T) f(t+\tau)\ \mathrm{rect}\ [(t+\tau)/T]\,dt\right\}\right\} \\ &= \mathcal{F}\left\{\frac{1}{T}\int_{-\infty}^{\infty} E\{f(t)f(t+\tau)\}\ \mathrm{rect}\ (t/T)\ \mathrm{rect}\ [(t+\tau)/T]\,dt\right\} \\ &= \mathcal{F}\left\{R_{ff}(\tau)\frac{1}{T}\int_{-\infty}^{\infty} \mathrm{rect}\ (t/T)\ \mathrm{rect}\ [(t+\tau)/T]\,dt\right\} \\ &= \mathcal{F}\{R_{ff}(\tau)[1 - |\tau|/T]\} \\ &= \mathcal{F}\{R_{ff}(\tau)\Lambda(\tau/T)\}. \end{aligned} \tag{8.93a}$$

In the discrete variable notation, this can be rewritten as

$$E\{S_N(\omega_n)\} = \mathrm{DFT}\{R_{ff}(k)\Lambda(k/N)\}. \tag{8.93b}$$

The second term in the braces in Eq. (8.93) arises as a result of the data truncation and is called a *window* function. As a result of the presence of the window function, the spectral estimator $S_N(\omega_n)$ provides a *biased* estimate of $S_f(\omega)$. Note that as $N \to \infty$ $(T \to \infty)$, $S_N(\omega_n)\,[S_T(\omega)]$ becomes *asymptotically unbiased*.

The effects of the window function are not unlike those discussed in Chapter 3 for the DFT and can be exhibited by using the frequency convolution property in Eq. (8.93):

$$E\{S_T(\omega)\} = S_f(\omega) \circledast T\ \mathrm{Sa}^2\ (\omega T/2). \tag{8.94}$$

The window function limits the frequency resolution to $1/T$ (or $1/N$) and also has some effect on the bias in the spectral estimate.

Another desirable property of an estimator is that the variance of the estimate should become small as T (or N) becomes large. An estimator with this property is said to be *consistent*. To derive a relation for the variance of the spectral estimator involves fourth-order moments and it is difficult to obtain a general result. If the data consists of sample values of a gaussian random process it can be shown that the standard deviation of the spectral-density estimate is on the order of that of the estimate and is not dependent on N.[†] This conclusion also holds approximately for nongaussian data and we conclude that the estimator described by Eq. (8.92) produces an *inconsistent* estimate of the power spectral density.

In practice, the variance of the spectral estimator can be reduced by smoothing. One way is to smooth over an ensemble. Given N data samples, we can use $N' = N/K$ data points for each of K periodograms and then the periodograms are av-

[†] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, Second ed. New York: Wiley-Interscience, 1986.

eraged for each frequency. Assuming statistical independence, the variance decreases by the factor K (cf. Example 8.9.2). Note that the frequency resolution has also been decreased by the same factor (that is, $1/N'$ instead of $1/N$).

A second way is to smooth over frequency. In this method a periodogram is computed from the N data samples and then several spectral estimates adjacent to each other in frequency are averaged. This is analogous to low-pass filtering on the periodogram and decreases the variance at the expense of decreased frequency resolution.

In practice, the proper choices for the computation of a power spectral density can be made more easily if some information is known about the characteristics of the spectral density. If little is known, a few trial spectral plots should be formed before the best compromises can be made as to what windowing, averaging, and data record lengths should be used.

DRILL PROBLEM 8.14.1

A possible estimate of the autocorrelation function $R_{ff}(\tau)$ based on finite observation data over $(-T/2, T/2)$ is

$$\hat{R}_{ff}(\tau) = \frac{1}{T}\int_{-T/2}^{T/2} f(t)f(t+\tau)\,dt\,.$$

(a) Is $\hat{R}_{ff}(\tau)$ an unbiased estimator of $R_{ff}(\tau)$? (b) Compare $\hat{R}_{ff}(\tau)$ and $R_{ff}(\tau)$ used in Eq. (8.93).

ANSWER: (a) Yes; (b) In $\hat{R}_{ff}(\tau)$ the overlap of $f(t)$ and $f(t+\tau)$ is truncated to $(-T/2, T/2)$; in Eq. (8.93) *each function* is truncated to this interval.

8.15 SUMMARY

Probability is a way to measure the number of favorable outcomes of an experiment relative to the total number of possible outcomes. The probability of an outcome is positive, real-valued, and bounded between 0 and 1. If the occurrence of one outcome precludes the occurrence of another, the two outcomes are mutually exclusive. When the outcomes are mutually exclusive, their probabilities are additive.

The probability of a given outcome based on knowledge of a second outcome is called a conditional probability. Two outcomes are statistically independent if the conditional probability is equal to the probability of occurrence. The joint probability of several outcomes is the probability of their joint occurrence. The joint probability of statistically independent outcomes is equal to the product of the probabilities of the individual outcomes.

A random variable is a rule (mapping) that assigns a number to each possible outcome. A random variable may be continuous or discrete.

The cumulative distribution function of a random variable is the probability of the random variable being less than a specified value (which then forms the argument) for a given experiment. The probability density function (pdf) is the derivative of the cumulative distribution function.

An expected value is a statistical average with a pdf weighting function. The expected value of X^n is called the nth moment of the random variable X. The first and second moments are called the mean and the mean-square value, respectively. The second moment about the mean is called the variance, and its square root is the standard deviation.

Examples of continuous random variables include the uniform and gaussian distributions. The binomial and Poisson distributions are examples of discrete random variables.

A histogram is a tabulation of the frequency of occurrence of a random variable within preset ranges. The histogram can be used to approximate the pdf of the random variable. The transformation of a random variable through a system can be handled as a mapping of the pdf using the differential slope of the system output-input gain characteristic.

Bivariate probability distributions describe two-dimensional random variables. The statistical dependence of two random variables is measured by the correlation between them. If the correlation coefficient between two random variables is zero, the random variables are said to be uncorrelated.

A random process is a collection or "ensemble" of random variables arranged in a time progression. Each member waveform versus time is called a sample function. The random process is stationary if its statistics do not vary with time. It is ergodic if all its statistics can be computed from a single sample function just as well as from the ensemble.

A joint moment is the expected value of the joint occurrence of two or more outcomes. The autocorrelation function is the second joint moment of two random variables. For a random process, these random variables are specified at two instants of time. If the random process is stationary, only the time difference between them is important.

The power spectral density of a stationary random process is given by the Fourier transform of the autocorrelation function. It may be estimated numerically by computing the squared magnitude of the discrete Fourier transform and dividing by the number of data samples.

Selected References for Further Reading

1. P. Beckmann. *Elements of Applied Probability Theory*. New York: Harcourt, Brace & World, 1968.
   A good general introduction to probability, random variables, and random processes with a number of interesting examples.
2. P. Z. Peebles. *Probability, Random Variables, and Random Signal Principles*. New York: McGraw-Hill, 1980.
   A good introductory treatment of probability, random variables, and random processes, with a large number of problems.

3. G. R. Cooper and C. D. McGillem. *Probabilistic Methods of Signal and System Analysis*. Second ed. New York: Holt, Rinehart & Winston, 1986.
An introduction to probability written for undergraduate engineering students, with applications in systems and filtering; additional references are given with comments at the end of the first chapter.

4. A. Leon-Garcia. *Probability and Random Processes for Electrical Engineering*. Reading, Mass.: Addison-Wesley, 1989.
Written for undergraduate engineering students, with applications to engineering problems and computer methods. The first chapter is interesting reading for students interested in communication systems.

5. C. W. Helstrom. *Probability and Stochastic Processes for Engineers*. New York: Macmillan, 1984.
A good introductory treatment of probability and random variables, with added emphasis on pairs of random variables; applications to engineering and physics.

6. J. S. Bendat and A. G. Piersol. *Random Data: Analysis and Measurement Procedures*. Second ed. New York: John Wiley & Sons, 1986.
A discussion of requirements and limitations of random data acquisition and measurements, as applied to physical problems.

7. H. Stark and J. W. Woods. *Probability, Random Processes, and Estimation Theory for Engineers*. Englewood Cliffs, N.J.: Prentice Hall, 1986.
A well-written senior or graduate text on probability, with some engineering applications, and an emphasis on detection and estimation.

8. A. Papoulis. *Probability, Random Variables and Stochastic Processes*. Second ed. New York: McGraw-Hill, 1984.
A widely used graduate text on probability, with some engineering applications. Chapter 1 has an interesting and easily read discussion of the various concepts of probability.

PROBLEMS

8.1.1 A card is drawn from a deck of 52 cards. What is the probability that the card is a) a diamond? b) a 2? c) the 2 of diamonds? d) a 2 or a 4 or a 6?

8.1.2 What is the probability that the last digit of a telephone number is a) odd? b) 5? c) 5 or 6? d) either odd or divisible by 4? (Assume all digits are equally likely.)

8.1.3 An honest coin is defined as one for which a head or a tail is equally likely.

a) Three honest coins are tossed. What is the probability of two heads and one tail?
b) Four honest coins are tossed. What is the probability that exactly one-half of them will show heads?

8.1.4 A box filled with black and white marbles is known to contain twice as many white marbles as black marbles. Assume they are mixed thoroughly and are indistinguishable except for color. Suppose that a certain number of marbles are drawn from the box without replacement and that all of them are found to be black. If the probability of drawing one additional marble is exactly 10% less than if the black marbles had been replaced, determine the minimum number of marbles that were in the box originally.

8.2.1 A certain remotely operated transmitter is programmed to shut down if both events A and B (e.g., voltages out of tolerance) occur simultaneously, or if event C occurs. The

probability that A will occur is 0.001, the probability of B occurring is 0.002, and the probability of C occurring is 1×10^{-5}.

a) What is the probability of transmitter shutdown for A, B statistically independent?
b) If $P(A|B) = 0.01$, are the events A and B statistically independent?
c) What is the value of $P(B|A)$?

8.2.2 An electronic part is purchased from three different vendors. Records show that average rates of defective parts are 0.02% from vendor A, 0.013% from vendor B, and 0.01% from vendor C. Of 50,000 parts purchased, 10,000 were from vendor A, 15,000 from vendor B, and 25,000 from vendor C. The inventories of parts are combined and used in production.

a) What is the probability that a given part in the combined inventory, selected at random, is defective?
b) What is the probability that a defective part was purchased from vendor A?

8.2.3 Find the probability of hitting an aircraft if three missiles are fired in succession, each having a probability of 1/3 of scoring a hit. Assume that the aircraft goes down when hit. Do this problem using two differing approaches, starting each as follows:

a) Find the probability of three misses, noting that the misses are statistically independent.
b) Write an expression, in terms of a chain of conditional probabilities, for the total probability of a hit.

8.2.4 A certain binary PCM system transmits the two binary states $+1$, -1 with equal probabilities (e.g., using phase modulation of a sinusoid). However, as a result of noise the receiver makes some errors in recognition; suppose that $P(+1|-1) = 0.10$ and $P(-1|+1) = 0.20$. Another possibility is that as a result of path distortion the receiver may lose necessary signal strength to make any decision. Thus, we identify three possible receiver states: $+1, 0, -1$, where 0 corresponds to "loss of signal." For the last condition, we let $P(0|+1) = P(0|-1) = p$, where p is the probability of an outage.

a) What are the probabilities of 1's and 0's at the receiver for $p = 0$?
b) If the receiver makes a decision for $+1$, and $p = 0$, what is the probability that a $+1$ was actually sent?
c) Repeat part (b) for the case $p \neq 0$.

8.3.1 Consider the experiment that consists in tossing of four honest coins; assign 1 to a "head," 0 to a "tail," and then add to form the number N. The random variable is defined by assigning $X = \log(1 + N^3)$.

a) Determine and plot the cumulative distribution function of X over $(0, 2)$.
b) From your result, determine $P(X \leq 1)$, and indicate it on your plot.

8.3.2 The cumulative distribution function for a certain random variable X is

$$F_X(x) = \begin{Bmatrix} 0 & x < -\pi \\ [1 + \sin(x/2)]/2 & -\pi \leq x < \pi \\ 1 & \pi < x \end{Bmatrix}.$$

Determine: (a) $P(X \leq 2)$; (b) $P(-1 \leq X < +1)$; (c) $P(1 \leq X < 3)$; (d) The minimum and the maximum values of X.

8.3.3 A random number generator produces the equally probable digits $N = 0, 1, 2, \ldots, 9$. Determine and plot the cumulative distribution function of

a) The random variable X defined by $X = -1$ if $N \leq 3$; $X = +1$ if $N \geq 7$; $X = 0$ otherwise.
b) The random variable Y defined by $Y = +1$ if $N \leq 3$; $Y = -1$ if $N \geq 7$; $Y = 0$ otherwise.
c) The random variable Z = integer part of $\{N/2\}$.

8.4.1 A certain random variable X has the following cumulative distribution function:

$$F_X(x) = x^2 u(x) + (1 - x^2)u(x - 1).$$

a) Find $P(1/4 \leq X)$.
b) Find $P(1/4 \leq X < 3/4)$.
c) Determine and sketch the pdf of X.

8.4.2 A random variable X has the following pdf:

$$p_X(x) = \begin{Bmatrix} kx(1 - x) & 0 \leq x < 1 \\ 0 & \text{elsewhere} \end{Bmatrix}.$$

a) Find the numerical value of k.
b) Find $P(1/4 \leq X)$.
c) Determine and sketch the cumulative distribution function of X.

8.4.3 A random variable X has the following pdf:

$$p_X(x) = k(1 - x^2) \operatorname{rect}(x/2).$$

a) Find the numerical value of k.
b) Find $P(-1/2 \leq X < 1/2)$.
c) Determine and sketch the cumulative distribution function of X.

8.4.4 The Rayleigh pdf is given by

$$p_X(x) = kx \exp[-x^2/(2\sigma^2)]u(x).$$

a) Find $P(\sigma \leq X)$ and $P(\sigma \leq X < 2\sigma)$.
b) Determine the cumulative distribution function of X.
c) Find the constant a such that $P(X > a) = 0.00033$.

8.4.5 The pdf of the amplitude X of a given signal $x(t)$ is given by $p_X(x) = k \exp(-|x|)$.

a) Sketch the pdf of X.
b) Determine and sketch the cumulative distribution function of X.
c) The signal $x(t)$ is applied to the input of an amplifier with unit gain; the corresponding output is $y(t)$. However, this amplifier cannot accommodate the (occasional) large peaks in the input waveform and saturates ("clips") at levels $Y = \pm A$. Determine and sketch the pdf and the cumulative distribution function of the output, Y.

8.5.1 Determine the mode, the median, and the mean value of the random variable whose pdf is given in

a) Problem 8.4.2;
b) Problem 8.3.2;
c) Problem 8.4.4.

8.5.2 Consider the triangular pdf shown in Fig. P–8.5.2.

a) Find $P(X > a/2)$.

b) Determine σ_X if $a = 1$.

c) Assume that $a = 1$, and divide the range of X into eight equal subranges (i.e., $0 \le X < 0.25$, $0.25 \le X < 0.50$, etc.). Replace each area defined by these subranges and $p_X(x)$ by an impulse function with the same area as the region it replaces, and at the midpoint of the subrange. This procedure of replacement of a continuous pdf with a discrete pdf is called "quantization." Note that quantization requires a definition of the subranges of X, and the points at which the discrete representations are to be placed. Sketch the resulting discrete pdf for this case.

d) Compute the standard deviation using the result of part (c), and compare with the standard deviation in part (b).

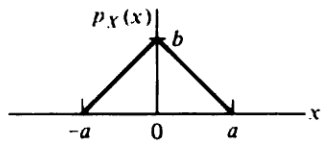


Figure P–8.5.2.

8.5.3 The pdf of a given random variable is $p_X(x) = 2 \exp(-2x)u(x)$.

a) Find the mean and the variance of X.

b) The random variable Y is defined by the transformation $Y = 2X + 5$. Find the mean and the variance of Y.

8.5.4 a) Find the first and second moments of the random variable X whose pdf is shown in Fig. P–8.5.4.

b) Determine the maximum value of the ratio $\gamma = \sigma_X/m_X$, and find the corresponding value of ρ.

Figure P–8.5.4.

8.5.5 The signal $x(t)$ in Problem 8.4.5 is applied to a half-wave rectifier, so that the output $y(t)$ is defined by

$$y(t) = \begin{Bmatrix} x(t) & X > 0 \\ 0 & X \le 0 \end{Bmatrix}.$$

a) Find and sketch the cumulative distribution function of Y.

b) Find and sketch the pdf of Y.

c) Find the mean and the variance of Y.

8.5.6 Sometimes the methods developed in probability can be adapted to advantage in linear systems. For example, the mean is a measure of the 50% point in the cumulative distribution function, and the standard deviation is a measure of the effective width of a pdf. Show that, by a proper analogy, the delay time t_0 and the rise time t_r of a linear time-invariant system with impusle response $h(t)$ can be written as [cf. Eq. (3.80)]:

$$t_0 = \frac{\int_{-\infty}^{\infty} t|h(t)|^2\,dt}{\int_{-\infty}^{\infty} |h(t)|^2\,dt}, \qquad t_r = \left[\frac{\int_{-\infty}^{\infty} (t - t_0)^2|h(t)|^2\,dt}{\int_{-\infty}^{\infty} |h(t)|^2\,dt}\right]^{1/2}.$$

Explain this analogy to the pdf, and to the cumulative distribution function.

8.6.1 The periodic waveform $x(t)$ shown in Fig. P–8.6.1 has a random starting time with a uniform distribution.

a) Sketch the cumulative distribution function and the pdf of the amplitude X.
b) Compute the mean and the variance of X.
c) Compute the rms value of $x(t)$, both from the pdf and from the time averages [cf. Eq. (4.24)].

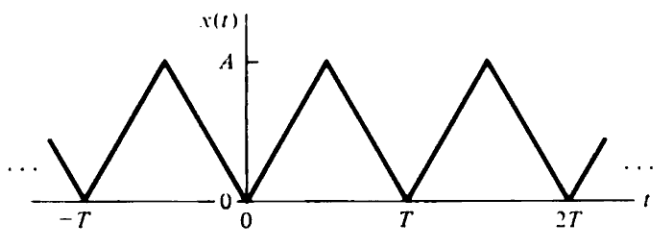


Figure P–8.6.1.

8.6.2 a) Using formulas for the partial sums given in Appendix A, determine the mean and the variance of the uniform discrete random variable defined for $x_i = 1, 2, \ldots, n$.
b) Repeat part (a), with the exception that the random variable is defined for $x_i = a, 2a, \ldots, na$.

8.6.3 a) Find the mean and the variance of a discrete random variable X having two possible values a, b with probability $P(X = a) = p$.
b) Under what conditions does your result check with that given for the binomial distribution?

8.6.4 a) Using the binomial distribution, find the probability that there will be exactly 50% heads when four honest coins are tossed.
b) Repeat part (a), with the exception that 16 coins are tossed.
c) Repeat part (a), with the exception that 64 coins are tossed.
d) Find $P(|X - m_X| \le \sigma_X)$ for each of the three cases above.

8.6.5 a) Write a pdf for a random variable X defined on the integers $x_i = 0, 1, 2, \ldots, n$ and with the Poisson distribution given in Eq. (8.49).
b) Show that the pdf in part (a) has unit area.
c) Show that the mean value of X is equal to ν.

8.6.6 Compare the Poisson approximation to the binomial probabilities for $\nu = 1$ and $i = 0, 1, 2, 3$ for each of the following cases: (a) $n = 10$; (b) $n = 100$.

8.6.7 The average time that a student must wait at a window for registration information is found to be 4 min.

a) What is the probability that a student selected at a random arrival will have to wait exactly 4 min? (Assume a Poisson random variable.)
b) What is the probability that the student in (a) will have to wait $4 \pm 1\sigma$ min?

8.6.8 A certain data transmission system has an average error probability of 1×10^{-6} per digit. Determine the probability that in a message of 2×10^6 digits there will be

a) exactly two errors;
b) less than two errors;
c) more than two errors.

8.6.9 Determine the constant k and the mean of the gaussian random variable X whose pdf is

$$p_X(x) = \{\exp[-(x-1)^2/3]\}/\sqrt{15} + k\,\delta(x).$$

[*Hint:* Put into standard form and then find answers without integration.]

8.6.10 A guideline for acceptable reception of baseband PPM is that the probability of the noise exceeding the pulse amplitude level (with no pulse present) should not exceed 1%.

a) Assuming that the pulses go from 0 to A volts and that the noise is gaussian with zero mean, determine the required peak-signal–to–mean-square noise (power) ratio in decibels.
b) Does this result agree with those in Chapter 7 [cf. Eqs. (7.22), (7.23)]?
c) What increase in peak signal power (in dB) is necessary to reduce the probability of the noise exceeding the pulse amplitude level, when compared with that in part (a), by one order of magnitude (i.e., to 0.1%)?

8.6.11 The output noise from a high-gain amplifier is gaussian-distributed with zero mean, and dissipates an averge power of 100 mW in a 50-Ω resistive load.

a) What is the probability of observing the output noise amplitude less than -1 V?
b) What is the probability of observing the output noise amplitude greater than $+1$ V?
c) Repeat parts (a) and (b) if the average value of the output noise is 1 V.

8.6.12 In the fall in the northern United States, a manufacturer of snowblowers makes the following advertised offer.† If you purchase a snowblower and then it snows less than 20% of average that winter you get a full refund of what you paid (and keep the snowblower). If it snows less than 50% of average you get 50% of your money back. (If it snows more than that you'll be glad you purchased the snowblower!) Assume here that the amount of snow is a random variable X which may be approximated by a gaussian distribution with mean m_X. Also assume that production costs on snowblowers are 50%, distribution costs are 40%, and usual profit margins are 10% of the selling price of snowblowers. Determine the expected percent profit to the company in the above conditions in this promotion if

a) $\sigma_X = 0.2m_X$;
b) $\sigma_X = 0.3m_X$.

† Based on an advertisement for the Toro Snowthrower in the fall of 1983.

8.6.13 A characteristic of importance in linear amplifiers for power output (e.g., in audio systems) is how well they can respond to sudden peaks in program material and yet not "clip" the output waveform. Assume here that the average power required of such an amplifier is 5 W. We are interested in finding the peak (instantaneous) power required if the signal waveform is clipped, on the average, not more than 0.01% of the time. For example, the peak power required to meet this condition for a sinusoidal input would be $(5)(\sqrt{2})^2 = 10$ W. Determine the peak power requirement if the input waveform distribution can be described by a random variable with zero mean and is

a) uniformly distributed;
b) gaussian;
c) Laplacian [i.e., $p_X(x) = k \exp(-\alpha|x|)$, $k = \alpha/2$].

√★ **8.7.1**† The subroutine library for most digital computers contains a random number generator that can be set to produce the numbers x_i uniformly distributed over (0, 1). Compute and plot a histogram over (0, 1) for this random number generator. Use 20 equally spaced intervals and generate a minimum of 2500 samples (x_i).

√★ **8.7.2**† For the x_i in Problem 8.7.1, let the random variable Y be defined by $y_i = x_i + x_{i-1}$. Plot the histogram of Y over (0, 2). This plot should resemble a triangular pdf (for the reason, see Problem 8.9.3.). Also, compute the mean and the variance of Y from your data.

√★ **8.7.3**† For the x_i in Problem 8.7.1, let the random variable Y be defined by

$$y_i = 1/(\sqrt{12N}) \sum_{n=0}^{N-1} (x_{i-n} - 0.5).$$

Compute the y_i for $N = 10$ and plot as a histogram over $(-0.25, 0.25)$. This plot should begin to resemble the pdf of a gaussian random variable, even with these limited computing resources. [You may wish to refer to the note in Problem 8.7.2 for $N = 1$ as a check in setting up the program for this problem.] Also, compute the mean and the variance of Y from your data.

★ **8.8.1** The random variable X representing a certain waveform $x(t)$ has a pdf as shown in Fig. P–8.8.1(a). This waveform is the input to an amplifier whose output-input gain characteristic is shown in Fig. P–8.8.1(b).

a) Letting Y be the amplifier output, find and sketch the pdf of Y.
b) Determine the rms value of X and the rms value of Y. Are they related linearly by the amplifier gain? Explain.

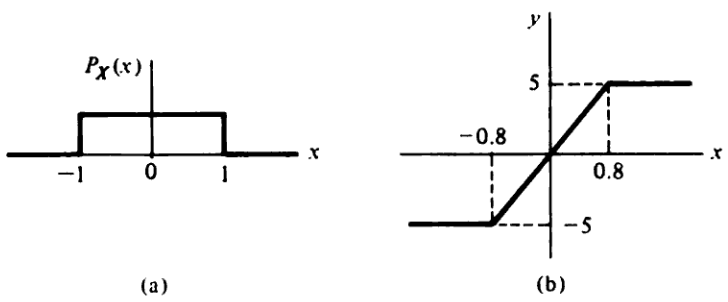


Figure P–8.8.1.

† More sample points may be run for better results in these problems, depending on the speed and memory size of your computer.

★ **8.8.2** A random variable X, uniformly distributed over $(0, 1)$, is applied to the input of a system that has an output-input gain characteristic: $y = \sin(\pi x/2)$. Determine and sketch the pdf of the output Y.

√★ **8.8.3** a) For the x_i in Problem 8.7.1, let $y_i = \sin(\pi x_i/2)$; compute and plot the y_i as a histogram over $(0, 1)$.
b) Compare this plot with the results of Problem 8.8.2.

★ **8.8.4** Determine and sketch the pdf of $y = A \cos x$, where X is uniformly distributed over $(-\pi, \pi)$. [*Hint:* Consider the ranges $(-\pi, 0)$ and $(0, \pi)$ separately.]

8.9.1 The joint pdf of the random variables X, Y is given as

$$p_{XY}(x, y) = k \exp[-(x + 2y)]u(x)u(y).$$

a) Are X, Y statistically independent? Explain.
b) Determine the value of the constant k.
c) Find $P(X < 2, Y < 2)$.
d) Determine $p_X(x)$ and $p_Y(y)$.

8.9.2 The joint pdf of the random variables X, Y is given as

$$p_{XY}(x, y) = x \exp[-x(1 + y)]u(x)u(y).$$

a) Find $p_X(x)$ and $p_Y(y)$.
b) Are X, Y statistically independent? Explain.
c) Find $P(X < 1, Y < 1)$.
d) Determine the pdf of Y given that $X = 1$.

8.9.3 In some situations, theory developed for linear time-invariant systems can be adapted to advantage in probability. For example, suppose that for the random variable X we identify $E\{e^{j\omega x}\}$ as the Fourier transform of its pdf. Let the random variable Z be the sum of the random variables X, Y with known joint pdf. Show that the pdf of Z is given by the convolution of the pdf of X with the pdf of Y [i.e., $p_Z(x) = p_X(x) \circledast p_Y(x)$] if X, Y are statistically independent.

8.9.4 Use the result of Problem 8.9.3 and repeated convolution to generate the pdf for the binomial distribution.

8.9.5 Let X be a random variable that takes on the values $\pm A$ with equal probabilities. Let Y be a gaussian random variable with zero mean and variance σ^2. Assume that X and Y are statistically independent, and $A > \sigma$.

a) Using the result of Problem 8.9.3, sketch the pdf of $Z = X + Y$.
b) The random variable Z is available to an observer to make an estimate if $X = A$ or $X = -A$ is present. The rule adopted is that if $Z > 0$ then the decision is that $+A$ is present and if $Z < 0$ then $-A$ is present. What is the probability of the observer making an error?

★ **8.10.1** Show that it is a sufficient condition that X, Y are uncorrelated in Example 8.9.1.

★ **8.10.2** Find the correlation of X, Y in a) Problem 8.9.1; b) Problem 8.9.2.

√★ **8.10.3** Suppose that in the exam in Example 8.5.3 there are two problems, equally weighted. A tabulation of the scores, identified by the random variables X, Y, Z, is given below.

N	Total Z	Problem 1 X	Problem 2 Y
1	33	10	23
2	47	20	27
3	58	18	40
4	67	30	37
5	75	35	40
6	82	38	44
7	88	38	50
8	94	49	45
9	100	50	50

a) Compute the mean and the standard deviation of X, Y, Z. Does the equality $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$ hold here? Explain.
b) Compute the correlation coefficient and the linear regression coefficients for (X, Z) and (Y, Z).
c) For student no. 5, find the best estimate of total score based only on a knowledge of (1) the score on problem 1; (2) the score on problem 2.
d) Based only on statistics, which do you believe is the better problem in this exam? Why?

√★ **8.10.4** a) Compute the correlation coefficient of successive sample values of the fractional part of $\sqrt{N}$ for $1 \leq N \leq 1000$.
b) Repeat for the fractional part of $10\sqrt{N}$. Does the removal of the first decimal make the successive values more uncorrelated?
c) What would you expect if $\sqrt{N}$ were used? (You may wish to try it!)

★ **8.11.1** Find an expression for $p(x \mid y)$ for the special case in which $\rho = 0$ in Eq. (8.78). Does $\rho = 0$ imply statistical independence for the jointly gaussian distribution? Explain.

★ **8.11.2** a) Write out the bivariate gaussian pdf, in as simplified a form as possible, for $m_X = m_Y = 1$; $\sigma_X = \sigma_Y = 1$; $E\{XY\} = 1 + \sqrt{3}/2$.
b) Write out $p_X(x)$ and $p_Y(y)$ for this case.

★ **8.11.3** Show how Eq. (8.79) follows from Eqs. (8.65) and (8.78).

8.12.1 Sunspot activity varies over time intervals of several hours to years, and affects high-frequency (HF) communications using skywave propagation in the ionosphere. Major sunspot activity is cyclic, with a full cycle on the order of 11.1 years. Magnetic storms associated with solar activity are cyclic in 27 day cycles, the rotation period of the sun. Would you classify these effects of solar activity on ionospheric propagation as stationary random processes over a) a minute; b) a week? Explain.

8.12.2 A gaussian random variable X in a measurement system has variance σ^2 and a mean value that drifts slowly in proportion to a thermal time constant τ. As a measure of stability, it is decided that the probabilities $P(X > k\sigma)$, $k = 1, 2, 3$, are to vary no more than 1% from their nominal values during a measurement duration Δt in order to make a valid measurement. Determine the maximum value of Δt, in terms of τ, that may be used.

8.12.3 In the random process $z(t) = X \cos \omega_c t - Y \sin \omega_c t$, X and Y are independent gaussian random variables each with zero mean and variance σ^2.

a) Determine m_Z and σ_Z^2.
b) Is your result valid if it is only known that X, Y are uncorrelated and have zero means and equal variances?
c) Show that this satisfies Eq. (5.66) if the random process describing $n(t)$ is ergodic.

8.12.4 Draw a block diagram of a system that will implement the operations depicted in Fig. 8.15 if $x(t)$ is ergodic and if only one threshold device is permitted in the design.

8.13.1 a) Find $R_{zz}(\tau)$ for Problem 8.12.3.
b) Show that the time average of $R_{zz}(\tau)$ for $\sigma_X \neq \sigma_Y$ yields the result given in Eq. (5.67).

8.13.2 Let $x(t)$ be a sample function from a real-valued stationary random process with autocorrelation function $R_{xx}(\tau) = \Lambda(\tau)$.

a) Determine and sketch the power spectral density of $x(t)$.
b) Determine the autocorrelation function and the power spectral density of $y(t) = x(t)e^{j\omega_c t}$.
c) Sketch the real part of the autocorrelation function determined in part (b).

8.13.3 A binary communication system transmits equally probable ones and zeros using the signal A over $(0, T_b)$ for a 1 and 0 over $(0, T_b)$ for a 0.

a) Determine the power spectral density of the transmitted waveform.
b) Determine the power spectral density of this binary transmission in the presence of the multipath described in Problem 3.10.2. (Assume that $\tau < T_b$.)

★ **8.13.4** A random binary source generating equally probable 1's and 0's with levels A, $-A$ over $(0, T_b)$ is filtered by the transversal filter in Problem 3.10.1. Determine the power spectral density of the output waveform by (a) first finding the autocorrelation function, and then taking a Fourier transform; and (b) first finding the system transfer function, and then multiplying the input power spectral density by the magnitude-squared transfer function.

★ **8.13.5** A binary communication system transmits equally probable 1's and 0's using the waveform $f_1(t) = \cos(\pi t/T_b)$ over $(-T_b/2, T_b/2)$ for a 1 and $f_2(t) = -f_1(t)$ for a 0.

a) Determine the power spectral density of the transmitted waveform.
b) Repeat part (a) if $f_2(t) = 0$.

★ **8.13.6** (Least-Squares Filtering) A linear time-invariant filter with impulse response $h(t)$ and frequency transfer function $H(\omega)$ is used to filter $y(t) = x(t) + n(t)$. The signal $x(t)$ and the additive noise $n(t)$ are sample functions of ergodic random processes with zero means; assume here that they are uncorrelated. The error $\epsilon(t)$ in filtering is taken as the difference between the filter output and the signal $x(t)$; i.e., $\epsilon(t) = x(t) - h(t) \circledast y(t)$. Determine the minimum value of $R_\epsilon(0)$ if $H(\omega)$ can be chosen to advantage, without any realizability constraints.

√★ **8.14.1** a) Compute a 256-point power spectral density for the uniformly distributed random number generator in your computer library, after removing the mean value, and plot your results over 0–127.
b) Perform frequency smoothing by taking each point in the computed power spectral density and replacing it with the arithmetic average of three successive points. Repeat this smoothing operation three times and plot the result to the same scale as in part (a).

c) Is the power spectral density of your random number generator approximately "white" over this interval?

√★ **8.14.2** a) Repeat Problem 8.14.1 for the random variable Y, where X is the zero-mean random variable produced by the random number generator and Y is defined by

$$y_i = (\sigma_X/2) \sin (100\pi i/N) + x_i .$$

b) Is identification of the sinusoidal signal enhanced by the method of frequency smoothing used? Explain.

√★ **8.14.3** a) Repeat Problem 8.14.1(a) for the random number generator in Drill Problem 8.7.1, but using 128 points.

b) Perform time smoothing by taking each point from the random number generator and replacing it with the arithmetic average of three successive points. Repeat this operation twice, take the power spectral density, and plot the result to the same scale as in part (a). Explain your results.

DESIGN PROBLEMS

Design Problem

8.1 Accuracy of Remote Sensor Location by Satellite

In the 1970s, a series of instrumented packages were attached to superpressure balloons (referred to here as "platforms"), which then drifted with the winds. These instrumented packages sent random transmissions to the Nimbus F satellite, which circled the earth in a 108-min polar orbit. Some specifications of this system are described by J. L. Coates, "The Nimbus F Random Access Measurement System (RAMS)," *IEEE Transactions on Geoscience Electronics*, vol. GE-13, no. 1 (January 1975): 18–26.

In this experiment, each platform transmits for 1 sec of every minute. The first portion of this transmission consists of unmodulated carrier (401.2 MHz). Because the satellite orbit parameters are known, the doppler shift on this unmodulated carrier can be used for platform location. The doppler increases as the satellite approaches that plane, normal to its trajectory, in which the platform is located, then decreases as it leaves. The rate of doppler shift near this plane is approximately linear (you can use a binomial series approximation to show this), with slope dependent on the distance between platform and spacecraft. Thus the platform can be located, within an ambiguity of which side of the satellite trajectory. This ambiguity can be resolved by directive antennas, effects of earth rotation, or by comparison of records on successive satellite passes. The last choices were used in the RAMS experiment.

Because details get complicated in spherical geometry, we assume here that within a limited scope the earth is flat and that the satellite path is parallel to this flat surface. The problem is to estimate how accurately a platform can be located if the trajectory of the satellite is known but the platform oscillator has a frequency error Δf and a random noise component (assume gaussian) with power spectral density $(k/2)\exp(-k|\omega|)$. Some details on satellites can be found in Reference 2 of Chapter 4. State any assumptions you make. Note that your results will involve both a bias error and a random error component.

CHAPTER

9 Information and Digital Transmission

Equipped with some knowledge of probability, we now return to a discussion of pulse code modulation (PCM). Of particular importance is the performance of PCM in the presence of noise and how it compares not only with other modulation systems but to the ideal communication system whose performance is only restricted by band-limited noise.

9.1 A MEASURE OF INFORMATION

The purpose of a communication system is, in the broadest sense, the transmission of information from one point in space and time to another. In preceding chapters we have examined several ways of accomplishing this goal by use of electrical signals. We have not, however, really specified what we mean by the term *information* even though we have referred to it and such related topics as messages, bandwidth, etc. In this and the two succeeding sections we shall briefly explore basic ideas about what information is, how it can be measured, and how these ideas relate to bandwidth, signal-to-noise ratio, etc. Exploration of these ideas becomes necessary in order to make meaningful comparisons of PCM with analog and analog pulse-modulation systems in the presence of noise.

The amount of information about an event is closely related to its probability of occurrence. Messages containing knowledge of a high probability of occurrence (i.e., those indicating very little uncertainty in the outcome) convey relatively little information. In contrast, those messages containing knowledge of a low probability of occurrence convey relatively large amounts of information. Thus a measure of the information received from the knowledge of occurrence of an event is inversely related to the probability of its occurrence.

To formulate a mathematical equation following the above reasoning, we note that if an event is certain (i.e., one occurring with probability 1), it conveys zero information. In the other extreme, if an event is improbable (i.e., one occurring with probability near 0), its occurrence conveys a large amount of information. We also desire our measure of information to be real-valued and monotonic, and to be additive for

events that are statistically independent. Based on these considerations, the measure of information associated with an event A occurring with probability P_A is defined as[†]

$$I_A = \log \frac{1}{P_A}. \tag{9.1}$$

The choice of base of the logarithm in Eq. (9.1) has purposely been left undefined at this time and will be considered in the following discussion.

Consider two equiprobable events, A and B, and a binary system to transmit the knowledge of their occurrence. A minimum of one binary pulse or signal is required, and this basic quantum unit of information is called a *bit*. The case of four equiprobable events requires four distinct binary pulse patterns, or two bits, while eight equiprobable events require three bits, etc. In general, any one of n equiprobable messages then contains $\log_2 n$ bits of information. Because we have assumed all n messages to be equiprobable, the probability of occurrence of each one is $P_i = 1/n$ and the information associated with each message is then

$$I_i = \log_2 n, \tag{9.2}$$

$$I_i = \log_2 \frac{1}{P_i} \qquad \text{bits.}$$

Equation (9.2) confirms what we had in Eq. (9.1) and supplies the units in terms of the base 2 logarithm. Even if the number of possibilities is not a power of 2, we can use Eq. (9.2) to express the information in bits, although the answer will have a fractional value. For numerical calculations it is often convenient to make use of the base 10 logarithm and then use the conversion,

$$\log_2 N = (\log_{10} N)(\log_2 10) = (\log_{10} N)/\log_{10} 2,$$

$$\log_2 N \cong 3.322 \log_{10} N. \tag{9.3}$$

The above results define our measure of information for the somewhat special case in which all messages are equally likely. To generalize, we define an *average information*, which is called the *entropy* H, of each message by taking the expected value of Eq. (9.2),[‡]

$$H \triangleq I_{\text{avg}} = E\{I_i\},$$

or

$$H = \sum_{i=1}^{n} P_i \log_2 \frac{1}{P_i} \qquad \text{bits.} \tag{9.4}$$

Thus each event or message is assigned an average information content even though the information for each event may fluctuate considerably.

[†] We have not proved that Eq. (9.1) follows from these conditions but have only stated it. Proof can be found in references on the subject (see, for example, R. Ash, *Information Theory*, New York: Wiley-Interscience, 1965, Chapter 1).

[‡] The name *entropy* and its symbol H are borrowed from a similar equation in statistical mechanics.

Now we encounter a problem in terminology. Previously we have always referred to binary pulses as bits, but now we are defining the bit as a measure of information content. For equiprobable events, the two definitions coincide, but this is admittedly a special case. One way to avoid this problem is by making use of the contraction *binit* for *binary digit* to distinguish it from the bit (which then is reserved for a measure of information content). In fact, we can easily show that the bit is always equal to or less than the binit because from probability theory we can write

$$E\{I_i\} \leq I_i\,|_{\max},$$

or

$$H \leq I_i\,|_{\max}. \tag{9.5}$$

The equality holds if all events or messages are equiprobable [this can be shown by using Eqs. (9.2) and (9.4) in Eq. (9.5)]. In other words, if all messages are equiprobable, then one binit is equal to one bit. This is an appealing approach, but the use of the term *bit* is so firmly entrenched in both senses that we bow to convention and use only the term *bit* for both meanings. Thus the term *bit* is used *both* as a binary digit and as a measure of information and we must rely on the context to determine which meaning is applicable. The saving fact is that both definitions coincide for the equally likely case.

EXAMPLE 9.1.1

The four symbols A, B, C, D occur with probabilities $1/2, 1/4, 1/8, 1/8$, respectively. Compute the information in the three-symbol message $X = BDA$, assuming that the symbols are statistically independent.†

Solution Because the symbols are independent, the measure of information is additive and we can write

$$I_X = \log_2 4 + \log_2 8 + \log_2 2,$$
$$I_X = 2 + 3 + 1,$$
$$I_X = 6 \text{ bits}.$$

† In construction of languages, one is constrained by certain rules of spelling and grammar that limit the choice in successive symbols. Therefore the choice of symbols is no longer completely independent. This can be accounted for by use of conditional probabilities and hence conditional entropy, topics that are interesting in their own right but are not covered here.

EXAMPLE 9.1.2

Determine and graph the entropy (i.e., average information) of the binary code in which the probability of occurrence of the two symbols is p and $q = (1 - p)$.

Solution Using Eq. (9.4), we have

$$H = \sum_{i=1}^{2} P_i \log_2(1/P_i),$$

$$H = p \log_2(1/p) + (1 - p) \log_2[1/(1 - p)].$$

A plot of H versus p is shown in Fig. 9.1. Note that the maximum entropy is one bit/symbol and occurs for the equiprobable case (i.e., $p = q = 1/2$)—cf. Eq. (9.5).

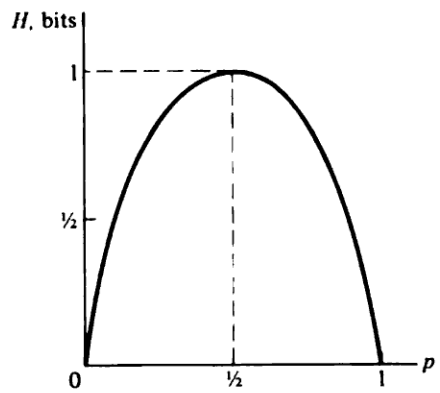


Figure 9.1 Entropy of a binary source.

DRILL PROBLEM 9.1.1

Determine the average information (entropy) per symbol for (a) the four-symbol code in Example 9.1.1; (b) the case in which four symbols are equiprobable.

ANSWER: (a) 1.75 bits; (b) 2 bits.

DRILL PROBLEM 9.1.2

An efficient binary code for Example 9.1.1 is $A = 0$, $B = 10$, $C = 110$, $D = 111$. By weighting each with the appropriate probability of occurrence, determine (a) the probability of occurrence of 0's in the code; (b) the average information per codeword.

ANSWER: (a) 1/2; (b) 1.75 bits.

The above concepts can be extended to the case of continuous messages, but the mathematics become quite involved and abstract. We shall prefer instead to use the discrete approach and employ the sampling theorem and quantization to reduce continuous analog messages to discrete ones.

In electrical communication systems we are particularly concerned with the rate of information transmission. We define the information-transmission rate R by

$$R = rH \qquad \text{bps}, \tag{9.6}$$

where r is the symbol rate (corresponding to the symbols over which H is determined), and bps is bits per second.

Although it is not generally within the realm of a communication system design to specify message probabilities (these usually either are known or can be approximated by measurements), a major concern is the maximization of the information-transmission rate of the system. This will, in fact, form the subject of the next section.

DRILL PROBLEM 9.1.3

A limit on the information-transmission rate R in the presence of thermal noise and assuming no bandwidth restrictions can be expressed as[†]

$$R_{\max} = \frac{1.44P_r}{kT_{eq}} \qquad \text{bps},$$

where P_r is the received power in watts, k is the Boltzmann constant, and T_{eq} is the equivalent receiver noise temperature in Kelvin. Calculate the minimum transmitter power required to send a standard black-and-white television picture with an information content of 10^6 bits from a deep-space probe within a time interval of 1 hour if the attenuation losses are 200 dB and $T_{eq} = 50$ K.

ANSWER: 13.3 W.

We conclude this section with a word of caution. The development of the concepts of information and maximum information-transmission rate have proceeded strictly from an engineering viewpoint. Thus the message content has been judged from its quality of uncertainty, reducing the transmission problem to a highly systematic basis. However, the validity of this approach may be open to question when we attempt to apply it to problems involving human behavior. The problem is not so much the concept of a measure of information but how it relates to human value and decision processes and what criteria are involved. For example, the artistic and aesthetic senses are not necessarily best served (if at all) by the criterion of a most efficient transfer of information. Failure to recognize some of these and other related factors has resulted in not a few engineers striving to milk the very last ounce of performance out of a communication system by using information theory only to find that no one could really use the output data after all!

[†] This relation is developed in Example 9.2.3.

9.2 CHANNEL CAPACITY

Previously we have referred to the link between transmitter and receiver, including noise sources, as the transmission medium. In information theory it is convenient to treat it as a more abstract and general noisy filter (which may be nonlinear, time-varying, etc.) called the *channel*. The limiting rate of information transmission through a channel is called the *channel capacity*.

Importance of the concept of channel capacity stems from a theorem first stated and proved by C. E. Shannon. Shannon's theorem states that if the entropy rate R is equal to or less than the channel capacity C, then there exists a coding technique that enables transmission over the channel with an arbitrarily small frequency of errors, or

$$R \leq C. \tag{9.7}$$

This restriction holds even in the presence of noise in the channel. A converse to this theorem states that it is not possible to transmit messages without errors if $R > C$. Thus the channel capacity is defined as the maximum rate of reliable information transmission through the channel.

Now consider a source with an available alphabet of α discrete messages. We assume that each message sent can be identified at the receiver; therefore this case is often called the "discrete noiseless channel." The maximum entropy of the source is $\log_2 \alpha$ bits, and if T is the transmission time of each message, the channel capacity is[†]

$$C = \frac{1}{T} \log_2 \alpha \qquad \text{bps}. \tag{9.8}$$

To attain this maximum, the messages must be equiprobable and statistically independent. These conditions form a basis for the coding of the information to be transmitted over the channel.

In the presence of noise, the capacity of the discrete channel decreases as a result of the errors made in transmission. Because entropy is a measure of randomness, this may be accounted for by subtracting the entropy rate of the symbols which were erroneously detected from the channel capacity of the discrete noiseless channel.

EXAMPLE 9.2.1

A binary source sends α equiprobable and identifiable messages using the binary symbols 0, 1 at a symbol rate r over a time T. However, as a result of noise, a 0 may be mistaken for a 1 and vice versa. For the case known as the "binary symmetric channel," the probabilities of making these two types of errors are equal and are designated here by p. Determine the resulting channel capacity.

[†] For unequal message transmission times,

$$C = \lim_{T \to \infty} \frac{1}{T} \log_2 \alpha(T).$$

Solution The number of equiprobable messages is $\alpha = 2^{rT}$. Using this in Eq. (9.8), we have

$$C = \frac{1}{T} \log_2 2^{rT} = r .$$

In the presence of noise, errors will be made at the entropy rate R_e which, from the results of Example 9.1.2, is

$$R_e = rH_e = rp \log_2(1/p) + r(1 - p) \log_2[1/(1 - p)] .$$

The channel capacity is then

$$\begin{aligned} C &= r - R_e \\ &= r[1 + p \log_2 p + (1 - p) \log_2(1 - p)] . \end{aligned}$$

Note that for $p = 0.5$, $C = 0$ and no information can be transmitted. For $p = 0.001$, $C = 0.989r$ so that C is about 989 bps for r of 1000 symbols per second.

In making comparisons between various types of communication systems, it is convenient to consider a channel that is described in terms of bandwidth and signal-to-noise ratio. Another theorem in information theory, known as the Hartley–Shannon theorem, states that the channel capacity of a white band-limited gaussian channel is

$$C = B \log_2(1 + S/N) \qquad \text{bps} , \tag{9.9}$$

where B is the channel bandwidth and S/N is the mean-square signal-to-noise ratio. Although restricted to the gaussian case, this result is of widespread importance to communication systems because many channels can be modeled by the gaussian channel. It is applicable to both discrete and continuous channels.

A proof of the Hartley–Shannon theorem is not easy and is quite involved. We shall follow an intuitive explanation based on the above reasoning for the discrete channel.

Consider a signal which is band-limited to B Hz. Using the sampling theorem, we can describe this signal by its samples taken at the rate

$$r = 2B . \tag{9.10}$$

How many levels shall we use to describe each sample in the presence of noise? A plausible choice is one in which each discernible level is spaced by one standard deviation of the noise. For zero-mean noise with power N, this gives level spacings of $\sigma_n = \sqrt{N}$. Using the rms value of the observed signal-plus-noise, $y(t)$, for the range of possible signals, we can formulate the number of discernible levels M as

$$M = \frac{\sigma_y}{\sigma_n} = \frac{\sqrt{S + N}}{\sqrt{N}} = \sqrt{1 + \frac{S}{N}} . \tag{9.11}$$

The channel capacity over $(0, T)$ is, from Eq. (9.8),

$$C = \frac{1}{T} \log_2 M^{rT} = r \log_2 M . \tag{9.12}$$

Use of Eqs. (9.10) and (9.11) in Eq. (9.12) gives

$$\begin{aligned} C &= 2B \log_2 \sqrt{1 + S/N} \\ &= B \log_2(1 + S/N) . \end{aligned}$$

An application of this result reveals that if $S/N = 1$, then $C = B$; in other words, the channel capacity in bps is equal to the channel bandwidth in Hz. However, if $S/N = 3$, then $C = 2B$, if $S/N = 7$, then $C = 3B$, etc. Thus the channel capacity increases as the S/N increases. Alternatively, for a fixed channel capacity, the bandwidth B can be reduced in exchange for an increase in signal-to-noise ratio S/N. In the ideal case, the results indicate that this trade-off is approximately exponential.

EXAMPLE 9.2.2

A black-and-white television picture may be considered as composed of approximately 3×10^5 picture elements. Assume that each picture element is equiprobable among 10 distinguishable brightness levels. Thirty (30) picture frames are transmitted per second. Calculate the minimum bandwidth required to transmit the video signal assuming that a 30-dB signal-to-noise ratio is necessary for satisfactory picture reproduction.

Solution Information per picture element $= \log_2 10 = 3.32$ bits.

Information per picture frame $= (3.32)(3 \times 10^5) = 9.96 \times 10^5$ bits.

Information rate $R = (30)(9.96 \times 10^5) = 29.9 \times 10^6$ bps.

Because R must be less than or equal to C, we let $R = C = 29.9 \times 10^6$ and Eq. (9.9) gives

$$B_{\min} = \frac{C}{\log_2(1 + S/N)},$$

$$B_{\min} = \frac{29.9 \times 10^6}{(3.32)(3.0004)} \approx 3 \text{ MHz} .$$

As noted in Appendix D, commercial television transmissions actually use a video bandwidth of 4 MHz.

DRILL PROBLEM 9.2.1

What is the minimum time required for the facsimile transmission of one picture over a standard telephone circuit? There are about 2.25×10^6 picture elements to be transmitted and 12 brightness levels are to be used for good reproduction. Assume all brightness levels are equiprobable. The telephone circuit has a 3-kHz bandwidth and a 30-dB signal-to-noise ratio (these are typical parameters).

ANSWER: 4.5 min.

Although the potential signal-to-noise trade-off with increased bandwidth is exponential in PCM, its performance in conveying information still falls short of that predicted by the Hartley–Shannon law. One reason for this is that the Hartley–Shannon law assumed no limitation on the number of pulse levels except that constrained by the signal-to-noise ratio of the channel. Thus such a system would use more levels as the signal-to-noise ratio improved. In contrast, practical PCM systems make use of a (fixed) finite number of pulse levels and their maximum information transmission is limited by this fact.

A second reason is that PCM, like the other wideband communication systems we have examined, suffers from a threshold effect. Therefore, an increase in the minimum signal-to-noise, on the order of 6–8 dB, is required for significant signal-to-noise improvement. We shall consider this threshold effect in a later section when the probability of error for PCM is discussed.

EXAMPLE 9.2.3

Starting with the Hartley–Shannon law, derive a relation for the channel capacity in the presence of white noise as the bandwidth is increased without limit.

Solution For white noise with power spectral density $\eta/2$, the Hartley–Shannon law is

$$C = B \log_2\left(1 + \frac{S}{\eta B}\right)$$
$$= \frac{S}{\eta}\left(\frac{\eta B}{S}\right)\log_2\left(1 + \frac{S}{\eta B}\right),$$

so that we wish to investigate the limit,

$$\lim_{B\to\infty}\left[\frac{S}{\eta}\log_2\left(1 + \frac{S}{\eta B}\right)^{\eta B/S}\right].$$

However, we can also write

$$\lim_{x\to 0}(1 + x)^{1/x} = e,$$

and, identifying x as $x = S/(\eta B)$, we have

$$\lim_{B\to\infty} C = \frac{S}{\eta}\log_2 e\,,$$

or

$$\lim_{B\to\infty} C = 1.44\frac{S}{\eta} = \frac{S}{0.69\eta}.$$

This is an interesting result because it gives the maximum information-transmission rate possible for a system of given power but no bandwidth limitations. This result is useful in the design of minimum-power systems, such as satellite and space probe communication systems, for which the bandwidth is not specified or is a secondary consideration. The power spectral density can be specified in terms of equivalent noise temperature by $\eta = kT_{eq}$.

9.3 IDEAL DEMODULATOR DETECTION GAIN

The Hartley–Shannon law provides us with a measure of the maximum information-transmission rate over a band-limited channel in the presence of white gaussian noise. But it also provides us with the relationship between bandwidth and signal-to-noise ratio for an optimum system. And because it holds for continuous as well as discrete channels, we can apply the results to both to make comparisons between types of systems.

Consider a message signal that is band-limited to f_m Hz having an information content of I bits. Before transmission of this information we may choose to use a modulator (or encoder). If performed correctly, this process of modulation (encoding) will not alter the information content of the message; it is designed to combat the degrading effects of the channel. Let the transmitted bandwidth be B Hz and assume that the information-transmission rate R is matched to the channel so that $R = C$, where C is the channel capacity. At the receiver, the transmitted signal is received within the bandwidth B in the presence of white gaussian noise, so we can write

$$C = B\log_2(1 + S_i/N_i)\,. \tag{9.13}$$

Here S_i/N_i designates the signal-to-noise ratio at the input of the receiver.

In the receiver, the modulated signal is applied to the input of an ideal demodulator (see Fig. 9.2). There is no loss of information in the ideal demodulator and there-

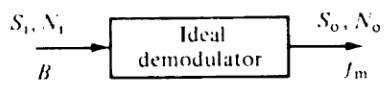


Figure 9.2 The ideal receiver.

fore the information-transmission rate at the output is identical to that at the input (assuming no time scaling). Thus we can also write

$$C = f_m \log_2(1 + S_o/N_o), \tag{9.14}$$

where S_o/N_o is the signal-to-noise ratio at the receiver output.

Setting the right-hand sides of Eqs. (9.13) and (9.14) equal, we have

$$B \log_2(1 + S_i/N_i) = f_m \log_2(1 + S_o/N_o),$$

or

$$(1 + S_o/N_o) = (1 + S_i/N_i)^{B/f_m}. \tag{9.15}$$

For S_i/N_i, $S_o/N_o \gg 1$, we obtain

$$\frac{S_o}{N_o} \approx \left(\frac{S_i}{N_i}\right)^{B/f_m}. \tag{9.16}$$

We conclude that in the optimum system the receiver output signal-to-noise ratio increases *exponentially* with the transmitted bandwidth B. In contrast, recall that for angle modulation (FM, PM) and pulse-timing modulation (PPM, PWM) the signal-to-noise (power) ratio increases proportional to the square of the bandwidth. Although superior in this respect to AM and PAM, which offer no signal-to-noise versus bandwidth advantage, such systems are far from optimum when $B \gg f_m$. As we shall see later, the signal-to-noise improvement in PCM systems also falls short of the optimum trade-offs.

9.4 QUANTIZATION NOISE

In PCM, the information to be transmitted is contained in the code. Noise is added in the transmission and this may cause errors in recognition of the code symbols sent. But if the signal pulses are received above a certain signal-to-noise threshold (which we will discuss later in this chapter), the average error rate can be kept very low.

However, in order to encode a continuous signal we must first quantize it into a finite number of discrete amplitude levels. Once quantized, the instantaneous values of the continuous signal can never be reconstructed exactly. The random errors introduced are called *quantization noise*. Quantization noise, in contrast to transmission noise, is artificially generated by the quantization process prior to being transmitted. It can be reduced to any desired degree by choosing the number and distribution of the quantizing levels.

To calculate the mean-square quantization noise, we assume equal amplitude increments between levels. Specifically, let the input signal be quantized into n levels, each spaced by an amplitude increment, a. (For binary PCM, n is chosen to be a power of 2.) If we assume that the signal amplitude is bipolar and there is no dc

level to offset the amplitudes, a reasonable distribution of the quantizing levels is $\pm a/2, \pm 3a/2, \pm 5a/2, \ldots, \pm(n-1)a/2$ (see Fig. 9.3).

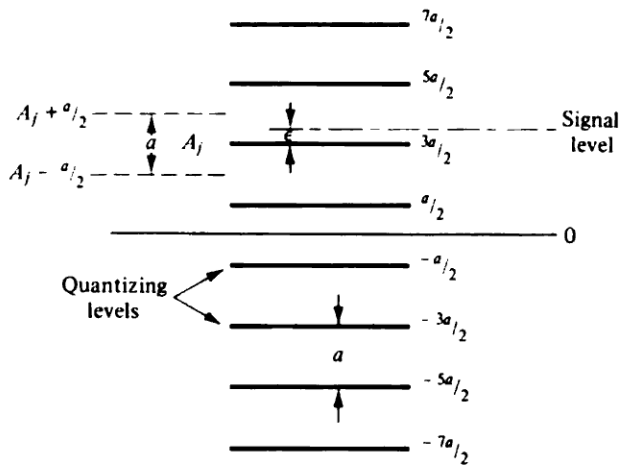


Figure 9.3 Quantization uncertainty for continuous input.

Assuming that the number of signal levels is even and that all signal levels are equiprobable, we find the probability density function of the levels is

$$p(x) = \sum_{\substack{i=-n \\ \text{odd}}}^{n} \left(\frac{1}{n}\right) \delta(x - ia/2). \tag{9.17}$$

The mean-square signal after quantization is

$$\begin{aligned} S &= \int_{-\infty}^{\infty} x^2 p(x)\, dx \\ &= 2\int_{0}^{\infty} x^2 p(x)\, dx \\ &= \frac{2}{n}\{(a/2)^2 + (3a/2)^2 + \cdots + [(n-1)a/2]^2\} \\ &= \frac{2}{n}\frac{n(n-1)(n+1)}{6}(a/2)^2 \end{aligned}$$

$$S = \frac{n^2 - 1}{12} a^2. \tag{9.18}$$

The same method can be applied if not all levels are equiprobable, but the result corresponding to Eq. (9.18) cannot be written in as compact a form.

In the quantization process, each sample of the continuous input signal is approximated to the nearest allowed level. This is shown in Fig. 9.3, where the input signal level is ϵ units from the nearest quantizing level A_j. At the eventual receiver output, the quantized level A_j could have been due to any signal amplitude in the range

$A_j - a/2$ to $A_j + a/2$. This uncertainty could just as well have been due to additive noise and therefore we treat it as an additive noise source.

The quantizing error ϵ is the difference between the signal level and the nearest allowed quantizing level, as shown in Fig. 9.3. Lacking any information to the contrary, we assume that all values of ϵ are equally likely anywhere in the range $-a/2 \leq \epsilon < a/2$ so that it can be described by the uniform probability density function,

$$p(\epsilon) = \begin{cases} \dfrac{1}{a} & -a/2 \leq \epsilon < a/2, \\ 0 & \text{elsewhere.} \end{cases} \tag{9.19}$$

Actually, the result is not highly dependent on this assumption as long as the probability density function is not strongly peaked somewhere in the interval (see, for example, Drill Problem 9.4.2). The mean-square quantization noise now follows readily:

$$\begin{aligned} \overline{\epsilon^2} &= \int_{-\infty}^{\infty} \epsilon^2 p(\epsilon)\, d\epsilon \\ &= \int_{-a/2}^{a/2} \frac{1}{a} \epsilon^2\, d\epsilon \\ \overline{\epsilon^2} &= \frac{a^2}{12}. \end{aligned} \tag{9.20}$$

Because the peak signal excursion (range) is $na/2$, the peak signal-to-noise (power) ratio is

$$\left(\frac{S}{N}\right)_{\text{pk qnt}} = \frac{(na/2)^2}{a^2/12} = 3n^2. \tag{9.21}$$

Expressed in decibels, this becomes

$$\left[\left(\frac{S}{N}\right)_{\text{pk qnt}}\right]_{\text{dB}} = 4.8 + 20 \log_{10} n. \tag{9.22a}$$

For a binary code, we have $n = 2^m$ so that, for the binary case, Eq. (9.22a) becomes

$$\left[\left(\frac{S}{N}\right)_{\text{pk qnt}}\right]_{\text{dB}} = 4.8 + 6m \qquad \text{(binary system).} \tag{9.22b}$$

Thus the peak-signal–to–rms-quantization-noise ratio increases by 6 dB for every additional bit used in a binary system.

For the case in which all signal levels are equiprobable, the mean-square signal is given by Eq. (9.18); dividing this by Eq. (9.20), we get

$$\left(\frac{S}{N}\right)_{\text{qnt}} = n^2 - 1 \approx n^2. \tag{9.23}$$

The performance quality, on a mean-square basis, increases as the square of the number of levels used and differs only by a constant from the peak S/N relation given in

Eq. (9.21). In both cases, the signal–to–quantization-noise ratio increases about 6 dB for each additional bit used. However, it is difficult to assign a criterion for performance in general. For speech, experimental results indicate that 8 to 16 levels are adequate for intelligibility, $n = 128$ to $n = 256$ (i.e., $m = 7$ to $m = 8$) is standard for voice telephony, while as many as 4096 ($m = 12$) to 65,536 ($m = 16$) are used in music-quality audio.

The calculation above assumed that all signal levels were equiprobable. However, this may not be the case (and often is not in practice). This results in some increments in signal amplitude having a poorer signal-to-quantization noise than others. For example, experiment indicates that typical instantaneous-speech amplitudes are less than one-fourth the rms value for 50% of the time.[†] Thus in the encoding of speech the small-amplitude signals, which occur most of the time, have relatively large quantization noise, resulting in a rather poor overall signal–to–quantization-noise ratio.

A remedy for this situation is to use a nonlinear (i.e., nonuniform) quantizer that has tapered quantization steps (e.g., see Fig. 7.25). A method often used in practice to achieve an equivalent effect is to weight the input signal before quantizing with a linear (uniform) quantizer. An inverse weighting is applied at the receiver so that the overall transmission is distortionless.

The input-output characteristic of a nonlinear predistortion system that could be used for signals such as speech is shown in Fig. 9.4. The effect of the nonlinear characteristic is to compress the larger signal amplitudes and therefore such a system is called a *compressor*.[‡] The inverse operation at the receiver is called an *expander*. The combination of a compressor and an expander is called a *compander*. Companding systems are often synthesized using active circuits whose gains are controlled by bi-

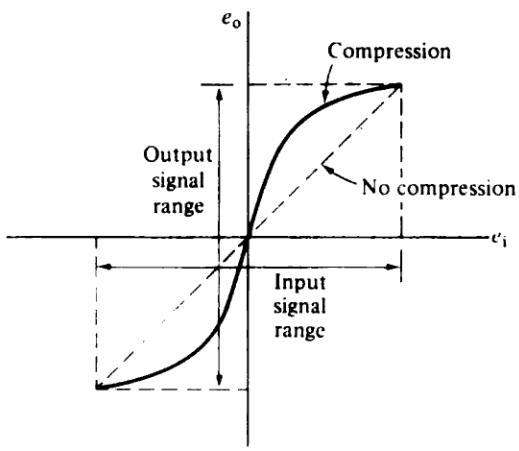


Figure 9.4 An input-output characteristic of a signal compressor.

[†] J. A. Betts, *Signal Processing, Modulation and Noise*, New York: American Elsevier Publishing Company, 1970, Chapter 7.

[‡] This type of signal compression is sometimes used in AM systems for voice communications to effectively obtain the effects of a higher-modulation index without exceeding the peak-modulation limit. A logarithmic taper is generally preferred for speech applications.

ased diode networks.† By this means, one can approximate the required nonlinear characteristic with a piecewise linear approximation.

EXAMPLE 9.4.1

Consider a continuous input signal having the following probability density function:

$$p(x) = \begin{cases} 1 - |x| & -1 \le x \le +1, \\ 0 & \text{elsewhere.} \end{cases}$$

a) Determine the quantizer step size and levels if a linear (uniform) eight-level ($n = 8$) quantizer is used.

b) Determine the nonlinear quantizer step levels required to make the quantized signal levels equiprobable. Also plot the required compressor characteristic to precede a linear quantizer.

Solution a) The net signal range is 2 so that the level step size, a, is $a = 2/n = \frac{2}{8} = \frac{1}{4}$. The quantizer levels are then set at $\pm\frac{1}{8}, \pm\frac{3}{8}, \pm\frac{5}{8}, \pm\frac{7}{8}$, and the quantizer round-off boundaries are set at $0, \pm\frac{1}{4}, \pm\frac{1}{2}, \pm\frac{3}{4}, \pm 1$.

b) As a result of the symmetry, we consider the case only for $x > 0$. To make all signal levels equiprobable we set the quantizer step levels such that the area under $p(x)$ is divided equally (see Fig. 9.5a). Designating the quantizer steps by a, b, c, d, respectively, we have

$$\int_0^a (1 - x)\,dx = \int_d^1 (1 - x)\,dx = \tfrac{1}{16},$$

and

$$\int_a^b (1 - x)\,dx = \int_b^c (1 - x)\,dx = \int_c^d (1 - x)\,dx = \tfrac{1}{8}.$$

Solving in succession, we get the quantizer steps,

$$a = 1 - (\sqrt{7/2})/2 = 0.065,$$
$$b = 1 - (\sqrt{5/2})/2 = 0.209,$$
$$c = 1 - (\sqrt{3/2})/2 = 0.388,$$
$$d = 1 - (\sqrt{1/2})/2 = 0.646.$$

† In PCM telephone systems in the United States, the companding law used prior to transmission is $e_o = [\log(1 + 255e_i)]/[\log(256)]$, where both e_i, e_o are normalized to a (0, 1) range. This is approximated using 15 linear segments in the gain characteristic of an amplifier.

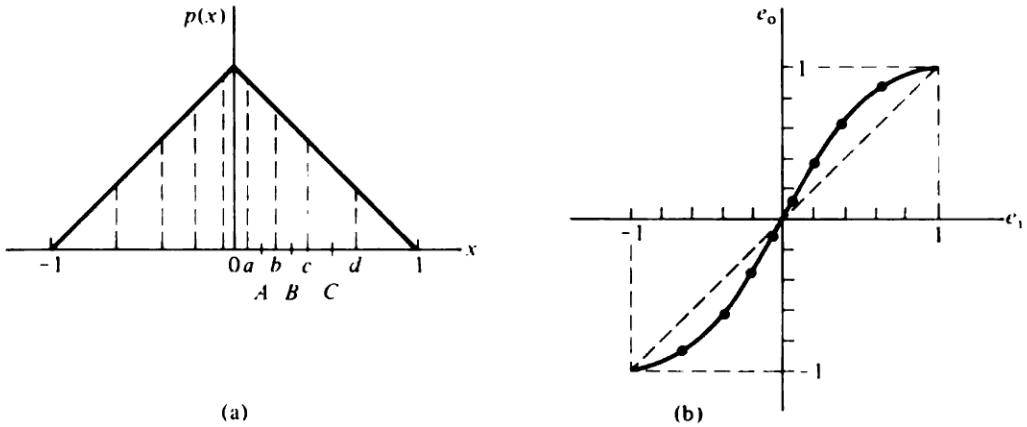


Figure 9.5 The probability density function and compressor characteristic for Example 9.4.1.

The corresponding compressor characteristic to precede a linear (uniform) quantizer to achieve the same results can be found by plotting the points (0.065, 0.125), (0.209, 0.375), etc. This is shown in Fig. 9.5(b). Similarly, the quantizer round-off boundaries are $0, \pm A, \pm B, \pm C, \pm 1$, where A, B, C are determined from

$$\int_0^A (1-x)\,dx = \int_A^B (1-x)\,dx = \int_B^C (1-x)\,dx = \int_C^1 (1-x)\,dx = \tfrac{1}{8},$$

or

$$A = 1 - \sqrt{3}/2 = 0.134,$$
$$B = 1 - \sqrt{2}/2 = 0.293,$$
$$C = 1 - 1/2 = 0.500.$$

DRILL PROBLEM 9.4.1

Compressor characteristics determined by a mapping to a uniform probability density, as illustrated in Example 9.4.1, may tend to compress too much if there are large differences in $p(x)$ from one region to another. This can be corrected by using the area under $[p(x)]^{1/3}$ instead of $p(x)$ between increments.† Recompute the quantizer points for Example 9.4.1 using this method (in this case, there are no large differences between adjacent regions, so the results may be expected to be similar to those in Example 9.4.1).

ANSWER: 0.095, 0.297, 0.521, 0.790.

---

† See, for example, J. J. Spilker, Jr., *Digital Communications by Satellite*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1977, p. 45.

DRILL PROBLEM 9.4.2

(a) Calculate the mean-square quantization noise for the signal described by the pdf of Example 9.4.1 if there were only two quantization levels (± 0.5). Do not assume that the pdf of the quantization error is constant over each level. (b) Calculate the mean-square quantization noise using Eq. (9.20), and compare.

ANSWER: $\frac{1}{12}$ (both cases).

The method of choosing quantization steps suggested in Example 9.4.1 is relatively easy to apply. It does not, however, guarantee optimum results even though the results may be satisfactory for many applications. A more fundamental method is to choose the quantization points in such a way as to minimize the quantization error in a mean-square sense. A solution to this problem can be found in a manner similar to that used for the orthogonal signal representations in Section 2.5.[†] However, the limits of integration are dependent upon the quantization points, and an iterative numerical procedure must be used to find a solution (cf. Problem 9.4.7).

9.5 PROBABILITY OF ERROR IN TRANSMISSION

We now devote some attention to the details of transmitting digitally encoded messages and begin by considering the use of the type of digital signals that are intended for direct transmission over a line. These are called *baseband* signals. The case of modulated pulses intended for transmission via radio-frequency propagation is considered in Chapter 10.

In a binary PCM system, the two pulse levels may be represented as an on-off signal—i.e., by an amplitude A and an amplitude 0 as the two possible levels—or by equal-amplitude pulses of opposite polarities—i.e., by amplitudes $-A/2$ and $+A/2$. The latter is known as a *polar* binary signal. The on-off signal results in a nonzero average level that is dependent on the relative proportions of ones and zeros being transmitted. In a random sequence of on-off pulses the slight variation in average level gives rise to a line spectrum in addition to the continuous frequency spectrum of the information. The spectral component at the bit rate can be used for receiver synchronization. On the other hand, the polar signal representation requires less transmitted power for a specified error rate in the presence of noise. Its use also makes setting the receiver decision threshold easier if the relative frequencies of occurrence of both binary signals are equal. Both of these topics will be discussed later in this section.

[†] See, for example, F. Haber, *An Introduction to Information and Communication Theory* (Advances in Modern Engineering Series, vol. 4), Reading, Mass.: Addison-Wesley, 1974, Chapter 6.

A block diagram of an overall system for baseband communication is shown in Fig. 9.6. Because the bandwidth of the transmission line is severely limited by wire capacitance and inductance, line equalization is usually employed at regular intervals along the line. This counteracts the capacitive and inductive effects and widens the usable bandwidth of the line. Even with equalization, however, the transmission frequency response does not closely approximate that of an ideal low-pass filter; and the methods of pulse shaping discussed in Section 7.3 are employed to minimize the intersymbol interference (ISI).

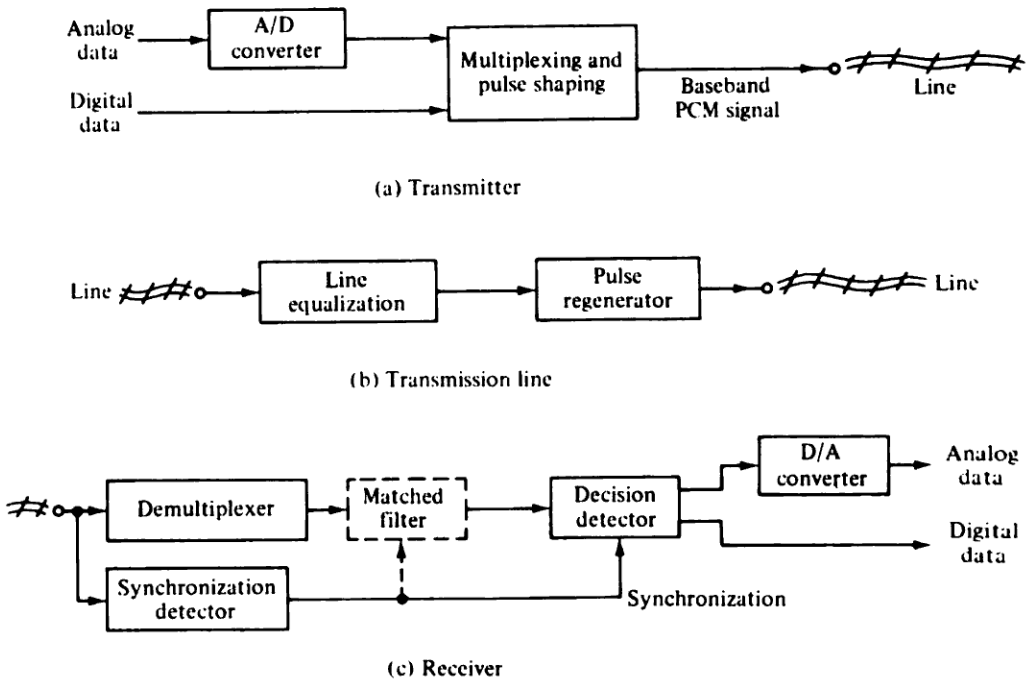


Figure 9.6 A baseband PCM communication system.

Note that in using the methods for pulse shaping to minimize ISI, we are more concerned with combating the bandwidth limitations of the transmission line than with the effects of transmission noise. The matched filter, which was discussed in Chapter 7, can be employed to maximize the signal-to-noise ratio in the receiver. However, the matched filter dictates a different type of pulse-waveform design. In the design of radio systems for digital signals, both of these types of requirements need to be considered and appropriate compromises can be made. In line systems the signal-to-noise ratio is generally sufficiently high so that the matched filter may be deleted. Instead, attention is concentrated more on widening the effective bandwidth using equalization to gain higher data rates and on minimizing the intersymbol interference by appropriate waveform shaping.

We begin our discussion of the probability of error by not using a matched filter receiver. Assume that the binary levels at the receiver are 0 and A; i.e., the received

signal $f(t)$ is either at the level 0 or at the level A. The observed waveform $y(t)$ at the receiver also contains additive noise $n(t)$ so that

$$y(t) = f(t) + n(t). \tag{9.24}$$

At a given time $t = t_1$, the possible receiver inputs are either

$$y(t_1) = A + n(t_1) \qquad \text{(signal present)}, \tag{9.25}$$

or

$$y(t_1) = n(t_1) \qquad \text{(signal absent)}. \tag{9.26}$$

The correct decision whether $f(t)$ is present or absent could easily be made from Eqs. (9.25) and (9.26) were it not for the fact that $n(t_1)$ is random. Therefore we use probability to describe the performance. Realizing that in general there is a nonzero probability that the noise amplitude can exceed that of the signal, we prefer to speak of a *probability of error* in making the correct decision.

To calculate the probability of error, we first make a rule for the decision. Let the decision threshold be designated by μ. At the time $t = t_1$, we use the following decision rule:

$$\left\{\begin{array}{ll} y(t_1) > \mu: & \text{"signal present"} \\ y(t_1) < \mu: & \text{"signal absent"} \\ y(t_1) = \mu: & \text{"guess"}^{\dagger} \end{array}\right\}. \tag{9.27}$$

The next step in determining the optimum-decision threshold value μ will depend on the shape of the probability density function of $n(t_1)$.

We assume that the noise $n(t_1)$ is gaussian-distributed with zero mean value and a mean-square value of

$$\overline{n^2(t_1)} = \sigma^2. \tag{9.28}$$

Using Eqs. (9.28) and (8.50), we find that the probability density function of the input when there is no signal present, designated here by $p_0(y)$, is

$$p_0(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-y^2/(2\sigma^2)}. \tag{9.29}$$

The probability density function of the input when the signal $f(t)$ is present, designated here by $p_1(y)$, is also gaussian, but it has a mean value equal to A [see Eq. (9.25)],

$$p_1(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(y-A)^2/(2\sigma^2)}. \tag{9.30}$$

These probability density functions are shown in Fig. 9.7.

† It will soon be evident that, except in highly unusual cases, this alternative occurs with probability 0, so it is not of major concern to us.

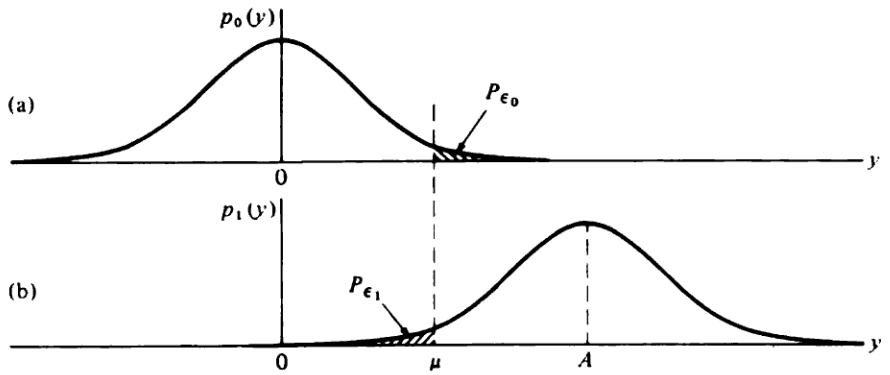


Figure 9.7 Probability density functions for the binary decision case.

The decision threshold is μ, and we see from Fig. 9.7 that there are instances when $y > \mu$ even though the signal is absent. This type of error is called a "false alarm" because the receiver decides that the signal is present even though it is not present. The probability of false alarm, designated here by P_{ϵ_0}, is given by the shaded area in Fig. 9.7(a). We can write this in equation form as

$$P_{\epsilon_0} = \int_{\mu}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-y^2/(2\sigma^2)}\, dy. \tag{9.31}$$

In a similar manner, we see from Fig. 9.7(b) that there are instances when $y < \mu$ even though the signal is present. This type of error is called a "false dismissal" because the receiver decides that the signal is absent even though it is present. The probability of false dismissal, designated here by P_{ϵ_1}, is given by the shaded area in Fig. 9.7(b), or

$$P_{\epsilon_1} = \int_{-\infty}^{\mu} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(y-A)^2/(2\sigma^2)}\, dy. \tag{9.32}$$

Although we can continue to speak of the probability of false-alarm errors and the probability of false-dismissal errors, it is more convenient to define a net probability of error to incorporate both types. For convenience, let us associate the binary digit 1 with the condition "signal present" and the binary digit 0 with the condition "signal absent" (we could just as well reverse the definition; the order is arbitrary). Now let P_0, P_1 be the source digit probabilities of zeros and ones, respectively; that is, P_0 is the probability that a binary 0 is sent. The usual case is for $P_0 = P_1 = 1/2$ (although they can be different as long as the sum is one—either a 0 or a 1 must be sent). In terms of these definitions, we define the net *probability of error*, P_ϵ, as

$$P_\epsilon = P_0 P_{\epsilon_0} + P_1 P_{\epsilon_1}. \tag{9.33}$$

If we assume the binary levels (zeros and ones) are equiprobable (that is, $P_0 = P_1 = 1/2$), the net probability of error is given by one-half the sum of the two shaded areas in Fig. 9.7:

$$P_\epsilon = \tfrac{1}{2}(P_{\epsilon_0} + P_{\epsilon_1}). \tag{9.34}$$

It is obvious from Fig. 9.7 that the sum of these two areas will be minimum if

$$\mu = A/2 . \tag{9.35}$$

This sets the decision threshold μ for a minimum probability of error. (Note that the value of μ for minimum probability of error is not so obvious if $P_0 \neq P_1$.) With this decision threshold the two shaded areas shown in Fig. 9.7 are equal as a result of symmetry, and we can write

$$P_\epsilon = \int_\mu^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-y^2/(2\sigma^2)}\, dy . \tag{9.36}$$

With the change of variable $z = y/\sigma$, this becomes

$$P_\epsilon = \int_{\mu/\sigma}^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz . \tag{9.37}$$

Referring to Eq. (8.56), we recognize that Eq. (9.37) can be rewritten as

$$P_\epsilon = Q(\mu/\sigma) . \tag{9.38}$$

Inserting the value of the decision threshold $\mu = A/2$, we obtain

$$P_\epsilon = Q[A/(2\sigma)] . \tag{9.39}$$

A graph of P_ϵ as a function of $A/(2\sigma)$ is shown in Fig. 9.8. Note that A is the (voltage) separation between the binary states and 2σ is the effective width of the proba-

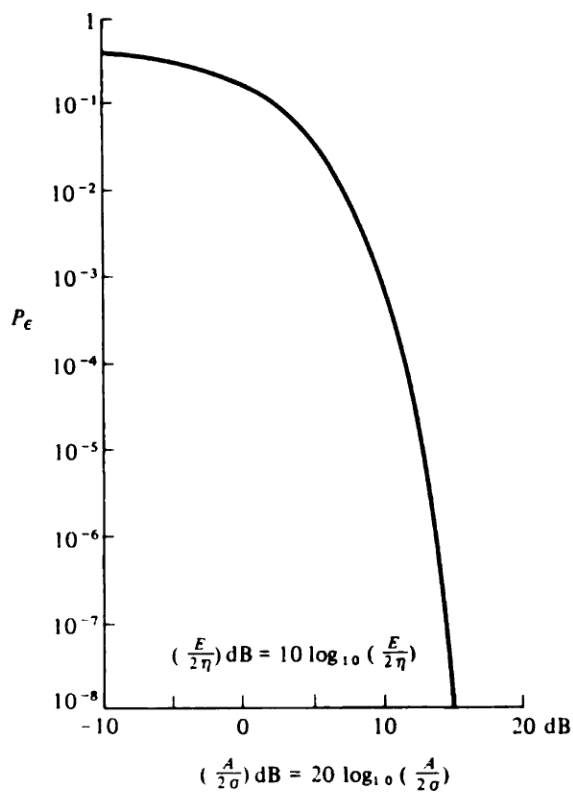


Figure 9.8 Probability of error versus $A/(2\sigma)$.

bility density of the additive noise. Thus we see that it is the separation between states, divided by the effective width of the noise pdf, that controls the probability of error in transmission.

These results can be put into more familiar units for the on-off case by letting the average signal power S be $S = (\frac{1}{2})(0)^2 + (\frac{1}{2})(A)^2 = A^2/2$. Also, we have that the average noise power is: $N = \sigma^2$ so that Eq. (9.39) can be rewritten for the on-off case as

$$P_\epsilon = Q\left(\sqrt{\frac{S}{2N}}\right) \qquad \text{(on-off binary)}. \tag{9.40}$$

The polar binary case follows in the same way except that $S = (A/2)^2$ so that

$$P_\epsilon = Q\left(\sqrt{\frac{S}{N}}\right) \qquad \text{(polar binary)}. \tag{9.41}$$

Therefore the average transmitted power for the on-off binary signal must be twice that of the polar binary signal to achieve the same net probability of error. The polar binary signal also has an advantage in that the optimum-decision threshold is at zero for $P_0 = P_1$, whereas the receiver for the on-off binary signal must adjust the decision threshold to be one-half the received signal amplitude.

The preceding equations can be rewritten explicitly in terms of the conditional probabilities

$$p_0(y) = p_{Y|0}(y) \tag{9.42}$$

and

$$p_1(y) = p_{Y|1}(y), \tag{9.43}$$

and the additional observation that

$$p_Y(y) = P_0 p_{Y|0}(y) + P_1 p_{Y|1}(y). \tag{9.44}$$

When written in this form, the probability density functions shown in Fig. 9.8, properly scaled by P_0, P_1, can be plotted together on one set of axes. If the appropriate areas in $p_Y(y)$ are identified as $P_0 P_{\epsilon_0}, P_1 P_{\epsilon_1}$, Eqs. (9.35)–(9.39) follow. This is illustrated in Fig. 9.9. It is often advantageous to use only one set of axes for a given set of conditions. The reader is cautioned, however, that the source digit probabilities P_0, P_1 must be included, as given in Eq. (9.44), before the two probability densities are combined.

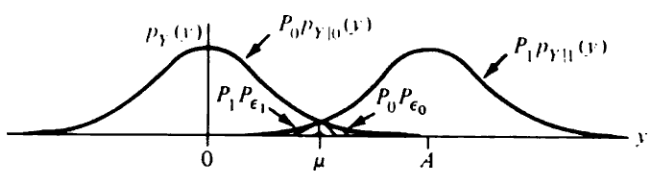


Figure 9.9 Conditional probability density functions for the binary case.

EXAMPLE 9.5.1

The concept of a net probability of error can be extended to the multilevel (M-ary) detection problem, and for gaussian noise the calculation is straightforward. Consider the polar trinary waveform in additive gaussian noise whose probability density functions are shown in Fig. 9.10, representing the trinary digits 0, 1, and 2. Calculate the net probability of error if the decision thresholds are set at $\pm A/4$ and the source probabilities of the digits are equal.

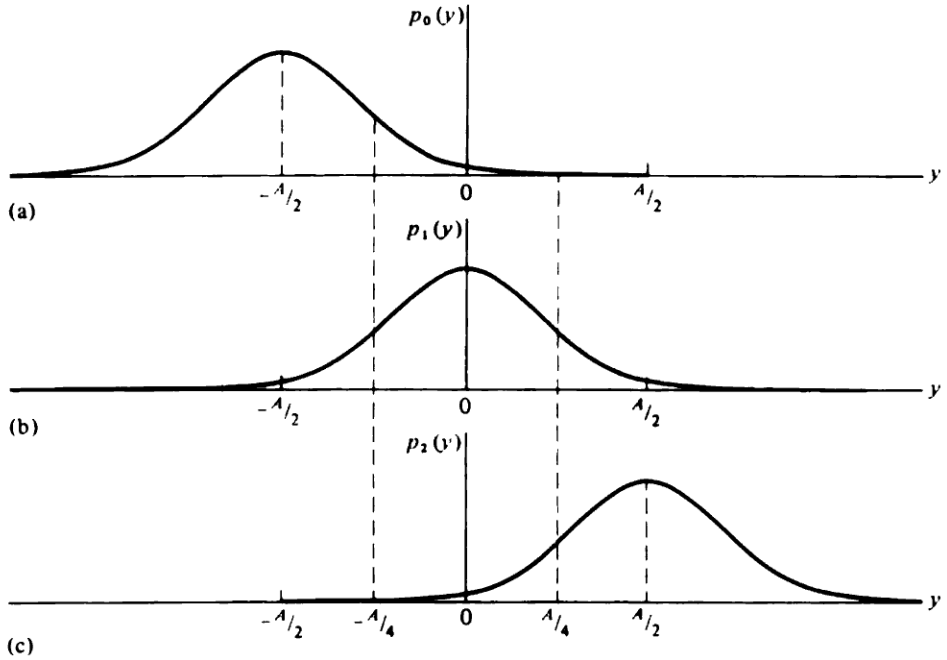


Figure 9.10 Probability density functions for the polar trinary waveform.

Solution Referring to Fig. 9.10, we can write

$$P_{\epsilon_0} = \int_{-A/4}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(y+A/2)^2/2\sigma^2}\,dy = Q[A/(4\sigma)].$$

In a similar manner, we have

$$P_{\epsilon_2} = Q[A/(4\sigma)] \quad \text{and} \quad P_{\epsilon_1} = 2Q[A/(4\sigma)].$$

We are given that $P_0 = P_1 = P_2 = 1/3$, so that

$$P_\epsilon = \tfrac{4}{3}Q[A/(4\sigma)].$$

EXAMPLE 9.5.2

Assuming that errors are made in mistaking adjacent levels only, calculate the net probability of error for the n-level quantizer illustrated in Fig. 9.3 for the case of additive gaussian noise with zero mean and variance $\sigma^2 = N$.

Solution Referring to the result of Example 9.5.1, we note that the factor 4/3 arose because there were $2(n-1)$ areas to be counted in determining the probability of error and each was weighted by $1/n$ because all were equiprobable. Generalizing the result of Example 9.5.1 and letting $a = A/2$, we can write

$$P_\epsilon = \frac{2(n-1)}{n} Q[a/(2\sigma)].$$

To write this in terms of S/N, we use Eq. (9.18) in the above to give

$$P_\epsilon = 2\left(1 - \frac{1}{n}\right) Q\left(\sqrt{\frac{3}{n^2-1} \cdot \frac{S}{N}}\right).$$

DRILL PROBLEM 9.5.1

An on-off binary system uses the signal levels 0 and A for transmission in the presence of gaussian noise. The probability of 1's and 0's is equal. Let the signal amplitude A be K times the standard deviation of the noise. (a) Calculate the net probability of error if $K = 10$. (b) Determine the required value of K to give a net probability of error of 10^{-5}.

ANSWER: (a) 2.87×10^{-7}; (b) 8.53.

DRILL PROBLEM 9.5.2

Assign the equiprobable binary digits 0, 1 to the probability density functions shown in Fig. 9.10(a), (c), reducing the polar trinary case to the polar binary case. With the decision threshold set at zero, repeat the calculations of Drill Problem 9.5.1 for the polar binary case.

ANSWER: Same as those of Drill Problem 9.5.1.

The probability of error is dependent on the signal-to-noise ratio at a given time. Because the matched filter discussed in Section 7.11 is the best linear time-invariant system to maximize the peak-signal–to–rms-noise ratio in the presence of white noise, we now consider this case.

Consider the on-off signal $f(t)$ over $(0, T)$ in the presence of white noise $n(t)$. Letting $y(t)$ be the output of the matched filter, we have

$$y(t) = f_o(t) + n_o(t), \tag{9.45}$$

where $f_o(t)$, $n_o(t)$ are the matched-filter outputs due to input signal and noise, respectively. The matched filter is designed to maximize the signal-to-noise ratio at the time

$t = t_m$ and we shall assume that $t_m = T$. From Eq. (7.53), the output of the matched filter at $t_m = T$ is E, the energy in $f(t)$. Therefore Eq. (9.45) for $f(t)$ present becomes

$$y(T) = E + n_o(T) \qquad \text{(signal present)}. \tag{9.46}$$

On the other hand, if $f(t)$ is absent, the matched-filter output is

$$y(T) = n_o(T) \qquad \text{(signal absent)}. \tag{9.47}$$

A comparison of Eqs. (9.46) and (9.47) with Eqs. (9.25) and (9.26) shows that we have already solved this problem with the exception that A is now replaced by E. Also, assuming that the output noise is gaussian-distributed with zero mean, we have, from Eq. (7.54),†

$$\sigma^2 = \overline{n_o^2(t)} = \frac{E\eta}{2}, \tag{9.48}$$

where $\eta/2$ is the (two-sided) power spectral density of the white noise input to the matched filter. Substituting E for A and $\sqrt{E\eta/2}$ for σ in Eq. (9.39) gives the probability of error for matched-filter detection as

$$P_\epsilon = Q\left(\sqrt{\frac{E}{2\eta}}\right) \qquad \text{(on-off binary)}. \tag{9.49}$$

This is also indicated in Fig. 9.8. From this graph, we observe that the net probability of error begins to decrease very rapidly with increases in signal energy beyond $E/(2\eta) \approx 10$ dB. Above this level, the probability of error decreases approximately by an order of magnitude for each dB increase in signal energy.

The polar binary case follows in the same way except that the output of the matched filter is

$$y(T) = \pm E + n_o(T). \tag{9.50}$$

Substituting $2E$ for A and $\sqrt{E\eta/2}$ for σ in Eq. (9.39) gives

$$P_\epsilon = Q\left(\sqrt{\frac{2E}{\eta}}\right) \qquad \text{(polar binary)}. \tag{9.51}$$

The choice of a matched filter receiver, as used here, is warranted for systems in which additive noise is a more dominant design factor than ISI. If ISI is the more dominant factor, then the overall frequency transfer characteristic of the transmitter filter, channel, and receiver filter should be chosen to minimize the ISI. If the channel characteristics are not completely known, or may vary, an adjustment is usually made at the receiver to correct the overall transfer function to yield minimum ISI. Such an adjustment is called equalization, the subject of a later section. If the effects of both ISI and additive noise must be considered, the following approach proves helpful.

† Because the matched filter is a linear time-invariant system, then $n_i(t)$ is also gaussian-distributed, although we shall not attempt to prove this.

Let $f(t)$ be the pulse waveform generated by a binary source and $g(t)$ is the received pulse waveform, which is assumed to be a Nyquist waveform to minimize ISI. We assume that the transmission system is linear and time-invariant, and that the noise is additive and gaussian (although not necessarily white) with zero mean. A model of the system is shown in Fig. 9.11. Here $H_T(\omega)$ and $H_R(\omega)$ are the transmitter and receiver filters, and $H_C(\omega)$ is the channel filter. We assume that the midband magnitude of $H_C(\omega)$ is normalized to unity, and that channel losses are accounted for by an overall gain factor.

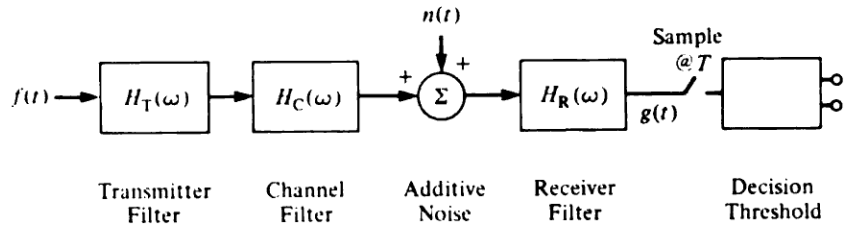


Figure 9.11 Linear channel model for pulse signaling.

We note that $g(0) = 1$ because $g(t)$ is a Nyquist waveform, and assign an amplitude A and a time delay t_d as parameters. Thus the Fourier transform of $g(t)$ is

$$AG(\omega)e^{-j\omega t_d} = F(\omega)H_T(\omega)H_C(\omega)H_R(\omega)\,. \tag{9.52}$$

Our objective here is to optimize the choice of $H_T(\omega)$ and $H_R(\omega)$, subject to the constraint of fixed transmitted energy per symbol:

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 |H_T(\omega)|^2\, d\omega\,, \tag{9.53}$$

while maintaining overall Nyquist pulse shaping. Use of Eq. (9.52) in Eq. (9.53) gives

$$E = \frac{A^2}{2\pi}\int_{-\infty}^{\infty} \frac{|G(\omega)|^2}{|H_C(\omega)|^2 |H_R(\omega)|^2}\, d\omega\,, \tag{9.54}$$

or,

$$A^2 = E\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{|G(\omega)|^2}{|H_C(\omega)|^2 |H_R(\omega)|^2}\, d\omega\right]^{-1}. \tag{9.55}$$

The mean-square noise at the decision threshold is

$$\sigma^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_n(\omega)\, |H_R(\omega)|^2\, d\omega\,. \tag{9.56}$$

The net probability of error for a polar binary system with equiprobable ones and zeros with additive gaussian noise is [cf. Problem 9.5.1]:

$$P_\epsilon = Q(A/\sigma)\,. \tag{9.57}$$

Because $Q(A/\sigma)$ is a monotonically decreasing function of A/σ, the probability of error is minimized if (σ/A) is maximized. Using Eqs. (9.55), (9.56) and finding it convenient to consider the square of this ratio, we have

$$\frac{\sigma^2}{A^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_n(\omega)\,|H_R(\omega)|^2\,d\omega\,\frac{1}{2\pi E}\int_{-\infty}^{\infty}\frac{|G(\omega)|^2}{|H_C(\omega)|^2|H_R(\omega)|^2}\,d\omega\,. \tag{9.58}$$

Identifying the two integrals on the right-hand side of Eq. (9.58) with the two in Eq. (7.46) and then using the Schwarz inequality, we get

$$\frac{\sigma^2}{A^2} \geq \left|\frac{1}{2\pi\sqrt{E}}\int_{-\infty}^{\infty}\frac{\sqrt{S_n(\omega)}\,|G(\omega)|}{|H_C(\omega)|}\,d\omega\right|^2, \tag{9.59}$$

with equality if, and only if [cf. Eq. (7.47)]:

$$|H_R(\omega)|^2 = k_R\frac{|G(\omega)|}{\sqrt{S_n(\omega)}\,|H_C(\omega)|}\,. \tag{9.60}$$

This gives the optimum receiver filter in terms of the square root of the energy density of $g(t)$, the square root of the power spectral density of the additive noise, and the magnitude frequency transfer function of the channel filter. Use of Eqs. (9.52), (9.60) give the corresponding magnitude frequency transfer function of the transmitter filter:

$$|H_T(\omega)|^2 = k_T\frac{|G(\omega)|\sqrt{S_n(\omega)}}{|H_C(\omega)|\,|F(\omega)|^2}\,. \tag{9.61}$$

In the case of white noise, $S_n(\omega) = \eta/2$ and the magnitude squares of the optimum receiver and transmitter filter frequency transfer functions are specified by

$$|H_R(\omega)|^2 = k_R\frac{|G(\omega)|}{|H_C(\omega)|}, \tag{9.62}$$

$$|H_T(\omega)|^2 = k_T\frac{|G(\omega)|}{|H_C(\omega)|\,|F(\omega)|^2}\,. \tag{9.63}$$

Note that if $|F(\omega)|$ does not vary much over the channel bandwidth, one filter design can satisfy both $|H_R(\omega)|$ and $|H_T(\omega)|$.

Finally, the minimum probability of error for optimum filter choices is found by substituting the square root of Eq. (9.59) into Eq. (9.57):

$$P_{\epsilon_{\min}} = Q\left\{\sqrt{E}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sqrt{S_n(\omega)}\,|G(\omega)|}{|H_C(\omega)|}\,d\omega\right]^{-1}\right\}. \tag{9.64}$$

For $S_n(\omega) = \eta/2$, $|H_C(\omega)| = 1$, and noting that $\frac{1}{2\pi}\int_{-\infty}^{\infty}|G(\omega)|\,d\omega = 1$, this gives

$$P_{\epsilon_{\min}} = Q\left(\sqrt{\frac{2E}{\eta}}\right), \tag{9.65}$$

which agrees with Eq. (9.51) for the polar binary case.

9.6 S/N PERFORMANCE OF PCM

We now examine the S/N performance of PCM in the presence of white gaussian-distributed noise (e.g., thermal noise). To simplify things, we make some assumptions that are usually valid in typical operating conditions. First we assume that the probability of error, P_ϵ, is small enough that the probability of more than one bit error in a code word is negligibly small. For example, if $P_\epsilon = 10^{-4}$ and a 10-bit code word ($m = 10$) is used, the probability of more than one error is on the order of (using the Poisson approximation) $1 - P_0 - P_1 = 1 - (1.001)\exp(-0.001) = 5 \times 10^{-7}$.

There are m binary digits to describe $n = 2^m$ levels and a is the amplitude-quantizing increment. We assume that the code words formed by the m binary digits are in order of the numerical significance of the codeword. Using a linear quantization rule, an error that occurs in the least significant bit corresponds to an amplitude error a, an error in the next higher significant bit corresponds to an error $2a$, etc. The mean-square output noise arising from one bit error is then

$$\overline{(\Delta m)^2} = \frac{1}{m}[(a)^2 + (2a)^2 + (4a)^2 + (8a)^2 + \cdots + (2^{m-1}a)^2]. \tag{9.66}$$

The sum of the geometric progression in Eq. (9.66) is

$$\overline{(\Delta m)^2} = \frac{2^{2m} - 1}{3m}a^2. \tag{9.67}$$

The probability that a given bit is in error is P_ϵ and there are m bits per codeword. Therefore the mean-square thermal noise per codeword is

$$\overline{n_{\text{th}}^2(t)} = mP_\epsilon\overline{(\Delta m)^2} = P_\epsilon(2^{2m} - 1)a^2/3. \tag{9.68}$$

Recalling Eqs. (9.18) and (9.20) for the mean-square signal and quantization noise present and using Eq. (9.68), we have

$$\frac{S_o}{N_o} = \frac{\overline{s^2(t)}}{\overline{n_{\text{qnt}}^2(t)} + \overline{n_{\text{th}}^2(t)}}, \tag{9.69}$$

$$\frac{S_o}{N_o} = \frac{2^{2m} - 1}{1 + 4P_\epsilon(2^{2m} - 1)} = \frac{n^2 - 1}{1 + 4P_\epsilon(n^2 - 1)}. \tag{9.70}$$

For the polar binary case, we use Eq. (9.41) in (9.70) to give

$$\frac{S_o}{N_o} = \frac{n^2 - 1}{1 + 4(n^2 - 1)Q(\sqrt{S_i/N_i})} \tag{9.71}$$

where S_i/N_i is measured at the receiver input. A plot of Eq. (9.71) is shown in Fig. 9.12. There are two distinct modes of performance, one limited by the (thermal) transmission noise and the other limited by the quantization noise. PCM systems usually are operated in the region where quantization begins to be the predominant factor.

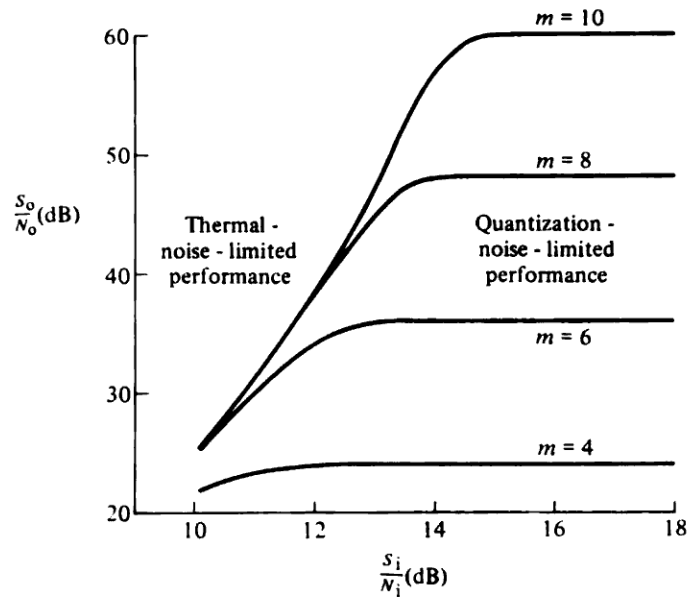


Figure 9.12 Performance of PCM in the presence of noise.

One way to obtain a fairly good qualitative indication of the performance of an operating PCM system is to display the bit stream on an oscilloscope. The time base of the oscilloscope is set to trigger on the bit rate with a sweep duration of a few bit intervals. For no additive noise or bandwidth limitations, the oscilloscope display would appear like Fig. 9.13(a). Figure 9.13(b) illustrates the effects of a limited bandwidth, and some random noise is added to obtain Fig. 9.13(c). The resulting pattern suggests an eye and is called an *eye pattern*. For a given system, the eye closes as more noise is added.

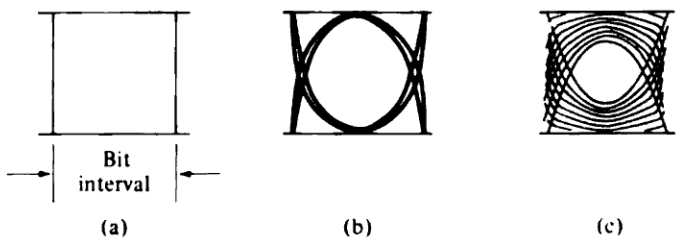


Figure 9.13 Eye patterns for PCM: (a) wideband, no noise; (b) bandwidth limitations, no noise; (c) bandwidth limitations with additive noise.

The best time to sample the received waveform is when the eye opening is largest. The presence of ISI also causes the eye to close, and the relative amount of closure at the best sampling time gives an indication of the degradation caused by ISI. For example, the eye closing by 20% due to the presence of ISI represents an equivalent S/N degradation of $-20 \log_{10} (1 - 0.20) = 1.94$ dB. Thus such a system would require an additional 1.94 dB of signal-to-noise ratio above that required for a specified error rate and zero ISI.

9.7 DELTA MODULATION AND DPCM

Up to this point, we have described PCM systems as those with fairly straightforward digital codes for the transmission of information. There are some alternative PCM systems that have features distinctive enough to warrant special notation. We briefly describe two such types of systems here.

In the transmission of messages having repeated sample values (i.e., an average value), the repeated transmission represents a waste of communication capability because there is little information content in the repeated values. One way to improve this situation is to send only the digitally encoded differences between successive sample values. This is known as *differential pulse-code modulation*, abbreviated as DPCM. For example, DPCM is advantageous in the transmission of sampled picture information because an appreciable portion of the code assigned to each level merely describes the average background level. Thus a picture that has been quantized to 6 bits (i.e., 64 brightness levels per picture element) can be transmitted with comparable fidelity using four-bit DPCM. Block diagrams of a DPCM modulator and demodulator are shown in Fig. 9.14.

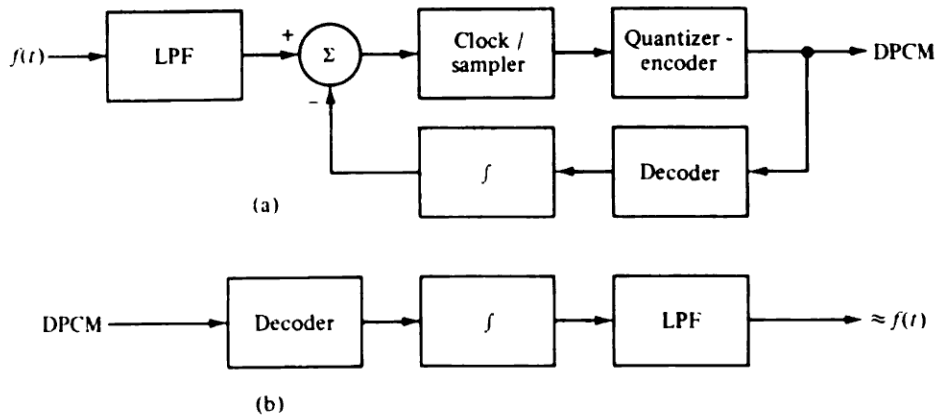


Figure 9.14 Differential pulse code modulation: (a) modulator; (b) demodulator.

Disadvantages of DPCM systems include the fact that if an error is made, an incorrect bias is retained until corrected. Also, such systems suffer from a possible rate-of-rise overloading as a result of the differencing and truncation operations. For instance, if two adjacent picture samples in the above example would differ by more than ± 7 levels (one bit in the four-bit code is reserved for sign), the system could only send a ± 7, and the resulting overload would cause an error in the reconstruction. This problem can be alleviated by placing a linear filter (say, an *RC* filter) before the encoder to limit the maximum rate of rise or fall in the waveform. (This also, of course, decreases the resolution of the reconstructed picture.) It is possible to detect the overload condition and vary the filtering when this condition arises (i.e., nonlinear filtering). Going still farther, one can make the procedure adaptive and dependent on the input data.

Viewed from a wider perspective, the purpose of the feedback path in the DPCM encoder is to perform a prediction of the next sample value. The reasoning behind this is that if the system can do a good job of predicting the next sample value, then there is not as much information needed for the encoder to send (i.e., the encoder has removed some of the redundancy in the data). Considerable research has centered on improving the prediction algorithms. Linear prediction systems using tapped delay lines with adjustable tap gains have been used with success to remove redundancy from the data prior to transmission, but at the expense of some added complexity.

An encoding system that is particularly simple to implement results when a DPCM system is used with only a one-bit output, that bit being used for the sign of the sample difference. This PCM system is known as *delta modulation*, abbreviated either as DM or ΔM. DM systems have an advantage in that the electronic circuitry required for modulation at the transmitter and demodulation at the receiver is substantially simpler than that required for other PCM systems.

A block diagram of a delta-modulation system for line applications is shown in Fig. 9.15(a). The transmitter consists of a comparator/modulator whose input is the difference between its accumulated output and the input signal $f(t)$. The modulator is

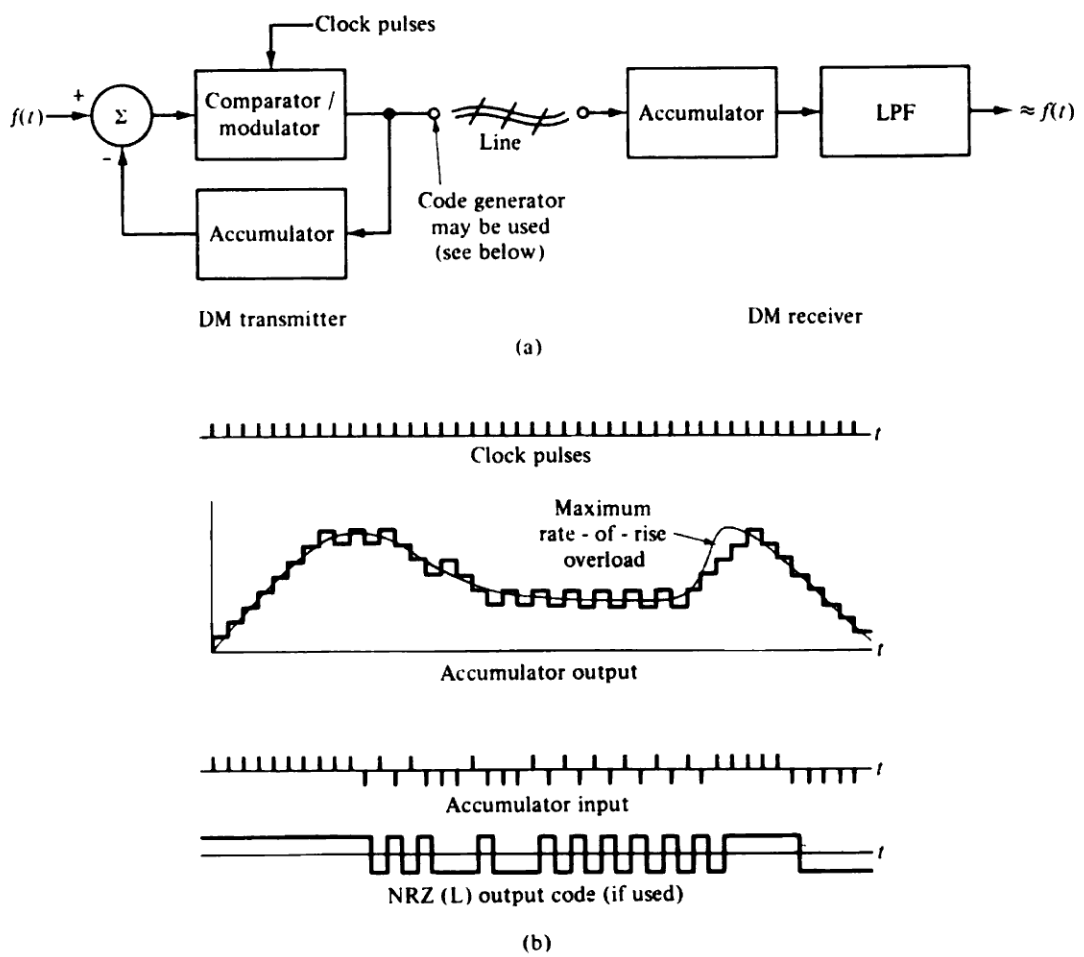


Figure 9.15 A delta-modulation system.

triggered by the clock pulses. If the input-signal amplitude $f(t)$ is greater than the accumulated output at the clock time, a positive output pulse results. Conversely, if $f(t)$ is less than the feedback signal, a negative output pulse is generated. Figure 9.15(b) shows that the output of the accumulator in the transmitter is a stepped approximation to the continuous input signal $f(t)$. Therefore a similar accumulator at the receiver, smoothed by a low-pass filter, will produce an output approximating the original signal $f(t)$. The DM system shown in Fig. 9.15 is intended for line applications where the signal-to-noise ratio is high. For higher noise conditions, the accumulator in the receiver can be preceded by a matched filter and decision threshold.

DM, like DPCM, is subject to rate-of-rise overload problems whenever the input changes too rapidly for the stepped waveform to follow it (see Fig. 9.15b). As a result of the accumulator, it eventually "catches up" and no long-term bias error results if the input is noise-free.

We observe from Fig. 9.15 that if the input-signal level remains constant, the reconstructed DM waveform exhibits a hunting behavior known as *idling noise*. This idling noise is a square wave at one-half the clock rate. If the clock rate is much greater than twice the highest frequency in the input signal, most of the idling noise can be filtered out at the receiver.

An estimate of the rate-of-rise limit, or *slope-overload*, condition for DM may be obtained quite easily for sinusoidal modulation. Let the input be $f(t) = b \cos \omega_m t$ so that

$$\left|\frac{df}{dt}\right|_{\max} = b2\pi f_m . \tag{9.72}$$

If the step size used in the DM system is a, then the maximum rate of rise that can be handled is $a/T = af_s$ so that

$$af_s \geq b2\pi f_m ,$$

$$f_s \geq \frac{2\pi f_m}{a/b} . \tag{9.73}$$

If the spectral weighting is uniform, Eq. (9.73) may be applied to band-limited signals by letting f_m be the highest frequency. If the spectral weighting is not uniform but decreases for higher frequencies, Eq. (9.73) may be used by replacing f_m with an equivalent frequency f_0. For voice signals in telephone applications it is found that $f_0 \approx 800$ Hz can be used satisfactorily.

Small step size is desirable to reproduce the input waveform accurately. However, small step size must be accompanied by a fast clock rate to avoid slope overload. Equation (9.73) states that if $af_s/(bf_m) \geq 2\pi$, then slope overload noise will not be a major contributor to the overall noise. We also must consider the effects of quantization noise.

We shall assume that the quantization noise in DM is uniformly distributed over $(-a, a)$ so that the mean-square quantization error is $a^2/3$. To obtain a measure of the

quantization noise, we also assume that its power spectral density is flat up to $f_s = 1/T$. Filtering this noise to a bandwidth f_m (where $f_m < f_s$), we have

$$\overline{n_{\text{qnt}}^2(t)} = \frac{a^2}{3}\frac{f_m}{f_s}. \tag{9.74}$$

The mean-square value of the signal is $\overline{f^2(t)}$ so that the signal-to-quantization noise ratio for DM is

$$\frac{S}{N} = \frac{3f_s}{a^2 f_m}\overline{f^2(t)}. \tag{9.75}$$

Because $\overline{f^2(t)} = b^2/2$ for sinusoidal signals, using Eq. (9.73) to eliminate a, and assuming a flat input-signal spectrum, we can rewrite Eq. (9.75) as

$$\frac{S}{N} = \frac{3}{8\pi^2}\left(\frac{f_s}{f_m}\right)^3. \tag{9.76}$$

Thus the signal-to-quantization noise ratio increases by $10 \log_{10} 2^3 = 9$ dB for every octave increase in the sampling frequency. This improvement, however, is realized only within the approximations made and cannot be expected to increase indefinitely with higher and higher sampling rates.

Finally, if we use $f_s = 2B$, Eq. (9.76) becomes

$$\frac{S}{N} = \frac{3}{\pi^2}\left(\frac{B}{f_m}\right)^3. \tag{9.77}$$

From Eq. (9.77) we conclude that the S/N performance of DM is superior to angle modulation (for example, FM) and pulse-timing modulation (for example, PPM) but falls short of the exponential characteristic of PCM. This is more noticeable when the number of levels used in PCM becomes large.

More extensive analyses and experimental tests have been made using various classes of signals in DM systems. A typical result for the net signal-to-noise ratio is shown in Fig. 9.16. Considerable research has produced adaptive DM methods that seek to minimize the effects of slope overload and idling noise at the expense of increased complexity. In these methods, the step size is varied automatically in some predetermined manner and in accordance with the time-varying characteristics of the input signal.

Delta modulation is used primarily for speech transmission in telephone and telemetry systems. It has been found that PCM is preferable for high-quality speech transmission, whereas DM is easier to implement and yields transmissions of acceptable quality. One interesting application is to use a DM encoder followed by an up-down binary counter (a counter that increments on positive pulses and decrements on negative pulses). With appropriate reset provisions, this combination can serve as a comparatively simple A/D converter for telephone and telemetry applications.

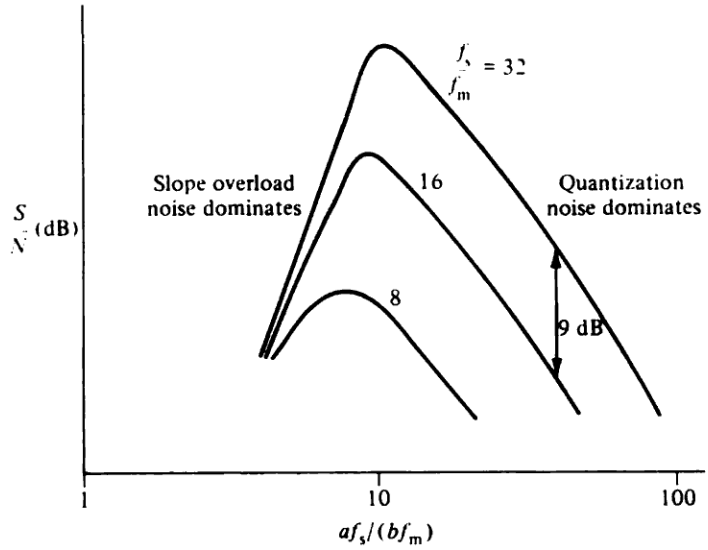


Figure 9.16 Typical output signal-to-noise ratio characteristics for delta-modulation systems.

DRILL PROBLEM 9.7.1

Starting with Eq. (9.75), develop an expression for the S/N performance of DM for typical voice transmissions and calculate for the case $f_0 = 800$ Hz; $f_m = 3.3$ kHz; $f_s = 32$ kHz.

ANSWER: 27.7 dB.

9.8 ★ ERROR ANALYSIS OF PCM REPEATERS

All communication channels cause some attenuation and distortion of the signals that pass through them. To ensure the satisfactory reception of signals at their final destination, repeaters are often provided at appropriate spacings along the transmission route. These provide the amplification and correction of channel distortion (i.e., equalization) necessary to enable satisfactory signal recognition at the receiver.

In analog repeaters, the signal at the end of each segment of the transmission path is equalized and amplified before being sent out again on the next segment of the transmission path. Repeater spacings along the transmission path are governed by path losses and the dynamic range of the repeater. If the repeater itself is not to be a primary source of distortion, its dynamic range is limited by possible nonlinear amplification characteristics at the higher levels.

Net repeater gain is limited also by possible feedback from output to input. In modulated analog repeater systems (e.g., FM) with antennas, feedback can be a primary consideration. To minimize feedback, repeaters may use separate highly directive antennas for reception and transmission. They also may shift the carrier

frequency slightly and filter out the original carrier frequency before transmission in order to obtain better isolation between input and output. (Systems with this operation are called *duplexers*.) Note that in the analog repeater the additive noise effects of all preceding transmission paths (and repeaters) are amplified, as well as the signal, before the composite signal-plus-noise is sent on its way.

A significant advantage for using all-digital transmission is that digital signals lend themselves to periodic conditioning and reshaping. Thus each repeater can be used to regenerate the pulse waveforms at each sample time. This type of repeater is called a *regenerative* repeater. Our primary interest in this section is to make a comparison of relative error rates of systems using analog repeaters and regenerative repeaters in digital transmission. Baseband transmission is assumed, and we ignore the possible modulation and demodulation processes required to transmit the signal through a channel.

Assume that the binary signals are NRZ-polar and that the additive noise present has zero mean value and variance σ^2. Let the signal plus noise at the input to the first analog repeater at a given sample time be

$$y_1 = \pm\frac{A}{2} + n_1 .$$

This signal is amplified to offset the attenuation and sent on to the second repeater. At the input to the second repeater, we have

$$y_2 = \pm\frac{A}{2} + n_1 + n_2 .$$

This continues until ultimately, after m transmission paths, the signal plus noise at the input to the receiver is

$$y = \pm\frac{A}{2} + n_1 + n_2 + \cdots + n_m ,$$

$$y = \pm\frac{A}{2} + n . \tag{9.78}$$

We assume that the additive noise terms are statistically independent so that (cf. Example 8.9.1)

$$\sigma_n^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_m^2 .$$

In addition, we assume that the amplifiers in the analog repeaters are all identical and that amplifier noise dominates the other noise sources. Under these conditions, we have

$$\sigma_n^2 = m\sigma^2 .$$

If, in addition, we assume that $n(t)$ is gaussian-distributed, then the probability of error is exactly the same as that calculated in Section 9.5 with σ^2 replaced by $m\sigma^2$:

$$P_e = \int_0^\infty \frac{1}{\sqrt{2\pi m}\,\sigma} e^{-(y+A/2)^2/(2m\sigma^2)}\,dy ,$$

$$P_e = Q[A/(2\sqrt{m}\,\sigma)] . \tag{9.79}$$

Next we consider a regenerative repeater. Assuming identical repeaters and gaussian noise, the probability of error at each repeater is

$$p = Q[A/(2\sigma)]. \tag{9.80}$$

The probability of making i errors in m trials is given by the binomial distribution [cf. Eq. (8.46)]

$$P_i = \binom{m}{i} p^i (1-p)^{m-i}. \tag{9.81}$$

However, an error is made at the receiver only if an *odd* number of incorrect decisions is made along the total transmission path. The net probability of error is obtained by summing over all values of i, giving

$$P_\epsilon = \sum_{\substack{i=1 \\ i\,\text{odd}}}^{m} \binom{m}{i} p^i (1-p)^{m-i} \tag{9.82}$$

$$= mp(1-p)^{m-1} + \frac{m(m-1)(m-2)}{3!} p^3 (1-p)^{m-3} + \dots$$

$$\approx mp, \tag{9.83}$$

where the approximation applies if $p \ll 1$ and $mp \ll 1$. Using Eq. (9.80) in Eq. (9.83), we have

$$P_\epsilon \approx mQ[A/(2\sigma)]. \tag{9.84}$$

We conclude that, within the approximations given, P_ϵ increases linearly with m for the use of regenerative repeaters.

A comparison of Eqs. (9.79) and (9.84) for several different error rates is shown in Fig. 9.17. Note that the relative advantage in the use of regenerative repeaters becomes more pronounced as p, the error rate for one repeater, decreases.

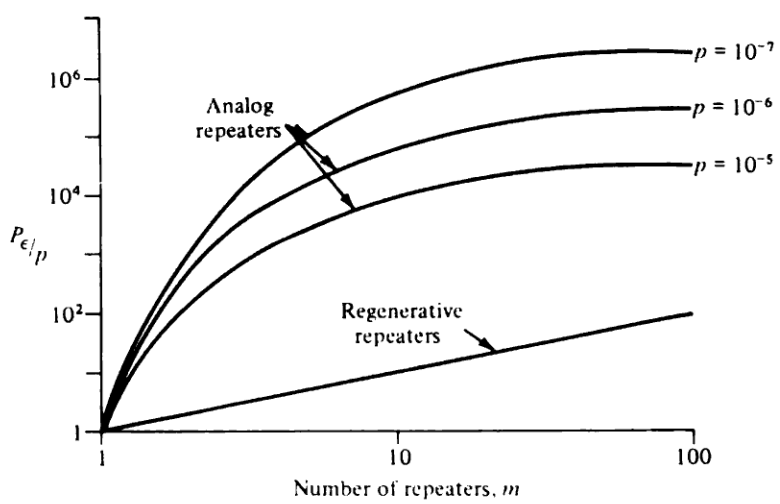


Figure 9.17 Net probability of error as a function of the number of repeaters.

9.9 ★ POWER SPECTRAL DENSITIES OF DATA WAVEFORMS

The power spectral density of the transmitted signal in a PCM system depends on the signaling format used. Major types of PCM signaling formats are discussed in Chapter 7 (see Fig. 7.33). Choice of signaling format is influenced by the transmission characteristics. Because these are often specified in terms of frequency, we shall examine the spectral densities of the PCM signaling formats that are commonly used.

For our purposes here, we assume that 0's and 1's are generated with equal probability by a source and are then encoded into one of the PCM signaling formats shown in Fig. 7.33. The pulse waveform $f_1(t)$ is used to signal a 0, the pulse waveform $f_2(t)$ a 1. The bit interval is $(0, T_b)$, and in each case the pulse waveform is limited to an interval of time equal to T_b. We rely on the results of Examples 8.13.2–8.13.4 to determine the power spectral densities within the above restrictions. In all cases, the power spectral densities are normalized to a 1-ohm resistance.

9.9.1 Nonreturn-to-Zero (NRZ)

In NRZ, the output level is at a specified constant value ($\pm A$) during the full bit interval. For equally probable 1's and 0's, the result of Example 8.13.2 for the transmitted signal $y(t)$ is

$$S_y(\omega) = A^2 T_b \, \mathrm{Sa}^2(\omega T_b/2) . \tag{9.85}$$

This is shown in Fig. 9.18. The $\mathrm{Sa}^2(x)$ function occurs often in digital data systems. For this power spectral density, approximately 77% of the power is contained in the frequency components $(-\pi/2, \pi/2)$, 90% in $(-\pi, \pi)$ and 94.8% in $(-2\pi, 2\pi)$.

Random NRZ data do not have discrete spectral components. Therefore either an external clock or nonlinear processing in the receiver is necessary for bit rate acquisition.

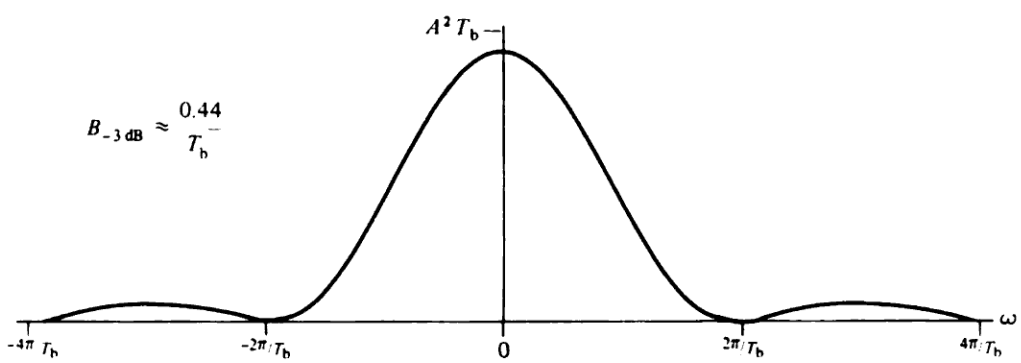


Figure 9.18 Power spectral density of random data (NRZ signaling format).

9.9.2 Return-to-Zero (RZ)

In this signaling format, we have

$$f_1(t) = 0,$$

$$f_2(t) = \begin{cases} A & 0 < t \le \dfrac{T_b}{2} \\ 0 & \dfrac{T_b}{2} < t \le T_b \end{cases}.$$

Using the result of Example 8.13.4 for equally probable zeros and ones, we can write

$$S_y(\omega) = \frac{A^2 T_b}{16}\,\mathrm{Sa}^2(\omega T_b/4) + \frac{\pi A^2}{8}\sum_{n=-\infty}^{\infty}\mathrm{Sa}^2(n\pi/2)\,\delta(\omega - 2\pi n/T_b). \qquad (9.86)$$

This is shown in Fig. 9.19.

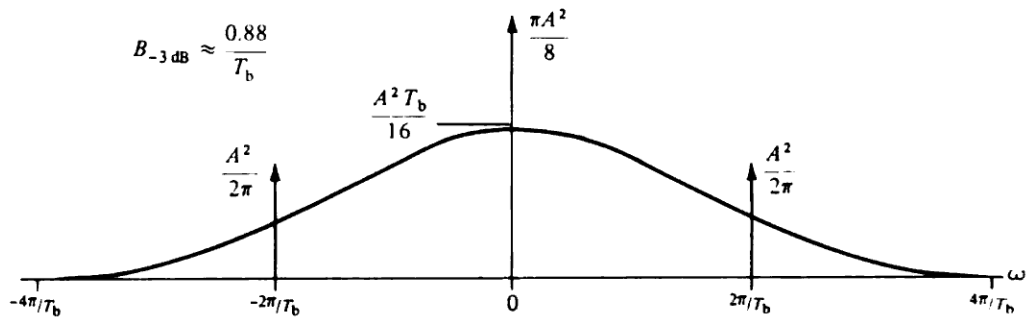


Figure 9.19 Power spectral density of random data (RZ signaling format).

The first zero crossings of the RZ power spectral density are at twice the frequency as those for the NRZ case. In addition to the continuous spectral density, random RZ data have a nonzero average value and discrete lines at odd multiples of the bit rate. Although it requires twice the bandwidth of NRZ, the presence of the discrete lines permits simple synchronization for the RZ signaling format.

9.9.3 Split-Phase (Manchester Code)

In this signaling format, we have

$$f_1(t) = \begin{cases} -A & 0 < t \le \dfrac{T_b}{2} \\ A & \dfrac{T_b}{2} < t \le T_b \end{cases}$$

$$f_2(t) = -f_1(t).$$

Using the result of Example 8.13.4 for equally probable zeros and ones, we can write

$$S_y(\omega) = A^2 T_b \, \mathrm{Sa}^2(\omega T_b/4) \sin^2(\omega T_b/4) . \tag{9.87}$$

This is shown in Fig. 9.20.

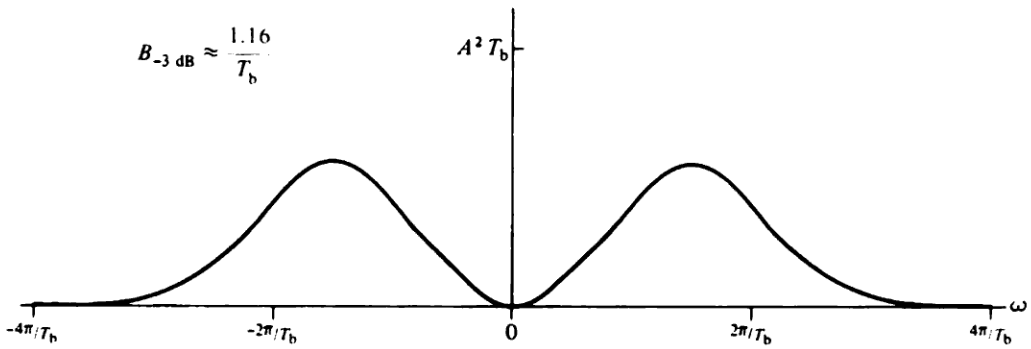


Figure 9.20 Power spectral density of random data (Split-Phase or Manchester code signaling format).

In contrast to that for both RZ and NRZ, the power spectral density for the split-phase signaling format has low relative weighting at low frequencies (and zero at zero frequency). The power spectral density is maximum at a frequency (in Hz) of $0.743/T_b$ and is down by 3 dB at a frequency of $1.16/T_b$. There are no discrete lines, and bandwidth occupancy is similar to that for RZ data at the higher frequencies.

9.9.4 Bipolar Return-to-Zero (BRZ)

The BRZ signaling format uses three amplitude levels: $A, 0, -A$ and is sometimes classified as a pseudoternary signal. For alternate mark inversion, we have

$$f_1(t) = 0 ,$$

$$f_2(t) = \begin{cases} \left.\begin{matrix} A & 0 < t \le \frac{T_b}{2} \\ 0 & \frac{T_b}{2} < t \le T_b \end{matrix}\right\} & \text{first 1 ,} \\ \left.\begin{matrix} -A & 0 < t \le \frac{T_b}{2} \\ 0 & \frac{T_b}{2} < t \le T_b \end{matrix}\right\} & \text{next 1 ,} \\ & \text{etc.} \end{cases}$$

This case can be modeled as a random binary generator whose output is filtered by a one-stage transversal filter to obtain a ternary output (see Example 8.13.3). For

convenience here we assume that the output can take any of the levels: $-1, 0, +1$. Assuming that $P(0) = \frac{1}{2}$, $P(-1) = P(+1) = \frac{1}{4}$,

$$R_{cc}(0) = \sum_{i=-1}^{1} i^2 P(c_k = i) = \tfrac{1}{2}.$$

Correlation between successive output symbols for the alternate mark inversion case can be evaluated with the aid of Table 9.1 to give

$$\begin{aligned} R_{cc}(1) &= \sum_{i=-1}^{1} \sum_{j=-1}^{1} ijP(c_k = i, c_{k+1} = j) \\ &= -\tfrac{1}{4}. \end{aligned}$$

Because the filter has only one stage ($N = 1$) (see Example 8.13.3),

$$R_{cc}(k) = 0 \quad \text{for } k > 1.$$

Using the result of Example 8.13.3 and designating the power spectral density of the filter input by $S_x(\omega)$, we have

$$\begin{aligned} S_y(\omega) &= S_x(\omega)[\tfrac{1}{2} - \tfrac{1}{2}\cos(\omega T_b)] \\ &= S_x(\omega)\sin^2(\omega T_b/2). \end{aligned}$$

Using the result of Example 8.13.2 and scaling the output to the range $(-A, A)$, we have

$$S_y(\omega) = \frac{A^2 T_b}{4} \operatorname{Sa}^2(\omega T_b/4) \sin^2(\omega T_b/2). \tag{9.88}$$

This is shown in Fig. 9.21.

There are no discrete lines in the power spectral density of the BRZ (AMI) signaling format. Bandwidth occupancy is that of NRZ at the higher frequencies.

TABLE 9.1 Enumeration of Joint Probabilities of Successive Symbols for BRZ (Alternate Mark Inversion)

Binary generator output		Transversal filter output		
b_k	b_{k+1}	c_k	c_{k+1}	Probability
0	0	0	0	$\frac{1}{4}$
0	1	0	1	$\frac{1}{4}$
1	0	1	0	$\frac{1}{4}$
1	1	1	-1	$\frac{1}{4}$

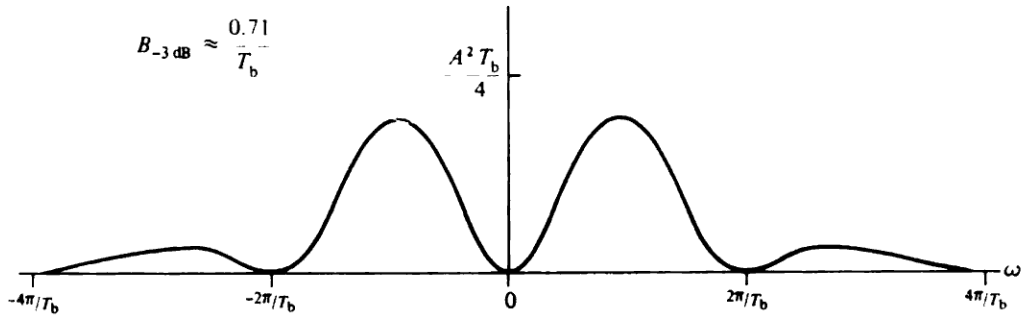


Figure 9.21 Power spectral density of random data (BRZ or alternate mark inversion signaling format).

9.9.5 Delay Modulation (Miller Code)

Derivation of the power spectral density of the delay modulation signaling format is beyond the level of this book.† A graph of the power spectral density is shown in Fig. 9.22. The major spectral weighting is in those frequency components that are greater than zero frequency but less than one-half the symbol rate ($1/T_b$). This makes delay modulation attractive for magnetic tape recording and similar applications that are severely restricted in bandwidth and require relatively low power at very low frequencies.

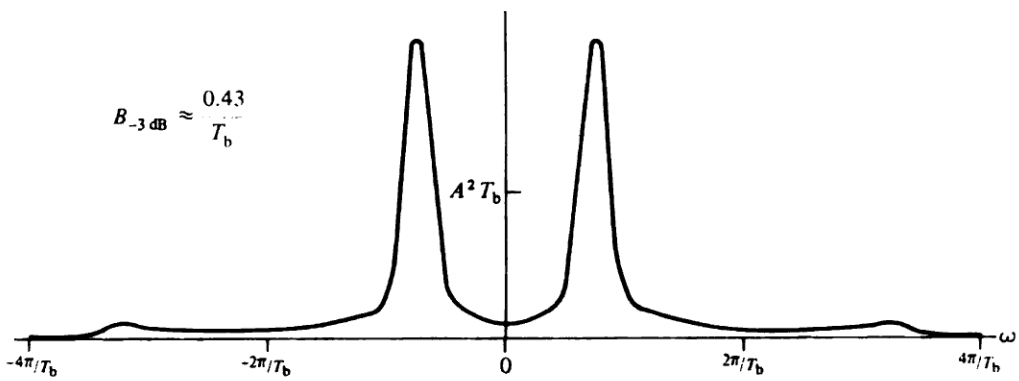


Figure 9.22 Power spectral density of random data (delay modulation or Miller code signaling format).

DRILL PROBLEM 9.9.1

Determine the power spectral density of the RB signaling format (see Fig. 7.33) in which the signal levels are $-A$, $0, A$ and 0's and 1's are equiprobable.

ANSWER: $S_y(\omega) = (A^2T_b/4)\ \mathrm{Sa}^2\ (\omega T_b/4)$.

† Some results are given in W. C. Lindsay and M. K. Simon, *Telecommunication Systems Engineering*, Englewood Cliffs, N.J.: Prentice Hall, Inc., 1973, Chapter 1.

9.10 PARTIAL-RESPONSE SIGNALING

In Chapter 7, we found that transmission rates up to $2B$ symbols per second could be attained within a bandwidth of B Hz. To avoid intersymbol interference (ISI), however, we assumed the use of an ideal low-pass filter to attain this rate. Introduction of raised-cosine spectral shaping to avoid ISI while providing for a more realizable filter characteristic resulted in transmission rates that decreased from $2B$ symbols per second to B symbols per second.

An alternative is to purposely use controlled amounts of ISI by combining a number of successive binary pulses prior to transmission. Because these are combined in a known way, the receiver can isolate the correct data stream again. This approach is called *partial-response signaling* or, alternatively, *correlative coding*. Basically, a partial-response system generates an ℓ-level signal from a binary input data stream x_k using a linear weighted superposition of the x_k. A transversal filter to accomplish this is shown in Fig. 9.23. The superposition memory extends over $(N + 1)$ bit periods and is determined by the number N of delay elements (shift registers) used. The rate of the input data stream is T^{-1}. A low-pass filter (LPF) restricts the bandwidth of the resultant signal to the Nyquist frequency $(2T)^{-1}$ and the ℓ-level output is designated as y_k.

Partial-response signaling permits control of the spectral properties of the signal to make efficient use of available bandwidth and to tailor these properties to the channel. It offers the capability of increasing the transmission rate above B symbols per second but does demand some additional required transmitted power. In addition, known redundancy in the received signal can be used for error detection.

We shall examine two specific cases of partial-response signaling.† In the first case, two successive binary input pulses are added so that

$$y_k = x_k + x_{k-1}. \tag{9.89}$$

This is termed *duobinary* signaling. A transversal filter ($N = 1$) to accomplish this operation is shown in Fig. 9.24(a). The resulting frequency transfer function is

$$H(\omega) = (1 + e^{-j\omega T})_{\text{LP}} \tag{9.90}$$

$$H(\omega) = \begin{cases} 2e^{-j\omega T/2} \cos \dfrac{\omega T}{2} & |\omega| \le \dfrac{\pi}{T} \\ 0 & \text{elsewhere} \end{cases}. \tag{9.91}$$

† For additional coverage, see A. Lender, "Correlative Level Coding for Binary Data Transmission," *IEEE Spectrum*, vol. 3, no. 2 (February 1966): 104–115; E. R. Kretzmer, "Generalization of a Technique for Binary Data Communication," *IEEE Transactions in Communication Technology*, COM–14 (February 1966): 67–68; S. Pasupathy, "Correlative Coding: A Bandwidth-Efficient Signaling Scheme," *IEEE Communications Society Magazine*, vol. 15, no. 4 (July 1977): 4–11.

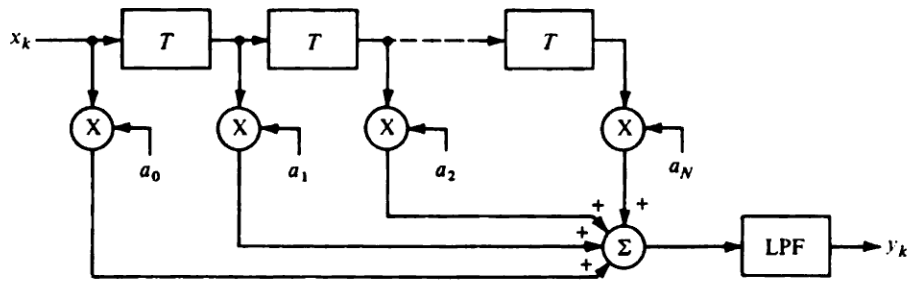


Figure 9.23 Transversal filter for generation of partial-response signaling.

The corresponding impulse response is

$$h(t) = \frac{4 \cos [\pi(t - T/2)/T]}{\pi T[1 - 4(t - T/2)^2/T^2]} \tag{9.92}$$

and this is shown in Fig. 9.24(c). From this result, we conclude that duobinary signaling introduces ISI, but it is controlled in such a way that the interference comes only from the immediately preceding symbol. Note that the operation of adding two successive symbols has provided an effective spectral shaping to the transmitted signal. The total bandwidth needed here, including that for spectral shaping, is $B = 1/(2T)$. This is in contrast to Nyquist pulse shaping in Chapter 7, in which additional bandwidth (up to $B = 1/T$) was needed for spectral shaping.

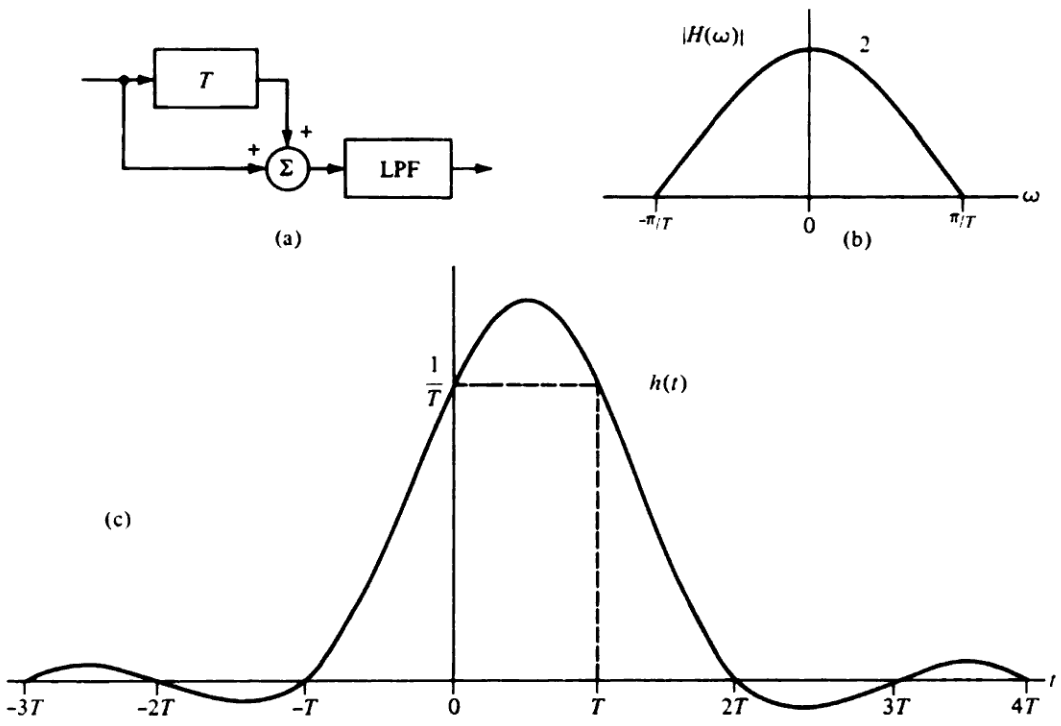


Figure 9.24 Duobinary signaling: (a) transversal filter implementation; (b) magnitude of frequency transfer function; (c) unit impulse response.

Suppose that the binary data stream x_k uses the levels 0, 1 for convenience; then the y_k will assume the levels 0, 1, 2. Also, let the transmitter use a polar signal representation y_k. Without additive noise or attenuation, the received signal, Y_k, will take on one of the three possible levels $-A, 0, A$. For a given amount of additive noise, it was shown in Section 9.5 that discrimination between logic levels at the receiver is proportional to the amplitude separation between levels. Therefore it follows that for a given error rate a duobinary system must transmit more power than a binary system for the same amount of additive noise present. However, the duobinary system permits the use of a higher transmission rate and also furnishes the proper spectral weighting to accomplish it.

The role of the receiver is to solve the equation [cf. Eq. (9.89)]

$$\tilde{x}_k = y_k - \tilde{x}_{k-1}, \tag{9.93}$$

where $\tilde{x}_k$ is the decoded binary data stream. Because $\tilde{x}_k$ is decoded according to the decoded value of $\tilde{x}_{k-1}$, decoding errors tend to propagate in the duobinary system described above. A method proposed by A. Lender eliminates this tendency by precoding the input binary data stream before transmission. In this precoding procedure for duobinary systems, the input binary data stream x_k is converted to another binary stream b_k according to the rule

$$b_k = x_k \oplus b_{k-1} \tag{9.94}$$

where, as before, the symbol $\oplus$ represents modulo-2 (EXCLUSIVE-OR) addition.

The binary stream b_k is applied to the input of the duobinary filtering operation described by Eq. (9.89), and yields

$$y_k = b_k + b_{k-1} = (x_k \oplus b_{k-1}) + b_{k-1}. \tag{9.95}$$

Examination of Eq. (9.95) shows that if $x_k = 1$, then $y_k = 1$ *regardless* of the binary value of b_{k-1}. If $x_k = 0$, then we have $y_k = 0$ or $y_k = 2$. These values are listed in Table 9.2. Thus the y_k can be decoded to find the x_k following the rule

$$\tilde{x}_k = y_k \text{ mod-2}. \tag{9.96}$$

A similar situation occurs if the polar signal representation $Y_k = -A, 0, A$ is used. If $x_k = 0$, then $b_k = b_{k-1}$ [see Eq. (9.94)], and Y_k will be $-A$ or A. If $x_k = 1$, then b_k will be the complement of b_{k-1}, and Y_k will be 0. The corresponding decoding rule is

$$\tilde{x}_k = \begin{cases} 0 & \text{if } Y_k = \pm A \\ 1 & \text{if } Y_k = 0 \end{cases}. \tag{9.97}$$

TABLE 9.2 Decoding of Precoded Duobinary Signals

x_k	b_{k-1}	y_k	Y_k
1	0	1	0
1	1	1	0
0	0	0	$-A$
0	1	2	A

In either case, precoding has enabled the receiver to make each binary decision based only on the current received sample; and, therefore, the possibility of error propagation has been eliminated.

The frequency response of the channel in many transmission systems (e.g., commercial telephone systems) does not extend down to zero frequency. It is advantageous in such systems to modify the duobinary signaling to obtain zero spectral weighting at zero frequency and yet retain the desirable property of allowing signaling up to the Nyquist rate. In one method, called *modified duobinary* signaling, binary pulses spaced by two sample periods are subtracted, giving [see Fig. 9.25(a)]

$$y_k = x_k - x_{k-2}. \tag{9.98}$$

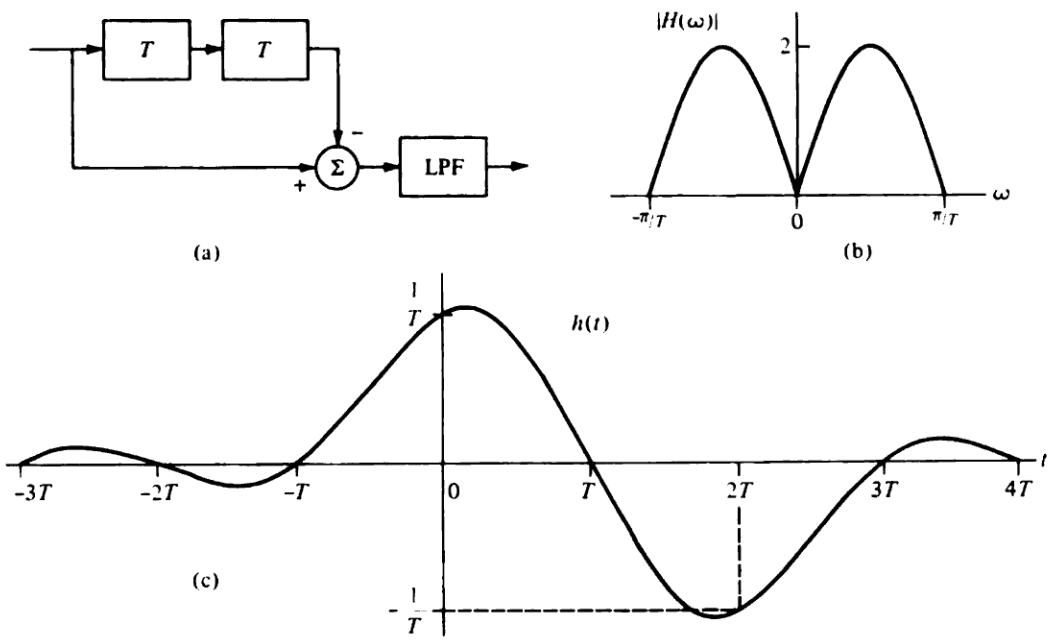


Figure 9.25 Modified duobinary signaling: (a) transversal filter implementation; (b) magnitude of frequency transfer function; (c) unit impulse response.

The resulting frequency transfer function is

$$H(\omega) = (1 - e^{-j2\omega T})_{\text{LP}} \tag{9.99}$$

$$H(\omega) = \begin{cases} 2je^{-j\omega T} \sin \omega T & |\omega| \le \dfrac{\pi}{T} \\ 0 & \text{elsewhere} \end{cases}. \tag{9.100}$$

The corresponding impulse response is

$$h(t) = (1/T)\{\text{Sa}(\pi t/T) - \text{Sa}\,[\pi(t - 2T)/T]\} \tag{9.101}$$

and this is shown in Fig. 9.25(c). From this result, we conclude that modified duobinary signaling introduces controlled ISI from only one symbol removed by two

sample periods. The relatively steep transition in $h(t)$ at $t = T$ makes this signaling method more sensitive to small timing errors.

Again, error propagation becomes a problem unless precoding is used. Precoding used for the modified duobinary case is

$$b_k = x_k \oplus b_{k-2}. \tag{9.102}$$

The binary stream b_k is applied to the modified duobinary operation described by Eq. (9.98), yielding

$$y_k = b_k - b_{k-2} = (x_k \oplus b_{k-2}) - b_{k-2}. \tag{9.103}$$

The various possibilities using Eq. (9.103) are listed in Table 9.3. From this table, we see that the y_k can be decoded to find the x_k if we follow the rule

$$\tilde{x}_k = |y_k| = y_k \text{ mod-2}. \tag{9.104}$$

Similarly, the receiver rule for modified duobinary signaling that corresponds to Eq. (9.97) for the duobinary case is

$$\tilde{x}_k = \begin{cases} 1 & \text{if } Y_k = \pm A \\ 0 & \text{if } Y_k = 0 \end{cases}. \tag{9.105}$$

Again, error propagation is eliminated because each binary decision depends only on the current received value.

Both duobinary and modified duobinary are cases of a wider selection of partial-response signaling methods. These and some others, often referred to by classes of partial response, are listed in Table 9.4. Note that duobinary and modified duobinary are referred to as Class 1 and Class 4 signaling, respectively.

TABLE 9.3 Decoding of Precoded Modified Duobinary Signals

x_k	b_{k-2}	y_k	Y_k
1	0	1	A
1	1	-1	$-A$
0	0	0	0
0	1	0	0

TABLE 9.4 Classes of Partial-Response Signaling

Class	Generating Equation	Magnitude Frequency Transfer Function, $\lvert H(\omega)\rvert$
1	$y_k = x_k + x_{k-1}$	$2 \cos \omega T/2$
2	$y_k = x_k + 2x_{k-1} + x_{k-2}$	$4 \cos^2 \omega T/2$
3	$y_k = 2x_k + x_{k-1} - x_{k-2}$	$[(2 + \cos \omega T - \cos 2\omega T)^2 + (\sin \omega T - \sin 2\omega T)^2]^{1/2}$
4	$y_k = x_k - x_{k-2}$	$2 \sin \omega T$
5	$y_k = x_k - 2x_{k-2} + x_{k-4}$	$4 \sin^2 \omega T$

DRILL PROBLEM 9.10.1

Determine the transmitted data stream corresponding to the input sequence $x_k = 0\,1\,1\,1\,0\,0\,0\,1\,0\,1\,1\,0\,0$, using (a) precoded duobinary signaling, and (b) precoded modified-duobinary signaling. Assume that the precoder b_k is initialized with 1's and that the transmitter levels are designated as $+$, 0, $-$. Also verify that the receiver decoding rule yields the x_k.

ANSWER: (a) $+000---0+00++$; (b) $0--+000-0++00$.

DRILL PROBLEM 9.10.2

Show that the sum of the weight coefficients in the transversal filter in Fig. 9.23 must be zero for a spectral null at zero frequency.

9.11 EQUALIZATION

In preceding sections, steps were described for the improvement of spectral weightings of a transmission channel to minimize intersymbol interference (ISI). In many systems, however, either departures from ideal characteristics are not known or they vary. For example, the characteristics of a telephone channel may vary as a function of a particular connection and line used, as well as possible variations with time. It is advantageous in such systems to include a filter that can be adjusted to compensate for the nonideal characteristics of the transmission channel. Because their purpose is to compensate for imperfect channel transmission characteristics, these filters are called *equalizers*. Although equalizers could be included in the transmitter, receiver, or channel, for the purposes of this discussion we include them in the receiver.

It is difficult to construct equalizers that will compensate exactly for all channel imperfections. Various criteria can be proposed for equalizer designs. We shall examine several types of equalizers that attempt to minimize ISI. However, we must keep in mind that additive noise is also a problem. Because cost is proportional to the number of adjustable parameters in an equalizer, the available number of adjustable parameters must be divided carefully between matched filtering (to minimize the effects of additive noise) and equalization (to minimize ISI). For example, in a high-noise environment with small ISI it might be best to concentrate on matched filtering. On the other hand, in a low-noise environment, such as a telephone channel, the most important consideration is the minimization of ISI.

Because of its versatility and relative ease in implementation, the transversal filter is a popular equalizer design. The transversal filter design lends itself to the compensation of channels for which the time dispersion of pulse signals is only over a small number of symbol intervals. Because this is typical of many data transmission channels (including telephone), the transversal filter is the only equalizer implementation we discuss here. The transversal equalizer consists of a tapped delay line, as shown

in Fig. 9.26. The delay line is tapped at intervals of T seconds, where T is the symbol interval of the input pulse stream.† The output of each tap is weighted by a variable gain factor a_n, and the weighted outputs are added and sampled to form the output. For convenience, we assume here that there are $(2N + 1)$ taps with the corresponding weights $a_{-N}, \ldots, a_0, \ldots, a_N$, as indicated in Fig. 9.26.

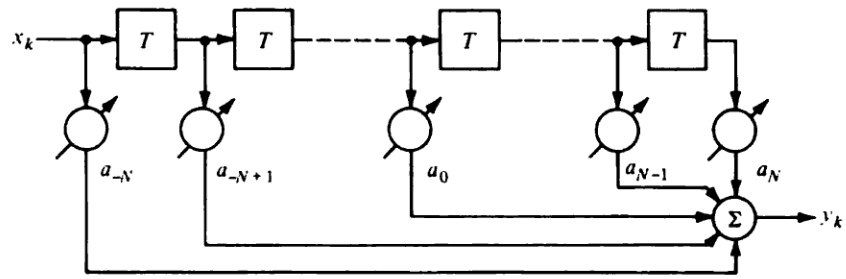


Figure 9.26 Block diagram of a transversal equalizer.

The problem of minimizing ISI is simplified by considering only those signals at correct sample times. Thus we represent the input to the transversal equalizer by $x(kT) = x_k$ and the corresponding output by $y(kT) = y_k$. For zero ISI, we require that

$$y_k = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{elsewhere} \end{cases}. \tag{9.106}$$

The output y_k can be expressed in terms of the input x_k and tap weights a_n as

$$y_k = \sum_{n=-N}^{N} a_n x_{k-n}. \tag{9.107}$$

Equation (9.107) yields $2N + 1$ independent equations in terms of the a_n. This limits us to $2N + 1$ constraints, and therefore Eq. (9.106) must be modified to

$$y_k = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \ldots, \pm N \end{cases}. \tag{9.108}$$

Thus the transversal equalizer can force the output to go to zero at N sample points to either side of the desired peak output (which has been normalized to unity for convenience). For this reason, this is called a *zero-forcing* equalizer. There is no guarantee that the output will be zero at T-second intervals beyond $\pm N$, where $2N + 1$ is the number of taps used. However, it can be shown that this equalizer is optimal in the sense that it minimizes the peak ISI.‡

† The transversal filter can be used for matched filtering also, by spacing the taps on the delay line at Nyquist intervals rather than at symbol intervals.

‡ R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York: McGraw-Hill, 1968, Chapter 6.

We can combine Eqs. (9.107) and (9.108) to exhibit the $2N + 1$ equations that can be solved for the tap weights, a_n:

$$\begin{bmatrix} x_0 & x_{-1} & \cdots & x_{-N} & \cdots & x_{-2N-1} & x_{-2N} \\ x_1 & x_0 & \cdots & x_{-N+1} & \cdots & x_{-2N} & x_{-2N+1} \\ \vdots & \vdots & & & & & \\ x_N & x_{N-1} & \cdots & x_0 & \cdots & x_{-N-1} & x_{-N} \\ \vdots & \vdots & & & & & \\ x_{2N-1} & x_{2N-2} & \cdots & x_{N-1} & \cdots & x_{-2} & x_{-1} \\ x_{2N} & x_{2N-1} & \cdots & x_N & \cdots & x_{-1} & x_0 \end{bmatrix} \begin{bmatrix} a_{-N} \\ a_{-N+1} \\ \vdots \\ a_0 \\ \vdots \\ a_{N-1} \\ a_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{9.109}$$

Computation of required tap weights using Eq. (9.109) is illustrated in the following example.

EXAMPLE 9.11.1

Determine the tap weights of a three-tap, zero-forcing equalizer for the input shown in Fig. 9.27(a), where $x_{-2} = 0.0$, $x_{-1} = 0.2$, $x_0 = 1.0$, $x_1 = -0.3$, $x_2 = 0.1$, $x_k = 0$ for $|k| > 2$.

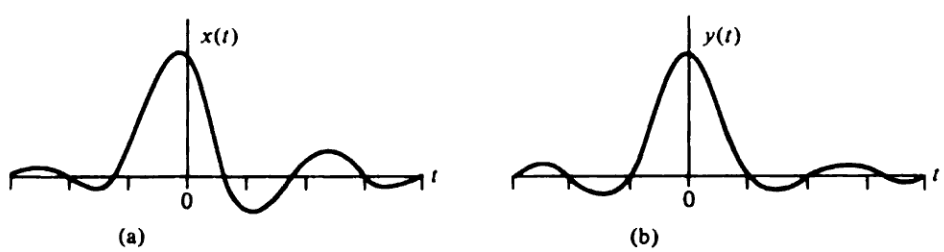


Figure 9.27 (a) Input pulse waveform and (b) output (equalized) pulse waveform for the three-tap, zero-forcing, transversal equalizer in Example 9.11.1.

Solution Writing the $2N + 1 = 3$ equations, the three tap weights a_{-1}, a_0, a_1 are solutions of the following set of simultaneous equations:

$$\begin{aligned} a_{-1} + 0.2a_0 \qquad &= 0 \\ -0.3a_{-1} + \quad a_0 + 0.2a_1 &= 1 \\ 0.1a_{-1} - 0.3a_0 + \quad a_1 &= 0. \end{aligned}$$

Solving, we obtain the required tap weights

$$a_{-1} = -0.1779 \qquad a_1 = 0.2847$$
$$a_0 = 0.8897.$$

Using this three-tap equalizer, we compute the values of the equalized pulse from Eq. (9.107) to find

$$\begin{aligned} y_{-3} &= 0.0 & y_1 &= 0.0 \\ y_{-2} &= -0.0356 & y_2 &= 0.0036 \\ y_{-1} &= 0.0 & y_3 &= 0.0285 \\ y_0 &= 1.0 & y_4 &= 0.0\,. \end{aligned}$$

The resulting equalized pulse waveform is shown in Fig. 9.27(b). Note that this pulse has the desired zeros to either side of the peak, but some ISI has been introduced at sample points farther from the peak. Except in special cases, the finite-N, zero-forcing equalizer cannot completely eliminate ISI. As the number of filter taps become large, however, the ISI can be made small if pulse-dispersion effects caused by the channel do not extend over very many symbol intervals.

DRILL PROBLEM 9.11.1

Let the input samples of a given pulse be $x_0 = 1$, $x_k = 0$ for $|k| > 1$. For this case, develop an approximation for the tap weights of a three-tap, zero-forcing transversal equalizer if $|x_{-1}x_1| \ll 1$.

ANSWER: $a_n \approx (-1)^n x_n \qquad \text{for } n = -1, 0, 1.$

Adjustment of tap weights of the transversal equalizer involves the solution of a set of $2N + 1$ simultaneous equations. Systems have been developed, by using iterative techniques, to adjust the weights automatically. Generally, these adjustments can be divided into two categories. The *preset equalizer* makes use of a special sequence of pulses prior to or during pauses in the transmission of data to determine and adjust the tap weights. In contrast, the tap weights in the *adaptive equalizer* are adjusted during the transmission of data by using the data signal itself. Hybrid systems make use of both of these approaches.

We discuss two methods of equalizer adjustment. The first method is better suited to preset equalizers, whereas the second method is applicable to both the preset and the adaptive equalizer. In the first method of adjustment, the preset equalizer uses widely separated test pulses to set the tap weights prior to data transmission. A simplified block diagram of a three-tap preset equalizer is shown in Fig. 9.28. Upon reception of a test pulse, a binary decision is made in the level slicer at each sample time, and these decision bits are loaded into a shift register, one at a time. The threshold in the level slicer is zero for all sampled values, except that the peak is detected by a peak detector and compared against $+1$ (or whatever reference level is desired). At the end of each test pulse, the binary contents of the shift register are

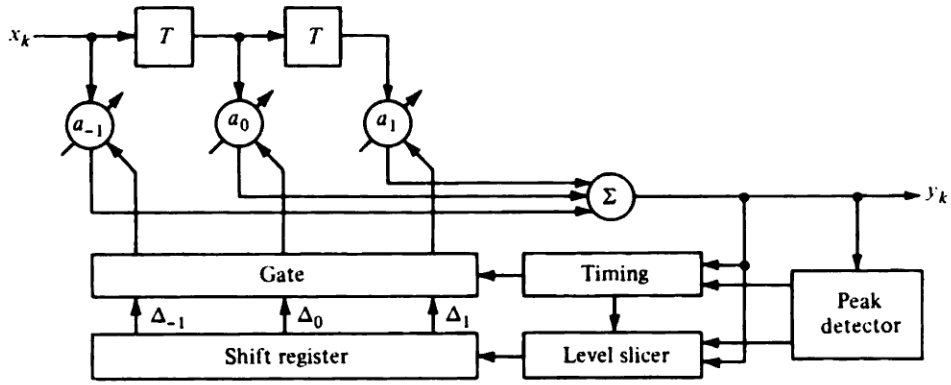


Figure 9.28 Example of a three-tap preset transversal equalizer.

used as correction signals to the respective tap weights in the transversal filter. Because each error signal is simply a binary value (i.e., ± 1), each tap weight is adjusted by a fixed increment, $\pm\Delta$, in each iteration. This iterative "training" algorithm continues, ideally, until all tap weights converge to their optimum values.

The accuracy to which the tap weights are set is proportional to Δ in the fixed-increment adjustment procedure. Equalizer settling time, on the other hand, is inversely proportional to Δ, so that the increment size must be chosen judiciously. The effects of additive noise can be decreased by averaging the generated error signals over several test pulses, but doing so increases the settling time required for convergence within a given accuracy constraint. Refinements to decrease the required settling time include use of the variable-increment adjustment of tap weights and the use of a pseudorandom noise (PN) pulse sequence instead of widely separated pulses in the training phase. A disadvantage of the preset equalizer is that once the training period is completed, the stability of the tap weights must be such that they can be held within the system's desired range of accuracy until the next training period occurs.

The second method for adjustment of the tap weights in a transversal equalizer is to minimize the mean-square error between the output sequence and a known (or desired) sequence. Recall that in the zero-forcing equalizer the tap weights are adjusted to yield N 0's to either side of the peak in the equalized output. In contrast, this second method seeks to minimize the mean-square error over all taps. In practice, the results of both methods are very similar when channel distortions are relatively small. However, they do yield different algorithms for the implementation of automatic equalizers.

To investigate a mean-square minimization criterion, let $c_k = c(kT)$ be a given pulse sequence and let $y_k = y(kT)$ be the output of a transversal filter. Defining an error $e(kT)$ as the difference between these two, we have

$$e(kT) = y(kT) - c(kT). \tag{9.110}$$

The mean-square error for K samples is

$$\overline{e^2} = \frac{1}{K}\sum_{k=1}^{K}[y(kT) - c(kT)]^2. \tag{9.111}$$

A necessary condition on a given tap weight a_n when adjusted to its optimal setting is

$$\frac{\partial \overline{e^2}}{\partial a_n} = 0 . \tag{9.112}$$

Using Eq. (9.111) in Eq. (9.112), we get

$$\frac{\partial \overline{e^2}}{\partial a_n} = \frac{2}{K} \sum_{k=1}^{K} [y(kT) - c(kT)] \frac{\partial y(kT)}{\partial a_n} . \tag{9.113}$$

Then using Eq. (9.107) in Eq. (9.113), we get

$$\frac{\partial \overline{e^2}}{\partial a_n} = \frac{2}{K} \sum_{k=1}^{K} [y(kT) - c(kT)] x(kT - nT) . \tag{9.114}$$

We rewrite Eq. (9.114) as

$$\frac{\partial \overline{e^2}}{\partial a_n} = 2R_{ex}(nT) , \tag{9.115}$$

where R_{ex} is the deterministic cross-correlation of the output error sequence $e(kT)$ and of the input sequence $x(kT)$. Combining Eqs. (9.112) and (9.115), we have

$$R_{ex}(nT) = 0 \quad \text{for } n = 0, \pm 1, \cdots, \pm N . \tag{9.116}$$

Therefore the tap weights are optimum, in a minimum mean-square error sense, when the cross-correlation between the output error sequence $e(kT)$ and the input sequence $x(kT)$ is zero for all integer multiples of the tap-delay increment T.

Now let the input sequence $x(kT)$ be equal to the sequence $c(kT)$; then Eq. (9.114) becomes

$$\frac{\partial \overline{e^2}}{\partial a_n} = \frac{2}{K} \sum_{k=1}^{K} [y(kT) - c(kT)] c(kT - nT) \tag{9.117}$$

$$= 2R_{yc}(nT) - 2R_c(nT) . \tag{9.118}$$

The procedure is that the nth tap weight, a_n, is to be adjusted so that $\partial \overline{e^2}/\partial a_n$ goes to zero. Noting that $R_c(0) = \overline{c^2}$, we form an estimator of the nth tap weight as follows:[†]

$$\hat{\epsilon}_n = \frac{1}{K\overline{c^2}} \sum_{k=1}^{K} [y(kT) - c(kT)] c(kT - nT) . \tag{9.119}$$

Because $y(kT)$ depends upon a_n, this adjustment is an iterative procedure. Also, the estimator must be checked for random inputs. It can be shown that this estimator is an unbiased estimator of the tap weight error if[‡]

$$R_c(nT) = 0 \quad \text{for } n \neq 0 . \tag{9.120}$$

[†] Ibid.

[‡] Ibid.

Equations (9.119) and (9.120) suggest that an efficient means of adjusting a preset equalizer is the use of the minimum mean-square error criterion. Instead of transmitting a sequence of widely separated pulses, a pseudorandom noise (PN) sequence is sent—which takes care of satisfying Eq. (9.120). At the receiver a maximal-length codeword generator can be used to generate the PN sequence. Then Eq. (9.119) can be used to generate an error signal for each tap weight. Use of the PN sequence in this manner, instead of widely separated test pulses, decreases the equalizer settling time required for a given tap-weight accuracy.

In contrast to preset equalizers, the error signals in adaptive equalizers are estimated continuously during the transmission of data. Thus adaptive equalizers can adapt to slowly varying changes during data transmissions and do not require long training periods.

In the presence of data transmission, the random bit sequence needed to adjust the tap weights in an adaptive equalizer can be derived from the binary decisions that the receiver makes on the data sequence. Such a learning procedure is called *decision-directed* because the receiver learns by employing its own decisions. A three-tap adaptive equalizer using this approach is shown in Fig. 9.29. Decision-directed adaptive equalizers do not establish initial equalization easily. Once the correct initial equalization is acquired, however, the error estimates are accurate and the equalization loop tracks changes in channel characteristics unless the changes are faster than the equalizer's time constants.

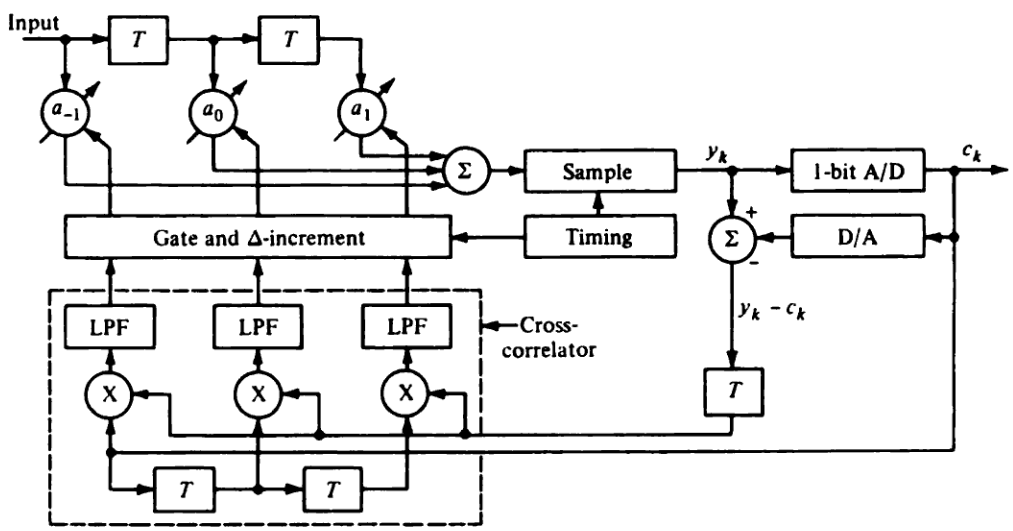


Figure 9.29 Example of a three-tap adaptive transversal equalizer.

A hybrid system—a combination of a preset equalizer and an adaptive equalizer—may be used to gain the best overall performance. In such a system, a PN sequence is sent (and also generated at the receiver) to enable the system to acquire the correct initial equalization values in a preset mode of operation. When the initial equalization is reasonably good, the system switches to an adaptive mode and data

transmission commences. The system may also provide a reset mode in which data transmission pauses temporarily to permit the sending of a PN sequence to verify the correctness of equalizer settings.

Equalizers based on the principles in the preceding discussion do a good job of minimizing ISI in channels where additive noise is not of major concern. They are essential parts of modern high-speed modems for the transmission of digital data using telephone channels.

9.12 *M*-ARY SIGNALING

In the PCM communication systems that we have considered up to this point, one of only two possible signals is transmitted during each T-second signaling interval. We call a signal waveform that might be sent in a T-second interval a *symbol*. In a binary transmission system, only two symbols are used (e.g., $\pm A$). The symbol rate (*baud*) is $(1/T)$ symbols per second. For example, in a 300-baud system, symbols are sent at a rate of 300 symbols per second.

In M-ary signaling, one of M possible symbols is transmitted during each T-second interval. In baseband transmission systems, each symbol corresponds to one of M distinct levels. However, M-ary transmissions are not restricted to amplitude-level variations. We could, for example, vary the amplitude, frequency, and phase of a sine wave to form the various symbols. The subject of digital modulation is introduced and discussed in Chapter 10.

It is convenient to envision all sampling and multiplexing operations being carried out by binary signals and then a composite bit stream being applied to a M-ary encoder prior to transmission. This also turns out to be the usual case for operating systems. For example, in a quaternary ($M = 4$) transmission we could assign the binary groupings 00, 01, 10, and 11 to levels A, B, C, and D, respectively. Each of possible output levels A, B, C, D, is transmitted within a T-second interval.

If a channel bandwidth of B Hz is available for transmission, the Nyquist rate is $1/T = 2B$ symbols per second. In M-ary signaling, each symbol represents $\log_2 M$ bits of information and may be uniquely coded into $\ell = \log_2 M$ levels ($M = 2^\ell$). It follows that the equivalent bit rate is $2B\ell$ bps. Thus the bandwidth required for M-ary signaling with a fixed information rate is inversely proportional to ℓ. For example, let B_2 be the bandwidth required for a binary PCM system; then $B = B_2/\ell$ is the bandwidth required for a M-ary system to transmit information at the same rate as a binary system.

Because the bandwidth required for transmission is proportional to $1/T$ (baud), but the information transmitted is proportional to the bit rate $(\log_2 M)/T$, we see that M-ary systems provide a means for increasing the rate-of-information transmission within a given bandwidth. However, this increased information rate comes at the expense of added transmitter power and increased system complexity. Moreover, intersymbol interference is also a problem in M-ary signaling systems, and the bandwidths

required for operating these systems are greater than the theoretical minimums discussed previously. Mathematical results are difficult to obtain for the effects of ISI in M-ary systems, however, and these effects are often studied using system simulation studies and computer programs to generate multilevel eye diagrams.†

In an analysis of M-ary signaling, we assume that the spacing of amplitude levels is uniform and that the M amplitude levels are centered at zero. For convenience, we assume also that M is an even integer (the usual case is that M is a power of 2, or $M = 2^\ell$, where ℓ is an integer). Designating the spacing between adjacent levels by A, we find the various levels are

$$A_j = \pm\frac{A}{2}, \pm\frac{3A}{2}, \pm\frac{5A}{2}, \ldots, \pm\frac{(M-1)A}{2}. \tag{9.121}$$

Assuming that all M levels are equiprobable, we obtain the average signal power [cf. Eq. (9.18)]

$$S = \frac{2}{M}\left\{\left(\frac{A}{2}\right)^2 + \left(\frac{3A}{2}\right)^2 + \cdots + \left[\frac{(M-1)A}{2}\right]^2\right\}$$

$$S = \frac{M^2-1}{3}\left(\frac{A}{2}\right)^2. \tag{9.122}$$

Therefore, for a given level spacing, the transmitted power increases approximately in proportion to the square of the number of levels.

Next, we assume that $M = 4$ (i.e., a quaternary system) and that the signaling levels are $-3A/2, -A/2, A/2, 3A/2$. For convenience, we let the corresponding symbols be labeled A, B, C, D, respectively. We also let the additive noise present be gaussian-distributed with zero mean value and variance σ^2. The probability density function for the case in which all four symbols are equiprobable is shown in Fig. 9.30. Because the symbols have been chosen to be equiprobable, the optimum threshold levels are midway between levels. Noting that there are six equal areas in Fig. 9.30 to be included in the probability-of-error calculation, we have

$$P_\epsilon = (6)\left(\tfrac{1}{4}\right)\int_A^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-[y-(A/2)]^2/(2\sigma^2)}\,dy. \tag{9.123}$$

With the change of variable, $z = [y - (A/2)]/\sigma$, this expression becomes

$$P_\epsilon = \frac{3}{2}\int_{A/(2\sigma)}^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-z^2/2}\,dz \tag{9.124}$$

$$P_\epsilon = \frac{3}{2}Q\left(\frac{A}{2\sigma}\right). \tag{9.125}$$

A comparison of Eq. (9.125) with Eq. (9.39) for the binary case reveals that the two expressions are identical except for the factor of 3/2. This factor arose from counting

† See, for example, K. Feher, *Digital Communications: Microwave Applications*, Englewood Cliffs, N.J.: Prentice Hall, 1980.

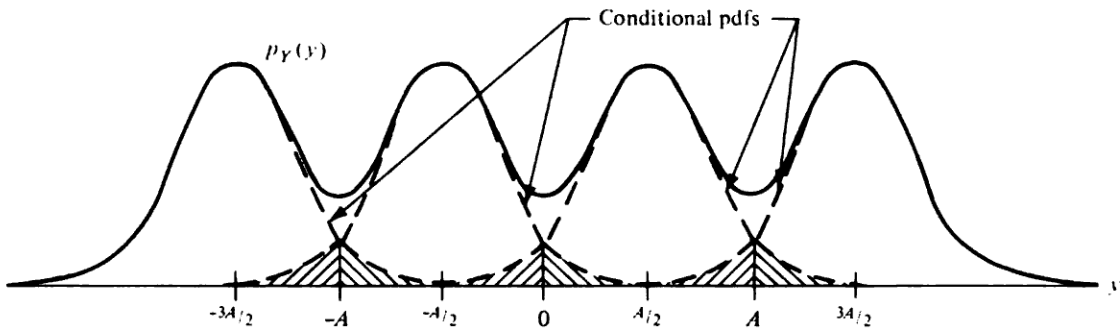


Figure 9.30 Probability density function for signal plus noise in an M-ary system, $M = 4$.

the areas indicated in Fig. 9.30 and multiplying this number by $1/M$ (because all symbols are equiprobable). Noting that in general there are $2(M - 1)$ areas to be counted, we see that Eq. (9.123) can be extended to the M-ary case to give

$$P_\epsilon = \frac{2(M-1)}{M} Q\left(\frac{A}{2\sigma}\right). \tag{9.126}$$

Therefore the probability of error increases slowly with an increase in number of levels.

A third comparison of interest is that of bandwidth. Designating the bandwidth of a binary system B_2, and the corresponding bit interval T_b, we have

$$B_2 = \frac{1}{2T_b}, \tag{9.127}$$

and

$$B = \frac{B_2}{\log_2 M} = \frac{B_2}{\ell}. \tag{9.128}$$

From Eqs. (9.122), (9.126), and (9.128), we conclude that, in M-ary signaling at a given information rate, the required bandwidth is inversely proportional to $\log_2 M$, the transmitted power is proportional to $M^2 - 1$, and the probability of error varies over a maximum of one to two times that of an equivalent binary system. These dependencies are illustrated in the following example.

EXAMPLE 9.12.1

Compare the required transmitter power and minimum bandwidth for those M-ary systems for which $1 \le \ell \le 5$. Assume additive white gaussian noise with two-sided power spectral density $\eta/2$ W/Hz and a probability of symbol error $P_\epsilon = 10^{-4}$. Normalize all results to the minimum bandwidth required for an equivalent binary system B_2.

Solution Noting that $M = 2^\ell$ and $\sigma^2 = \eta B$, we see that the equations that must be solved are

$$B = \frac{B_2}{\ell}$$

$$10^{-4} = \frac{2(M-1)}{M} Q\left(\frac{A}{2\sigma}\right)$$

$$S = \frac{M^2 - 1}{3}\left(\frac{A}{2\sigma}\right)^2 \eta B .$$

Use of the equations above and Appendix I gives the results that are listed in Table 9.5. Note that in this range the average power required rises by about 4–5 dB for every twofold increase in the number of levels used.

TABLE 9.5 Comparison of M-ary Signaling for $P_\epsilon = 10^{-4}$

ℓ	M	B/B_2	$A/(2\sigma)$	$[S/(\eta B_2)]$ dB
1	2	1	3.72	11.4
2	4	$\frac{1}{2}$	3.82	15.6
3	8	$\frac{1}{3}$	3.86	20.2
4	16	$\frac{1}{4}$	3.88	25.0
5	32	$\frac{1}{5}$	3.88	30.1

If the input to the M-ary encoder is binary, it is more meaningful to use for comparison the bit-error rather than the symbol-error probability. While there are no unique relationships between the bit-error probability and the M-ary symbol-error probability, we can obtain relationships for two special cases.

In the first case we assume that whenever the receiver is in error for an M-ary symbol, the output is equally likely to be any one of the $M - 1$ erroneous ℓ-bit codewords. Using a standard binary code (cf. Table 9.6), we find that the sum of the Hamming distances from one ℓ-bit word to all of the others in the code is $\ell 2^{\ell-1}$. The total number of bits in all of the erroneous words is $(M-1)\ell = (2^\ell - 1)\ell$. If we assume all are equally likely, the average probability of bit error, P_{be}, is

$$P_{be} = \frac{\ell 2^{\ell-1}}{(2^\ell - 1)\ell} P_\epsilon$$

$$P_{be} = \frac{2^\ell}{2(2^\ell - 1)} P_\epsilon . \tag{9.129}$$

Note that

$$\frac{1}{2} \le \frac{P_{be}}{P_\epsilon} \le \frac{2}{3}$$

for this case.

TABLE 9.6 Four-Bit Binary Code and Four-Bit Gray Code

Decimal No.	Level	Binary Code		Gray Code	
0	$15A/2$	0000		0000	
1	$13A/2$	0001		0001	
2	$11A/2$	0010		0011	
3	$9A/2$	0011		0010	
4	$7A/2$	0100		0110	
5	$5A/2$	0101		0111	
6	$3A/2$	0110		0101	Reverse-order
7	$A/2$	0111	Image, except for first bit.	0100	image, except for first (sign) bit.
8	$-A/2$	1000		1100	
9	$-3A/2$	1001		1101	
10	$-5A/2$	1010		1111	
11	$-7A/2$	1011		1110	
12	$-9A/2$	1100		1010	
13	$-11A/2$	1101		1011	
14	$-13A/2$	1110		1001	
15	$-15A/2$	1111		1000	

One difficulty with the standard binary code is that the Hamming distance varies (from 1 to ℓ) between one adjacent level and another. In M-ary signaling, we would prefer to use a code in which only one binary digit varied as the amplitude changed from one adjacent level to the next. In other words, the code words for adjacent levels should be separated by a Hamming distance of 1. The Gray code shown in Table 9.6 is an example of such a code.†

In the second case of the relationship between bit-error and symbol-error probability, we assume that the input to the M-ary encoder is coded so that the binary codewords for adjacent amplitude levels differ in only one binary digit (as in the Gray code). We assume also that the symbol probability of error is low, so that it is very unlikely that an error of more than one level will be made. Within these assumptions, we have

$$P_{be} = \frac{1}{\ell} P_e \qquad \text{(Gray level coding)}. \tag{9.130}$$

Equations (9.129) and (9.130) can be used to give some approximate answers for comparison purposes.

† The binary digits in the Gray code, g_k, can be found from those in the corresponding binary code b_k from

$$g_k = \begin{cases} b_k & k = 1 \\ b_k \oplus b_{k-1} & k > 1 \end{cases}.$$

DRILL PROBLEM 9.12.1

Repeat Example 9.12.1, except use $P_{be} = 10^{-4}$ and a Gray code.

ANSWER: See Table 9.7.

TABLE 9.7 Comparison of M-ary Signaling for $P_{be} = 10^{-4}$ and Gray Code

ℓ	M	B/B_2	$A/(2\sigma)$	$[S/(\eta B_2)]$ dB
1	2	1	3.72	11.4
2	4	$\frac{1}{2}$	3.65	15.2
3	8	$\frac{1}{3}$	3.58	19.5
4	16	$\frac{1}{4}$	3.52	24.2
5	32	$\frac{1}{5}$	3.47	29.2

9.13 ★ CODING FOR RELIABLE COMMUNICATION

In binary pulse code modulation, each sample of the signal is represented by a codeword of, say, k bits. These bits are transmitted and the role of the receiver is to recognize each codeword in order to reconstruct the samples. However, errors may occur in transmission as a result of noise. One way to improve the reliability of communication in the presence of noise is to increase the signal-to-noise ratio. An alternative is to add extra bits, at the expense of an increase in bandwidth, to detect and possibly even to correct the errors. This section is intended to introduce the interested reader to the latter approach. We shall restrict our attention to the binary case.

Suppose that we wish to transmit a signal that has been quantized to 16 equiprobable levels. In this case each codeword will consist of $k = 4$ bits. To each codeword we add r bits, which will be used to check for errors, and possibly to correct them. These added bits are redundant because they bear no additional information about the signal. Each codeword is then composed of n bits in the following manner:

$$k = \text{number of message digits per word;}$$

$$r = \text{number of check digits per word;}$$

$$n = k + r = \text{total number of digits per word.} \tag{9.131}$$

Such a code is referred to as a (n, k) code. The code rate efficiency is defined as k/n and is an indicator of the information rate relative to the bit rate in the code.

9.13.1 Algebraic Codes†

One type of code can be formed by adding extra binary digits at the end of each codeword for error detection—and, possibly, error correction. These extra bits are called parity-check bits.

A general codeword for a parity-check code can be written in the form

$$a_1 a_2 a_3 \cdots a_k c_1 c_2 \cdots c_r$$

where a_i is the ith bit of the message codeword and c_j is the jth parity-check bit. The check bits are chosen to satisfy the $r = n - k$ linear equations:

$$\begin{aligned} 0 &= h_{11}a_1 \oplus h_{12}a_2 \oplus \cdots \oplus h_{1k}a_k \oplus c_1 \\ 0 &= h_{21}a_1 \oplus h_{22}a_2 \oplus \cdots \oplus h_{2k}a_k \oplus c_2 \\ &\vdots \\ 0 &= h_{r1}a_1 \oplus h_{r2}a_2 \oplus \cdots \oplus h_{rk}a_k \oplus c_r . \end{aligned} \tag{9.132}$$

These equations can be expressed more conveniently in terms of an $n \times 1$ column matrix $[T]$ representing the codeword

$$[T] = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \\ c_1 \\ \vdots \\ c_r \end{bmatrix} \tag{9.133}$$

and a rectangular $r \times n$ parity-check matrix $[H]$,

$$[H] = \begin{bmatrix} h_{11} & h_{12} \cdots h_{1k} & 1 & 0 \cdots 0 \\ h_{21} & h_{22} \cdots h_{2k} & 0 & 1 \cdots 0 \\ \vdots & \vdots \quad \vdots & & \vdots \\ h_{r1} & h_{r2} \cdots h_{rk} & 0 & 0 \cdots 1 \end{bmatrix} \tag{9.134}$$

so that Eq. (9.132) can be written as

$$[H][T] = 0 . \tag{9.135}$$

Now let the received codeword be $[R]$, which may or may not be equal to $[T]$. If $[H][R] = 0$, we know that $[R]$ is a codeword—and most likely the transmitted

† This section requires some knowledge of matrix notation.

codeword. On the other hand, if $[H][R] \neq 0$, then $[R]$ is not a codeword and at least one error has been made.

The codeword $[R]$ may be written in terms of the transmitted codeword by introducing an $n \times 1$ error matrix $[E]$:

$$[R] = [T] \oplus [E]. \tag{9.136}$$

If $[E]$ contains all zeros, no error has been made.

In order to correct the errors we need to determine $[E]$. As a first step we determine a matrix $[S]$—called the *syndrome*—from the received codeword and the parity-check matrix:

$$\begin{aligned}[S] &= [H][R] \\ &= [H][T] \oplus [H][E] \\ &= [H][E]. \end{aligned} \tag{9.137}$$

The syndrome $[S]$ has dimensions $r \times 1$ and can be any of 2^r sequences (including the all-zero sequence). The error sequence $[E]$ can be any of $2^n = 2^{r+k}$ sequences. Therefore the syndrome cannot unambiguously determine the error sequence. The task of the decoder is to select one of several error sequences which are associated with a given syndrome. Ordinarily, this selection is based on a minimum-distance criterion. If there is a single error, the error matrix $[E]$ will have one nonzero entry and the multiplication of Eq. (9.137) will result in an exhibition of the corresponding column of $[H]$. The position of the error in the codeword is then known by recognition of the appropriate column in $[H]$. This is illustrated in the examples below.

Hamming codes are those parity-check codes for which the columns of $[H]$ consist of all the distinct nonzero r sequences of binary numbers. Thus a Hamming code has as many parity-check matrix columns as there are single-error sequences; i.e.,

$$r + k = n = 2^r - 1. \tag{9.138}$$

Hamming codes are capable of correcting all single-error sequences, and every possible received sequence is either a codeword or is a distance of one from a codeword.

The codes that we have discussed are examples of *block* codes. In an (n, k) block code, the $r = n - k$ redundant bits are combined with the k message bits to give an n-digit codeword. Each n-digit codeword is encoded and decoded independently from any other codeword; n is called the *block length* of the codeword.

EXAMPLE 9.13.1

A given (7, 4) Hamming code has the following parity-check matrix:

$$[H] = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

a) Determine the codeword for the message code 0011.

b) If the received codeword is 1000010, determine if an error has been made. If it has, find the correct codeword.

Solution a) The codeword is 0011 $c_1c_2c_3$ and we can write (all operations are modulo–2)

$$[H][T] = \begin{bmatrix} 1 + c_1 \\ 1 + c_2 \\ 0 + c_3 \end{bmatrix}.$$

Using Eq. (9.135), we find that the codeword is 0011110.

b) Computing the syndrome, we have

$$[S] = [H][R] = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Because the syndrome is the third column of the parity-check matrix, the third position of the received codeword is in error and the corrected codeword is 1010010.

EXAMPLE 9.13.2

The columns of the parity-check matrix for the Hamming code may be chosen in any order as long as each is distinct and no column has all zeros. Repeat Example 9.13.1 for the (7, 4) Hamming code, which is chosen so that the columns of the parity-check matrix are the binary representations of the successive integers $1, 2, 3, \ldots, n$.

Solution a) The parity-check matrix is found by writing the binary representations of $1, 2, \ldots, 7$ in successive columns so that

$$[H] = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

Therefore a general codeword is $c_3c_2a_1c_1a_2a_3a_4$. The parity equations are

$$[H][T] = \begin{bmatrix} c_1 + 0 \\ c_2 + 0 \\ c_3 + 1 \end{bmatrix} = 0,$$

so that the codeword is 1000011.

b) Computing the syndrome, we have

$$[S] = [H][R] = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Thus the error is in the seventh position, and the decoded codeword is 1000011. Note that for a single error the syndrome in this case is the binary representation of the position in the codeword which is in error.

9.13.2 Convolutional Codes

In contrast to block codes, the parity-check information in convolutional codes is distributed over a span of message symbols, called the *constraint span* of the code. In this way, long streams of message bits can be encoded continuously without the necessity of grouping them into blocks. This is accomplished by using shift registers whose outputs are combined in a preset manner to give certain constraints within the encoded bit stream.

A shift register is a cascaded series of one-stage binary memories. The binary state of each memory is transferred to the next at the command of a clock that is synchronized to the input bit stream. Thus the output of a four-stage shift register is the input delayed by four clock intervals, and at any one time this shift register will contain the four successive most recent bits in the input data stream.

A convolutional code of constraint span K can be generated by combining the outputs of K shift registers with v modulo–2 adders. At each clock time the outputs of the v adders are sampled by a commutator. Thus v output symbols are generated for each input symbol, giving a code of rate $1/v$.

An example of a convolutional encoder for $K = 4$, $v = 3$ is shown in Fig. 9.31. The binary state is available at the output of each stage until it is shifted into the next stage. The equations for these three adders are

$$\begin{aligned} v_1 &= S_1 , \\ v_2 &= S_1 \oplus S_2 \oplus S_3 \oplus S_4 , \\ v_3 &= S_1 \oplus S_3 \oplus S_4 . \end{aligned} \tag{9.139}$$

Note that not all stages are connected to all adders. The connections influence the code that is generated and their selection is beyond the level of this discussion.

We demonstrate the convolutional encoding by assuming that the input data stream begins with 10110 and that all shift registers are initially at zero. The first message digit to enter is 1 and the corresponding output is 111 (if the first message digit were 0, the output would be 000). The second message digit is 0 and the corresponding output is 010 (if the second message digit were a 1, the output would be 101). Note, however, that if the first message digit had been 0, then the output would have been 000 if the second digit were 0, or 111 if the second digit were 1. Therefore the generated code depends on previous message digits within the constraint span of the code.

Next we consider the decoder problem. At each increment in the message code, the three-bit output code in the above example yields eight possible choices. Some of these choices can be ruled out immediately as impossible for a given encoder. The

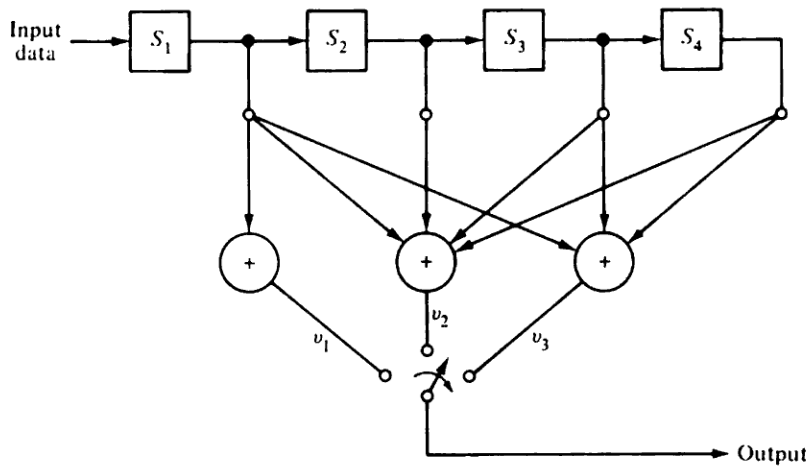


Figure 9.31 Example of a rate $\frac{1}{3}$ convolutional encoder.

remaining ones can be examined in reference to the previous digits within the constraint span of the code. In the absence of noise, this procedure is quite straightforward and involves making decisions after each group of $v = 3$ input digits. This procedure involving successive branches can be diagrammed as a "code tree." A code tree for the convolutional encoder in Fig. 9.31 is shown in Reference 2.

In the presence of noise, the optimum procedure is similar. It examines the possible code choices in terms of the conditional probabilities between each choice within the constraint span of the code. A decoding algorithm using this approach, known as the Viterbi algorithm,† is popular for codes of short constraint spans. Other decoding algorithms have also been devised. However, this section is intended as a brief introduction to the subject and we shall not pursue the topic further here.

DRILL PROBLEM 9.13.1

Determine the first 18 output bits in the convolutional code for the message 110011. Use the encoder shown in Fig. 9.31 and assume that all shift registers are initially at zero.

ANSWER: 111101001000100101.

9.14 SUMMARY

Communication is the transmission of information from one point in space and time to another. Information conveyed by a knowledge of the occurrence of an event is proportional to the logarithm of the inverse of the probability of its occurrence. The

† See, for example, G. D. Forney, "The Viterbi Algorithm," *Proceedings of the IEEE*, vol. 61 (March 1973): 268–273.

binary unit of information is the bit. The information rate is the product of the average information (entropy) per symbol and the symbol rate.

The filter-like characteristics of the propagation medium and transmitting and receiving stages can be described by a mathematical model known as the channel. The limiting rate of information transmission through a noisy channel is called the channel capacity. The Hartley–Shannon law gives the channel capacity C in the presence of band-limited white gaussian noise as $C = B \log_2 (1 + S/N)$. Use of this result shows that the optimum trade-off between the signal-to-noise ratio, S/N, and bandwidth B is exponential.

Quantizing a signal introduces noise. This quantization noise is inversely proportional to the number of quantizing levels. A nonlinear characteristic known as companding is used to make all given signal levels more equiprobable in order to minimize the effects of quantization noise.

Direct transmission of a low-frequency signal is called baseband communication. Two types of binary baseband PCM are the on-off signal and the polar signal. The polar baseband PCM signal is more efficient and uses one-half as much average power for a given error probability. Differential PCM (DPCM) is used for possible bandwidth savings; delta modulation (DM) uses simpler circuitry to send and receive coded transmissions.

Overall S/N performance in PCM systems is limited by both transmission (thermal) noise and quantization noise. Usually, PCM systems are designed so that the latter is the controlling factor.

In long-distance transmission, signals often are equalized and amplified by repeaters spaced at distances determined by expected signal distortions and attenuations. Effects of transmission noise are compounded in successive repeaters in analog systems, whereas these effects increase approximately in proportion to the number of repeaters in digital systems.

A random binary waveform's power spectral density is of the form $(\sin x/x)^2$. Various coding formats have some effect on this form. For PCM signals having transitions predominantly toward a given polarity in direct relationship to the clock rate, the power spectral density will contain discrete frequency components.

Partial-response signaling uses controlled amounts of intersymbol interference to transmit data at increased rates in nonideal band-limited channels. This increased rate is at the expense of some additional transmitted power, and precoding is necessary to prevent undesired error propagation.

Filters adjusted to compensate for channel imperfections are called equalizers. Transversal filters are popular equalizer designs because they permit adjustment of parameters either prior to data transmission (preset types) or during data transmission (adaptive types).

Transmission at increased information rates can be achieved using one of more than two possible levels within each symbol duration. In M-ary signaling, one of M levels is sent and, thus, each symbol represents $\log_2 M$ bits of information. Because bandwidth is inversely proportional to symbol duration, such systems can transmit at increased data rates within a given bandwidth by increasing M. This is accomplished at the expense of an increase in transmitted power and complexity of equipment.

In error-control coding, redundancy is introduced by adding extra bits to the transmitted message, making it possible to detect and, possibly, to correct errors that may have occurred. Potential increases in system performance are at the expense of bandwidth. Algebraic codes are examples of block codes in which redundancy is introduced as parity-check digits in each codeword. In contrast, the parity-check digits in convolutional codes are continuously interleaved in the coded bit stream.

Selected References for Further Reading

1. W. R. Bennett and J. R. Davey. *Data Transmission*. New York: McGraw-Hill, 1965.
   Presents digital communication techniques with particular emphasis on telephone systems.
2. H. Taub and D. L. Schilling. *Principles of Communication Systems*. Second ed. New York: McGraw-Hill, 1986.
   Chapters 5, 12, and 13 contain good discussions of quantization, delta modulation, and block codes.
3. A. B. Carlson. *Communication Systems*. Third ed. New York: McGraw-Hill, 1986.
   Baseband transmission is covered in Chapters 11–13.
4. K. S. Shanmugam. *Digital and Analog Communication Systems*. New York: John Wiley & Sons, 1979.
   Chapter 5 has a good treatment of baseband data transmission; various coding techniques are discussed briefly in Chapter 9.
5. R. W. Lucky, J. Salz, and E. J. Weldon, Jr. *Principles of Data Communication*. New York: McGraw-Hill, 1968.
   Chapters 4–6 cover baseband data transmission and equalization with particular emphasis on telephone systems.
6. K. Feher (ed.). *Advanced Digital Communications*. Englewood Cliffs, N.J.: Prentice Hall, 1987.
   There are good explanations of partial-response signaling and adaptive equalization in Chapters 8 and 12.
7. S. Lin. *An Introduction to Error-Correcting Codes*. Englewood Cliffs, N. J.: Prentice Hall, 1970.
   Well-presented coverage of various methods of error-correcting codes.
8. R. Steele. *Delta Modulation Systems*. New York: John Wiley & Sons (Halsted Press), 1975.
   Discusses various forms of delta modulation and their S/N performance; comparisons with PCM. Advanced.
9. J. G. Proakis. *Digital Communications*. Second ed. New York: McGraw-Hill, 1989.
   The topic of digital signaling over a bandwidth-constrained channel, with an emphasis on adaptive equalization, is covered in Chapter 6.
10. R. G. Gallager. *Information Theory and Reliable Communication*. New York: John Wiley & Sons, 1968.
    An advanced treatment of coding is presented in Chapter 6.

PROBLEMS

9.1.1 a) Compute the information in the experiment when a fair die is thrown, assuming that the events (dots on each face) are equally likely to occur.

b) Repeat part (a) if the following probabilities are given for the faces of the die:

$$P(1) = 0.1, \quad P(2) = 0.2, \quad P(3) = 0.25, \quad P(4) = 0.3, \quad P(5) = 0.05.$$

9.1.2 In a university that uses the letter grades A, B, C, D, F, it is known that a certain professor in a required course gives 40% of the students a C grade and 5% F. How much information, in bits, is transmitted when a student's parents are informed (having no prior information) that the grade in the course is

a) C; b) not an F; c) F.

9.1.3 Residential units in a certain city each have meters for electricity, water, and gas consumption. Suppose that the electricity meters each have six decimal read-outs (i.e., 0–9), water meters have five, and gas meters have four.

a) Calculate the amount of information (in bits) to make each of the three meter readings.
b) It is proposed that this meter information be sent automatically, in a binary format, via a data transmission. As a minimum, each residential unit would need a unique identification number in addition to a sequence of three codes for the meter readings (in a known sequence). If the capacity of such a system is to be 65,000 residential units, determine the number of bits required to automate the meter readings.

9.1.4 The entropy of "average" English text is 4.16 bits with a letter frequency (per 1000):

E	124	S	66	C	32	W	20	K	5
T	96	R	60	U	31	Y	19	Q	2
A	81	H	51	P	23	B	16	X	2
O	79	L	40	F	23	G	16	J	1
N	72	D	37	M	22	V	9	Z	1
I	72								

a) What would be the entropy of English text if all letters were equiprobable?
b) If you are asked to guess the name of an object based on its first letter, which is the more helpful clue, "E" or "Z"? Explain.
c) Suppose you set out to design a matching game in which letters are arranged to make words. You decide to assign 1 point to the easiest letter and 10 points to the most difficult. You decide to use entropy as a measure of difficulty and then scale linearly to the points given to each (rounding off to whole numbers). Neglecting the fact that use of some letters may be dependent on others, how much value should you assign "N" and "K"?

9.1.5 a) Assuming an information transmission rate of 10 bits per second for a keypad (containing digits 0 through 9) of the keyboard of a computer, and given that each digit is equally likely to be transmitted, compute the average symbol duration time (i.e., inverse symbol rate).
b) Repeat part (a) if a 0 is three times more likely than any of the other digits.

9.1.6 A given code has six messages with probabilities $P(\mathrm{A}) = 1/2$, $P(\mathrm{B}) = 1/4$, $P(\mathrm{C}) = 1/8$, $P(\mathrm{D}) = 1/16$, $P(\mathrm{E}) = 1/32$, $P(\mathrm{F}) = 1/32$.

a) Determine the entropy of the code.
b) Assuming statistical independence, find the information in the messages X = BABC and Y = FDEF.

9.1.7 Repeat Problem 9.1.6, with $P(\mathrm{A}) = 0.25$, $P(\mathrm{B}) = 0.22$, $P(\mathrm{C}) = 0.18$, $P(\mathrm{D}) = 0.15$, $P(\mathrm{E}) = 0.12$, $P(\mathrm{F}) = 0.08$.

9.1.8 (Huffman Coding) It is advantageous to make the average number of digits in a code as nearly equal as possible to the source entropy. This can be done by assigning shorter codewords to more probable messages, but at the expense of variable-length codewords. An algorithm for generating such a code is outlined, and consists of three steps. The first step is to list the messages in a column in descending order of their probabilities of occurrence. Next, combine the last two (i.e., those two with the lowest probabilities) and list again in a second column in descending order of probabilities. Continue this procedure until there are only two left in the list.

In the second step, list the final two in the first step in a row, then place each pair beneath its reduction, row-by-row, retracing the steps above. In the third step, assign a 0 and 1 to the top pair in the list in step two (it makes no difference if 0 or 1), then assign a 0 or 1 as the second digit for entries in the second row, and continue row-by-row. This procedure systematically assigns shortest codewords to messages with highest probabilities of occurrence. Using this procedure, generate the codewords:

a) in Drill Problem 9.1.2;
b) for Problem 9.1.6;
c) for Problem 9.1.7.

9.2.1 A standard voice-grade (analog) telephone line has a usable bandwidth for data transmission of about 2400 Hz.

a) If $S/N = 30$ dB can be maintained, what is the channel capacity?
b) Calculate the S/N to maintain the following data rates on the line, allowing a 10 dB margin over the minimum theoretical S/N required: 1200, 2400, 4800, 9600 bps.

9.2.2 a) What is the required signal-to-noise ratio for a peripheral device (e.g., a printer), if one assumes a channel capacity of 3 kbps, in a bandwidth of 500 Hz?
b) Repeat part (a) for a bandwidth of 1 kHz.

9.2.3 Compute the required channel capacity for the transmission of digitized TV pictures, assuming a signal-to-noise ratio of 50 dB and a transmission bandwidth of 5 MHz.

9.2.4 Engineering management requests a design on a facsimile machine that will transmit a 8×10 inch typewritten page at a resolution of 150 dpi (dots per inch) on a standard telephone circuit (cf. Drill Problem 9.2.1) within one minute. Can it be done? Explain your answer.

9.3.1 At the input of an ideal receiver the transmitted bandwidth is 50 kHz while the bandwidth at the output is 5 kHz.

a) If the input has a S/N ratio of 10 dB, what is the S/N ratio at the demodulator output?
b) Compare your answer in part (a) with what you would expect in FM under similar conditions.

9.3.2 Suppose that the information in the ten digits of a telephone number is to be transmitted digitally.

a) How many pulses are necessary if each pulse has (1) two levels? (2) four levels? (3) eight levels?
b) Estimate the bandwidth required, relative to that required for the binary case, for the other two cases in part (a).
c) Estimate the increase in signal power (in dB) required, relative to that required for the binary case, for the other two choices in part (a). Assume a fixed variance of the additive noise and the criterion used for Eq. (9.11).

9.4.1 A signal voltage that is band-limited to 3 kHz and amplitude-limited to a peak-to-peak amplitude of 4 V is converted to a binary PCM code using 128 equally spaced levels.

a) Calculate the minimum bandwidth required.

b) Calculate the peak-signal–to–rms-quantization ratio, and the mean-square quantization S/N ratio.

c) Repeat parts (a), (b) if 256 levels were used.

9.4.2 a) A logarithmic companding law used for telephone systems is the μ-law defined by

$$e_o = \frac{\log(1 + \mu e_i)}{\log(1 + \mu)},$$

where e_i, e_o are normalized to (0, 1) and μ is a constant. (Odd symmetry is used for negative arguments.) Plot the compressor characteristic for $\mu = 255$. (PCM telephone systems in the United States use μ-law companding with $\mu = 255$.)

b) A logarithmic compression used in telephone service in Europe is the A-law defined by

$$e_o = \begin{cases} \dfrac{Ae_i}{1 + \log(A)} & 0 \le e_i < 1/A \\ \dfrac{1 + \log(Ae_i)}{1 + \log(A)} & 1/A \le e_i < 1 \end{cases}.$$

Plot the compressor characteristic for $A = 100$, and compare with the $\mu = 255$ characteristic in part (a).

9.4.3 A signal with a uniform distribution over $(-1, 1)$ is quantized using four equally spaced levels.

a) Compute the mean-square value of the quantized waveform.

b) This quantized waveform is applied to the input of the $\mu = 255$ compressor described in Problem 9.4.2(a). Compute the mean-square value of the compressor output.

c) Repeat part (b) for the $A = 100$ compressor.

9.4.4 A given signal has a probability density

$$p(x) = \begin{cases} ke^{-|x|} & -2 \le x < 2, \\ 0 & \text{elsewhere.} \end{cases}$$

a) Determine the mean-square quantization noise for four equally spaced quantization levels over $(-2, 2)$. Do not assume that the distribution is constant within each interval.

b) Compare your answer to part (a) with that assuming a uniform distribution.

9.4.5 Repeat Example 9.4.1 for a four-level quantizer. Also, determine the mean-square value of the quantized output.

9.4.6 The method of dividing the range into a number of discernible levels prior to Eq. (9.11) is sometimes used in the design of data transmissions requirements for instrumentation. For example, the noise generated in a photomultiplier tube (PMT) has zero mean value and a variance proportional to the light (signal) level. If the lowest signal level is 0.1 V and the highest level is 2 V, determine the appropriate levels for eight-level quantization. [*Hint:* Assume a triangular decreasing pdf over (0.1, 2), integrate to find $F_X(x)$, then divide $F_X(x)$ into eight equal portions to solve for x.]

9.4.7 Quantization levels can be chosen to minimize the error on a mean-square basis. If b_k is the kth quantizing level and a_k, a_{k+1} are the adjacent round-off boundaries, then the error for n quantizing levels is

$$\overline{\epsilon^2} = \sum_{k=1}^{n} \int_{a_k}^{a_{k+1}} (x - b_k)^2 p(x)\, dx .$$

We assume here that the a_k's are chosen to be the arithmetic means between the respective neighboring b_k points. (This can be shown to be an optimum choice.)

a) Derive an expression for b_k that will minimize the mean-square error over the interval between the adjacent a_k's.

b) Because the end points (a_1, a_{k+1}) are known, the quantizing levels (b_k) can be found. Solve Example 9.4.1 for a four-level quantizer using these results.

9.5.1 a) For the polar binary case, let the two signal levels be $-A, +A$. Assume additive gaussian noise with zero mean and variance σ^2. Show that for equiprobable ones and zeros the net probability of error is given by [cf. Eq. (9.39)]:

$$P_e = Q(A/\sigma) .$$

b) Generalize the result in part (a) for the case in which the difference between the signal levels is Δ.

9.5.2 A received polar binary signal is either $+1$ mV or -1 mV over $(0, T)$. Additive white gaussian noise with zero mean is present with a (two-sided) power spectral density of 1 fW/Hz across 50 ohms.

a) Determine the noise-equivalent bandwidth B required for a probability of error of 1×10^{-4}. (See Problem 9.5.1.)

b) If the composite signal-plus-noise is processed by a matched filter, determine the minimum time T for a probability of error of 1×10^{-4}.

9.5.3 An on-off binary system uses the pulse waveform

$$f(t) = \begin{cases} A \sin(\pi t/T) & 0 \le t < T \\ 0 & \text{elsewhere} \end{cases}$$

for a one and absence of a pulse for a zero. Let $A = 0.2$ mV and $T = 2$ μsec. Additive white gaussian noise is present with a power spectral density of 1 fW/Hz across 1 ohm.

a) Determine the net probability of error for equiprobable ones and zeros.

b) Repeat part (a) for a binary system which sends $[-f(t)]$ for zeros and transmits the same average power.

9.5.4 A binary transmission uses the Manchester code representation of ones and zeros, as shown in Fig. P–9.5.4. Assume the use of matched-filter detection in the presence of additive white gaussian noise with (two-sided) power spectral density of 1 fW/Hz across 1 ohm. The waveform parameters at the receiver are $A = 0.1$ mV and $T = 2$ μsec.

a) Sketch a block diagram of the optimum receiver.

b) Sketch a diagram of the autocorrelation of the input waveform.

c) Sketch the pdf's for ones and zeros, assuming that $B = 1/(2T)$.

d) Find the net probability of error, assuming equiprobable ones and zeros.

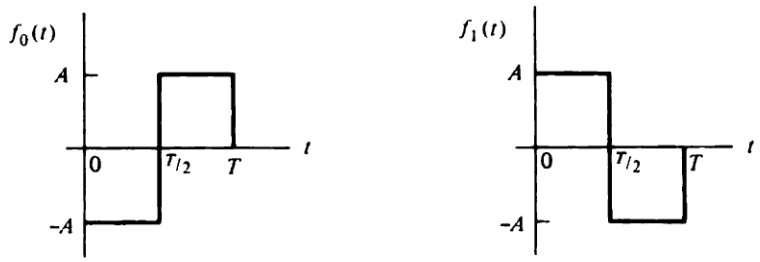


Figure P–9.5.4.

9.5.5 A binary system operating at 1200 bps transmits a pulse waveform $f(t)$ for a one and $-f(t)$ for a zero, where $f(t)$ is a rectangular pulse with duration of 0.2 msec. The additive noise is white gaussian noise with zero mean. The system is to be designed so that the received pulse waveform has full-cosine Nyquist pulse shaping with a -6 dB frequency of 1200 Hz. Determine the magnitude frequency transfer functions (normalized to unit gain at dc) of the transmit and receive filters that give zero ISI and optimum detection if the channel characteristics are described as follows.

a) $H_C(\omega) \approx 1$ over the frequencies of interest.
b) The impulse response of the channel is of the form:

$$h_C(t) = (1/\tau_C) \exp(-t/\tau_C)u(t), \quad \text{where } \tau_C = 0.1 \text{ msec.}$$

9.5.6 a) Repeat Problem 9.5.5, but with additive gaussian noise that has a (two-sided) power spectral density

$$S_n(\omega) = \tfrac{1}{2}N_0[1 + (\omega/\omega_1)^2]^{-1};$$

b) Examine the particular case in which $\omega_1 = 1/\tau_C$.

9.6.1 Calculate the mean-square error in the reconstruction of a three-bit binary codeword resulting from a random bit error in transmission in each of the following conditions. Assume that the probability of error, P_e, is small enough that the probability of more than one-bit error in a codeword is negligibly small.

a) Use a linear quantization rule, and normalize so that reconstruction is over the amplitude range $(-1, 1)$.
b) Repeat part (a) using the companding rule developed in Example 9.4.1.

9.6.2 Let a PCM threshold be defined at the point for which S_o/N_o decreases by 1 dB from the quantization-noise-limited performance of Fig. 9.12. This number can be used as a guideline for minimum S_i/N_i to maintain quantization-noise-limited performance.

a) Determine the S_i/N_i threshold values for $m = 6, 8, 10, 12$.
b) Show that a reasonable approximation over this range is given by $S_i/N_i \geq (m/2 + 10)_{dB}$ for quantization-noise-limited performance.

9.6.3 At the optimum sample times in the receiver, the presence of ISI can be treated as an additive random variable. Consider the case of a polar-NRZ waveform with ISI having the sampled values

$$y(t_k) = \pm A + n(t_k) + z(t_k)$$

where $n(t_k)$ is additive noise and $z(t_k)$ represents ISI. The pdf of $z(t_k)$ is given here as

$$p_Z(z) = \tfrac{1}{4}\delta(z + \alpha) + \tfrac{1}{2}\delta(z) + \tfrac{1}{4}\delta(z - \alpha).$$

a) Assuming zero-mean gaussian noise $n(t)$ with variance σ^2, obtain an expression for the net probability of error in terms of A, α, and σ for equiprobable ones and zeros. (It may be helpful to make use of the result of Problem 8.9.3.)
b) For $A/\sigma = 3$, compute the equivalent degradation in S/N (in dB) for $\alpha/A = 0.1$ and $\alpha/A = 0.2$.
c) Interpret the ISI qualitatively in terms of an eye diagram.

9.6.4 It is given that the normalized response of a certain channel to one pulse of a polar binary waveform can be approximated by (except for a time delay):

$$g(t) \approx \begin{cases} \sin(\pi t/T) & 0 \le t < T \\ 0.2 \sin(\pi t/T) & T \le t < 2T \end{cases}.$$

a) Sketch an eye pattern for an alternating 1010 sequence.
b) Sketch an eye pattern for a random sequence. [See Problem 9.6.3(a).]
c) Identify both the amplitude and the timing effects of the ISI in your eye patterns.

9.7.1 A given DM system operates with a sampling frequency $f_s = 1/T$ and a fixed step size a. If the input to the system is $f(t) = kt$ for $t > 0$, determine (a) the value of k above which a slope overload will occur; (b) the value of k that minimizes the mean-square quantization noise.

9.7.2 A DM system that is tailored to handle signals of the form $f(t) = t \exp(-kt)u(t)$ uses a sampling frequency of $f_s = (10k)$. Determine the maximum step size a above which a slope overload will occur.

9.7.3 A DM system is designed for a low-pass signal with bandwidth of 3 kHz (assume flat spectral response). The maximum peak-to-peak input voltage is 2 V. If the S/N improvement in demodulation is 30 dB, calculate
a) the required sampling frequency, f_s;
b) the step size a;
c) the ratio $af_s/(bf_m)$ [cf. Fig. 9.16].

√ **9.7.4** Write a computer program for the DM system shown in Fig. 9.15(a). Let the input be $f(t) = \sin 2\pi t$ and use sample spacings of 0.05 sec. Run your program and plot both $f(t)$ and the accumulator output for a step size of a) 0.10; b) 0.20; c) 0.40.

√ **9.7.5** Convert the DM system of Problem 9.7.4 into a two-bit DPCM system by adding the following statement at the appropriate point: "If the magnitude of the difference is greater than one step size, double the step size." Repeat Problem 9.7.4 for this system with sample spacings of 0.10 sec.

★ **9.8.1** Modify the results of Section 9.8, assuming the gains of all repeaters are not identical.

★ **9.8.2** a) Compute the net probability of error in a transmission system using two analog repeaters, assuming additive gaussian noise with zero mean and a probability of error of 1×10^{-8} for each repeater.
b) Suppose one repeater has 1 dB more noise than it should [under the conditions of part (a)]. What is the resulting probability of error under this condition?

★ **9.8.3** NRZ polar binary signals are transmitted over a transmission path using four regenerative repeaters. Assume additive gaussian noise with zero mean value and a probability of error of 1×10^{-8} for each repeater. Compute
a) the net probability of error;
b) the net probability of error if analog repeaters were used;
c) the additional transmitted power, in dB, required in each analog repeater to give the same net probability of error as a system in which regenerative repeaters are used.

★ **9.9.1** Determine the power spectral density of the following transmissions for equiprobable ones and zeros: a) the binary system in Problem 9.5.3(a); b) the binary system in Problem 9.5.3(b).

★ **9.9.2** Suppose that a certain binary system transmits the waveform in Problem 9.5.3(a) for a one and a fixed amplitude $-A/2$ over $(0, T)$ for a zero. Assume equiprobable 1's and 0's.

a) Do you expect any impulses in the power spectral density of the transmitted data waveform? Explain.
b) Determine the power spectral density.

★ **9.9.3** The waveform $f_1(t)$ is sent over $(0, T)$ to signal a one and $f_2(t)$ is sent over $(0, T)$ to signal a zero. Determine the power spectral density, assuming equiprobable ones and zeros, if $f_1(t) = A$ and $f_2(t) = \gamma A$, where $0 \le \gamma < 1$.

★ **9.9.4** Repeat Problem 9.9.3 if

a) $f_1(t) = A \sin^2(\pi t/T), f_2(t) = 0$.
b) $f_1(t) = A \sin^2(\pi t/T), f_2(t) = -f_1(t)$.

9.10.1 a) Beginning with the block diagram in Fig. 9.24(a), derive the impulse response expression given in Eq. (9.92), and verify the plot given in Fig. 9.24(c).
b) Removing the relative time delay, compare Fig. 7.9(c) with Fig. 9.24(b), (c) for $B_x = 1/(2T)$.

9.10.2 Derive the net probability of error for a duobinary system and compare it with a polar binary system (cf. Problem 9.5.1) within the following constraints. Both systems are optimized to operate in the presence of additive white gaussian noise and with a flat channel response; i.e., $|H_C(\omega)| = 1$. The amplitudes in the duobinary system are $-2A$, 0, $+2A$, the decision thresholds are set at $\pm A$, and the pulse shaping is $(1/T)h(t)$ [cf. Eq. (9.92)].

9.10.3 Consider the partial-response signaling described by

$$y_k = x_k + 2x_{k-1} + x_{k-2}.$$

a) Sketch a block diagram of the transversal filter needed to implement y_k.
b) Determine the frequency transfer function of the filter.
c) Determine the impulse response of the filter.

9.10.4 Use the sequence of binary digits in Problem 7.6.4.

a) Determine the corresponding sequence of duobinary and modified-duobinary digits y_k. Let the levels be represented as $0, 1, 2$ and $-1, 0, 1$, respectively.
b) Decode the y_k found in part (a), assuming no knowledge of the input sequence (x_k) except that the two bits prior to this sequence can be assumed to be 1's.
c) Repeat part (b), assuming that an error occurs so that the first $y_k = 0$ occurring beyond the first three digits is interpreted by the receiver to be 1. Trace the effect of this error on the decoding of the signal.
d) Repeat part (a) and part (c) using precoding. Assume that the precoder is initialized with 1's.

9.10.5 Investigate the required receiver rule for use of the precoding operation $b_k = [(x_k + b_{k-1})\text{mod-}3]$ in duobinary signaling. For a binary input data stream x_k, how many levels are required for y_k?

9.11.1 In the presence of multipath interference, replicas of the transmitted signal arrive at the receiver with various attenuations and delays. In this problem, assume that there is only one predominant source of multipath. If the transmitted signal in such a system is $x(t)$, then the (noise-free) received signal $y(t)$ is

$$y(t) = x(t) + \alpha x(t - \tau_m),$$

where α is the attenuation of the multipath component and τ_m is its delay. Design a transversal equalizer (cf. Fig. 9.26) for this particular problem, assuming that $T = \tau_m$ and $\alpha \ll 1$. [*Hint:* Work with frequency transfer functions.]

9.11.2 Repeat Example 9.11.1, with the exception that $x_{-1} = 0.4$.

9.11.3 Show that under the conditions of Drill Problem 9.11.1 and $|x_{-1}| = |x_1|$, the three-tap transversal equalizer provides (a) amplitude equalization if $x_{-1} = x_1$, and (b) phase equalization if $x_{-1} = -x_1$.

9.11.4 A digital transmission system is considered for sending data through a channel that is modeled as a low-pass filter with frequency transfer function:

$$H_C(\omega) = A(1 + a\omega^2)e^{-j\omega t_0} \operatorname{rect}[\omega/(2W)].$$

The transmitted waveform $f(t)$ is approximated by a trapezoidal waveform shown in Fig. P–9.11.4. Design a three-tap transversal equalizer to force one zero on either side of the main pulse response. It is given that $t_r = T/8$, and that $|H_C(W)|/|H_C(0)|$ is -3 dB.

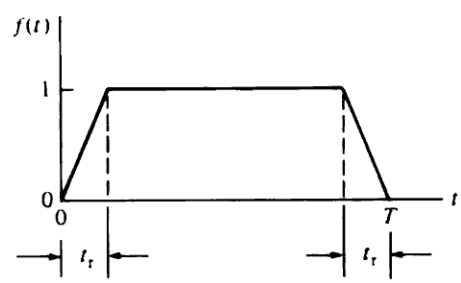


Figure P–9.11.4.

√ **9.11.5** Determine the coefficients of a five-tap transversal equalizer filter that will force two zeros on either side of the main pulse response. The sample values of the distorted pulse waveform are: $x_{-4} = -0.02$, $x_{-3} = 0.05$, $x_{-2} = -0.1$, $x_{-1} = 0.4$, $x_0 = 1.0$, $x_1 = -0.2$, $x_2 = 0.1$, $x_3 = -0.1$, $x_4 = 0.05$, $x_k = 0$ for $|k| > 4$. For convenience in the numerical solution, use Eq. (9.109) in the matrix form $[x][a] = [y]$, where $[x]$ is the $(2N + 1) \times (2N + 1)$ channel response matrix, $[a]$ is a column matrix of the tap gains, and $[y]$ is a column matrix satisfying Eq. (9.108). Find $[x]^{-1}$ using numerical methods, and note that the tap gains are given by the center column in $[x]^{-1}$. (Why?)

9.12.1 Repeat Example 9.12.1 for $P_\epsilon = 1 \times 10^{-6}$, and compare your results with those in Table 9.5.

9.12.2 Repeat Drill Problem 9.12.1 for $P_\epsilon = 1 \times 10^{-6}$, and compare your results with those in Table 9.7.

9.12.3 Design a M-ary signaling system to transmit 9600 bps over a 3.2 kHz channel, and calculate the probability of bit error assuming Gray level coding. Use Nyquist pulse shaping and assume additive white gaussian noise with two-sided power spectral density of 0.5×10^{-10} W/Hz. Expected losses are 30 dB, and a maximum peak output power level of 0 dB may be used in the transmitter.

9.12.4 For Eq. (9.123) it was assumed that all levels are equiprobable. Investigate the effect on P_e if the levels at $\pm 3A/2$ were more probable than the two at $\pm A/2$, and compare the results with Eq. (9.125). [*Hint:* It may be helpful to recall Leibniz's rule in integral calculus.]

★ **9.13.1** The parity-check matrix for a repetition code with code symbols 000 and 111 is

$$[H] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

a) Show that Eq. (9.135) is satisfied.
b) Demonstrate by way of an example that this repeated code will correct single errors. Does this conclusion agree with Eq. (7.42)?

★ **9.13.2** Explain why a column of all zeros is not allowed in the parity-check matrix $[H]$.

★ **9.13.3** Find all possible codewords in the parity-check code that has the parity-check matrix of Example 9.13.1.

★ **9.13.4** a) Using shift registers and modulo-2 adders, sketch a diagram for a decoder for error detection of the (7, 4) Hamming code in Example 9.13.1.
b) This section emphasized the *decoding* aspects of block codes. To generate a (n, k) block code we define a $k \times n$ generator matrix $[G] = [I_k \,|\, P]$, where $[I_k]$ is the $k \times k$ identity matrix and $[P]$ is a $k \times r$ submatrix of binary digits that form distinct binary code words from those already in $[I_k]$. The identity matrix $[I_k]$ in $[G]$ is used to reproduce a message codeword for the first k bits and the $[P]$ submatrix generates the check digits. The parity-check matrix used in decoding is given by $[H] = [P^T \,|\, I_r]$, where I_r is the $r \times r$ identity matrix and $[P]^T$ is the transpose of $[P]$. Find the generator matrix $[G]$ for the (7, 4) Hamming code in Example 9.13.1.
c) Sketch a block diagram for the encoder in part (b).

★ **9.13.5** A Hamming code ($r \geq 3$) maintains a distance 3 between codewords, and thus it can be used for single-error correction or for double-error detection. Suppose that a parity-check matrix for a (15, 11) Hamming code is arranged so that the binary column words decrease in numerical value from left to right, except that all words containing a single one are placed in the 4×4 identity matrix. Demonstrate that the general statements above are true for this case.

★ **9.13.6** A given convolutional encoder is constructed using shift registers and modulo-2 adders. The equations for the adder outputs are: $v_1 = S_1 \oplus S_3$, $v_2 = S_1 \oplus S_2 \oplus S_3$.
a) Sketch a block diagram of the encoder.
b) Determine the output bits for the input conditions of Drill Problem 9.13.1, but as applied to this encoder.

★ **9.13.7** A given convolutional encoder is constructed using shift registers and modulo–2 adders. The equations for the adder outputs are

$$v_1 = S_1 \oplus S_2 \oplus S_3, \qquad v_2 = S_1, \qquad v_3 = S_1 \oplus S_2.$$

Repeat Drill Problem 9.13.1 for this encoder.

★ **9.13.8** An alternative rate-$\frac{1}{3}$ convolutional encoder to that in Problem 9.13.7 is suggested in which the equations for adder outputs are: $v_1 = S_1 \oplus S_2$, $v_2 = S_2 \oplus S_3$, $v_3 = S_2 \oplus S_3$.

a) Draw a block diagram of each of the two encoders.

b) To make a comparison between encoders, prepare a code tree diagram, such as shown in Fig. P–9.13.8, for each encoder. The (0, 1) binary numbers on the tree show the inputs; place the three-bit code generated in each of the parentheses as the encoder progresses. (Assume that the shift registers start with all zeros.)

c) Based on the results of part (b), which encoder would you prefer to use in a communications system, and why?

d) If the output sequence 101010 were received in the presence of additive noise, what is the most probable input sequence that was sent by the encoder in Problem 9.13.7? Explain your answer.

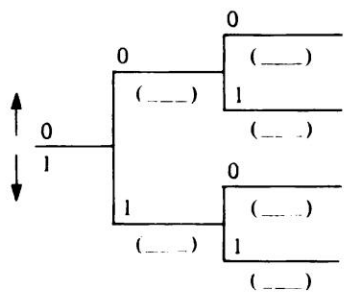


Figure P–9.13.8.

DESIGN PROBLEMS

Design Problem

9.1 Geostationary Satellite Visual Channel Design

This problem is for a baseband design of a satellite subsystem that senses and transmits visual information on cloud patterns from a geostationary satellite orbit. Its parameters are typical of those for the Geostationary Orbiting Environmental Satellite (GOES) series in use in the 1980s.

This cylindrical satellite is spin-stabilized (for example, see Reference 2 of Chapter 4) and rotates at 100 revolutions per minute; the spin axis is aligned with the earth's rotational axis. The sensor chosen for the visual channels is the photomultiplier tube (PMT), which yields an output voltage proportional to input light intensity. A lens mounted perpendicular to the satellite spin axis and a mirror behind it, tilted by a stepping motor, focus light on a linear array of eight PMT's. These PMT's sense eight parallel tracks of visual data. When the lens is pointed away from the earth, the stepping motor alters the tilt of the mirror for the next scan. In this manner the satellite scans the hemisphere in a systematic progression from pole to pole, at which time the mirror is reset and the system is ready to repeat another scan pattern.

Outputs of the PMT's are quantized to 64 levels, companded using a square-root law [cf. Problem 9.4.6], and time multiplexed. The desired spatial resolution of each PMT is 0.5 mile by 0.5 mile at the subsatellite point (i.e., that point on the earth's surface on a line joining the satellite position and the earth's center.)

Design this satellite subsystem. Probability of error goal for the system is 1×10^{-6}. We are not concerned here with the transmitting and receiving systems; base your design on use of polar NRZ binary data waveforms at baseband. Several items of interest to the design project are listed below. For parts a–c, assume that there are no provisions for mass data storage on the satellite. Allow 25% above Nyquist rate for sampling. State any assumptions you make.

a) Prepare timing diagrams and a companding characteristic plot. Determine the minimum bandwidth you will need. How long will it take to make one complete scan of the hemisphere?

b) For cases in which the link S/N budget may become marginal, it is advantageous to build in a provision in which the satellite is commanded to scan at a reduced resolution. Investigate possibilities for building into your design several "back-up" modes by combining channels to give 1-, 2-, 4-mile scan resolutions. Prepare timing diagrams and estimate relative S/N trade-offs.

c) Investigate requirements and possible benefits in use of a single-error-correcting code in this system.

d) Only a portion of a revolution of the satellite is used to collect data. Investigate requirements for data storage and read-out if the satellite were to transmit data to ground stations all of the time, but at a reduced bandwidth.

CHAPTER

10 Digital Modulation

In the preceding chapter, we covered various aspects of the digital transmission of information at baseband frequencies. For digital communication systems employing bandpass channels, it becomes advantageous to modulate a carrier signal with the digital data stream prior to transmission. Three basic forms of digital modulation, corresponding to AM, FM, and PM, are known as *amplitude-shift keying* (ASK), *frequency-shift keying* (FSK), and *phase-shift keying* (PSK). In this chapter, we consider each of these modulation methods briefly from the perspective of their probability-of-error advantage and bandwidth efficiency. Emphasis in the early part of the chapter is on binary systems. Variants of the three basic modulation methods, some using more than two possible signaling states, are described later in the chapter. The concept of a spread spectrum, in which the bandwidth is purposely spread before transmission and then restored to its original in reception, is introduced. The chapter closes with two sections that discuss more generalized signaling representations and considerations.

10.1 AMPLITUDE-SHIFT KEYING (ASK)

In amplitude-shift keying, the amplitude of a high-frequency carrier signal is switched between two or more values in response to the PCM code. For the binary case, the usual choice is *on-off keying* (sometimes abbreviated as OOK). The resultant amplitude-modulated waveform consists of RF pulses, called *marks*, representing binary 1, and *spaces* representing binary 0. An ASK waveform is shown in Fig. 10.1 for a given PCM code (the same code used in Fig. 7.35). As in AM, the baseband bandwidth is doubled in ASK.

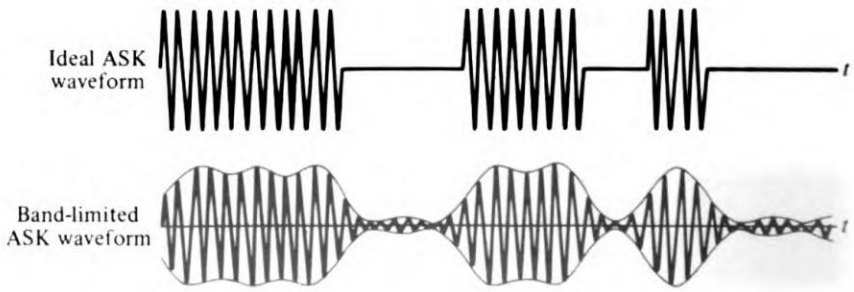


Figure 10.1 Binary ASK waveforms.

The ASK waveform for one pulse (i.e., a binary 1) can be written as

$$\phi(t) = \begin{cases} A \sin \omega_c t & 0 < t \le T, \\ 0 & \text{otherwise.} \end{cases} \tag{10.1}$$

The impulse response of the matched filter for optimum detection of this ASK waveform in the presence of white noise is, within an arbitrary constant,

$$h(t) = \phi(T - t).$$

The matched-filter output for the (noiseless) input $\phi(t)$ is

$$\begin{aligned} y(t) &= \phi(t) \circledast h(t) \\ &= \int_{-\infty}^{\infty} \phi(\tau)\phi(T - t + \tau)\, d\tau \\ &= r_\phi(T - t), \end{aligned} \tag{10.2}$$

where $r_\phi(t)$ is the time-autocorrelation function for the finite-energy signal $\phi(t)$ [cf. Eq. (4.43)]. The optimum decision time is for $t = T$; thus,

$$y(T) = r_\phi(0) = E. \tag{10.3}$$

A sketch of the matched-filter output is shown in Fig. 10.2. Using Eq. (10.1), we find the signal energy is

$$E = \int_0^T A^2 \sin^2 \omega_c t\, dt = A^2 T/2. \tag{10.4}$$

The receiver must make a decision at $t = T$ based on the two possibilities $y(T) = n_0(T)$ and $y(T) = E + n_0(T)$. For equal source probabilities of ones and zeros and noise which has a symmetrical probability density function, the optimum-decision threshold is set at $E/2$. Thus the net probability-of-error computation reduces to that of the on-off baseband system. For gaussian-distributed noise, we have found that this gives [(cf. Eq. (9.49)]

$$P_e = Q\left(\sqrt{\frac{E}{2\eta}}\right). \tag{10.5}$$

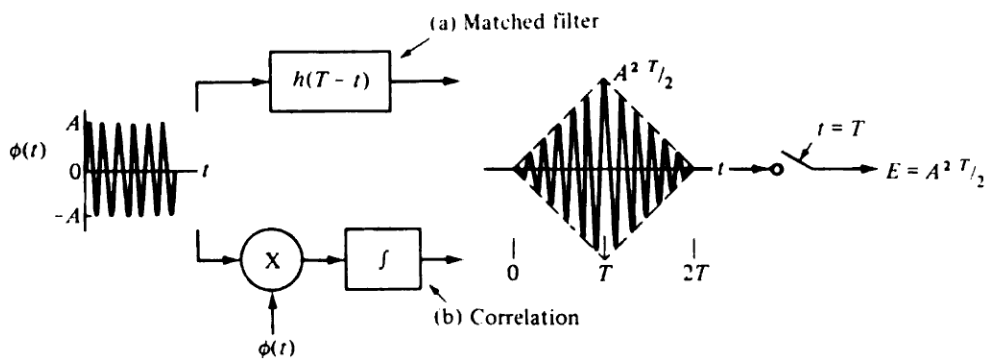


Figure 10.2 Matched-filter detection of ASK waveforms.

For purposes of comparison with other systems, we express the probability of error in terms of the average signal energy per bit, $E_{avg} = ST$ so that Eq. (10.5) can be rewritten as

$$P_\epsilon = Q\left(\sqrt{\frac{E_{avg}}{\eta}}\right). \tag{10.6}$$

The average signal power is $S = (1/2)(A^2/2)$; as before, $N = \eta B$, and if we assume Nyquist sampling, $B = 1/(2T)$[†] so that we can rewrite Eq. (10.5) in terms of the average signal-to-noise ratio:

$$P_\epsilon = Q\left(\sqrt{\frac{S}{2N}}\right). \tag{10.7}$$

From this result, we conclude that the ASK system is equivalent to an on-off baseband system in terms of the average signal-to-noise ratio required for a given probability of error. Note, however, that the *peak* signal-to-noise ratio required is increased by a factor of 2 (i.e., by 3 dB).

The matched-filter detection of ASK is essentially a synchronous detection, as shown in Fig. 10.2(b). Envelope detection can be used and is much simpler to implement, but the mathematics needed to derive the net probability of error is more involved and is not covered here. For equiprobable 1's and 0's, the result is[‡]

$$P_\epsilon = \frac{1}{2}\exp\frac{-E}{4\eta} + \frac{1}{2}Q\left(\sqrt{\frac{E}{2\eta}}\right). \tag{10.8}$$

It turns out that for $P_\epsilon < 10^{-4}$, the required signal-to-noise ratios are high enough that there is only about a 1-dB penalty (or less) for the use of envelope detection.

The power spectral density of ASK is centered at ω_c and has an identical shape to that of the corresponding on-off–keyed baseband signal. Because the bandwidth has been doubled in the modulation process, the theoretical maximum bandwidth efficiency is 1 bps/Hz. Operating systems typically use two to three times this amount of bandwidth.

10.2 FREQUENCY-SHIFT KEYING (FSK)

In frequency-shift keying, the instantaneous frequency of the carrier signal is switched between two (or more) values in response to the PCM code. Figure 10.3(a) shows an idealized FSK signal corresponding to the binary PCM code of Fig. 7.35.

[†] The bandwidth B used here is referred to baseband; the bandpass bandwidth is doubled, but after synchronous detection it can be reduced again, yielding a detection gain of 2 [cf. Eq. (5.73)].

[‡] See, for example, R. E. Ziemer and W. H. Tranter, *Principles of Communication Systems, Modulation, and Noise*. Second ed. Boston: Houghton Mifflin, 1985, pp. 383–386.

This suggests that we can consider the FSK waveform as composed of two ASK waveforms of differing carrier frequencies, as shown in Fig. 10.3(b). Thus to convey either of the binary symbols, we have a choice of the two waveforms:

$$\phi_1(t) = \begin{cases} A \sin m\omega_0 t & 0 < t \le T, \\ 0 & \text{elsewhere}, \end{cases} \tag{10.9a}$$

$$\phi_2(t) = \begin{cases} A \sin n\omega_0 t & 0 < t \le T, \\ 0 & \text{elsewhere}. \end{cases} \tag{10.9b}$$

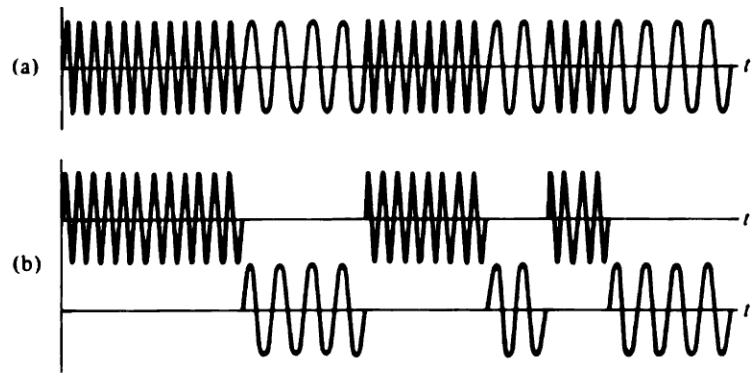


Figure 10.3 (a) An idealized FSK waveform and (b) its decomposition into two ASK waveforms.

The two received signal waveforms are now different so we use two matched filters, one for each waveform. Two possible matched-filter receivers for FSK are shown in Fig. 10.4.

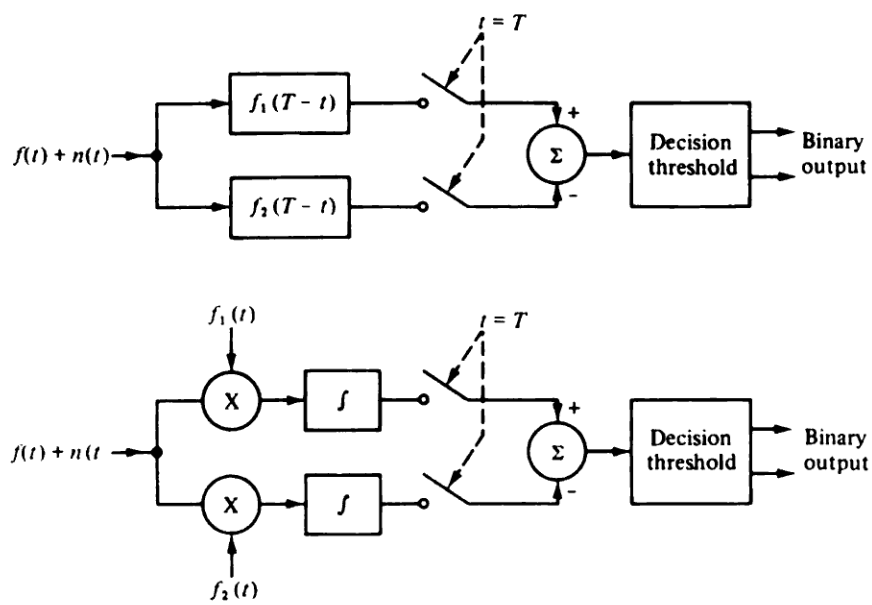


Figure 10.4 Matched-filter detection of FSK waveforms.

The average energy per binary digit is

$$E = \int_0^T A^2 \sin^2 m\omega_0 t\, dt = A^2T/2. \tag{10.10}$$

If one signaling frequency is present in the absence of noise, we assume that the one matched-filter output is zero and the other output is at E.† Conversely, if the second signaling frequency is present, the first matched-filter output is zero and as a result of the subtraction the net output is at $-E$. This is illustrated in Fig. 10.5.

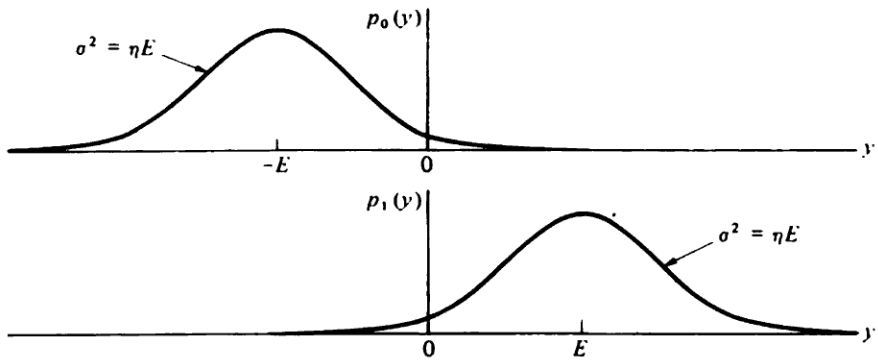


Figure 10.5 Probability density functions for binary FSK waveforms.

This would appear to be analogous to the polar-baseband case were it not for the fact that the noise voltages are subtracted at the outputs of the two matched filters. If we assume that the two matched-filter frequency responses do not overlap, the output-noise voltages are statistically independent and hence add on a power (mean-square) basis. If the two filter bandwidths are the same (the usual case), we can then just double the variance from $\sigma_n^2 = \eta E/2$ to $\sigma_n^2 = \eta E$ (cf. Fig. 10.5). The remainder of the analysis proceeds in the same manner as before. For gaussian-distributed noise and equiprobable ones and zeros, we get

$$\begin{aligned} P_e &= \int_0^\infty \frac{1}{\sqrt{2\pi\eta E}} e^{-(y+E)^2/(2\eta E)}\, dy \\ &= Q\left(\sqrt{\frac{E}{\eta}}\right), \end{aligned} \tag{10.11}$$

where E is the (average) energy per binary digit. Therefore we conclude that, on an average bit energy-to-noise basis, the net probability of error for FSK is the same as that for ASK. On the other hand, for the same peak-power requirements FSK has a 3-dB advantage over ASK because the latter is off approximately one-half of the time.

As noted before, matched-filter detection is really a synchronous detection. This requires both frequency and phase synchronization for two oscillators, one at each signaling frequency. One way to implement this is to use phase-locked loops, as

† This implies orthogonality over $(0, T)$; Problem 10.2.4.

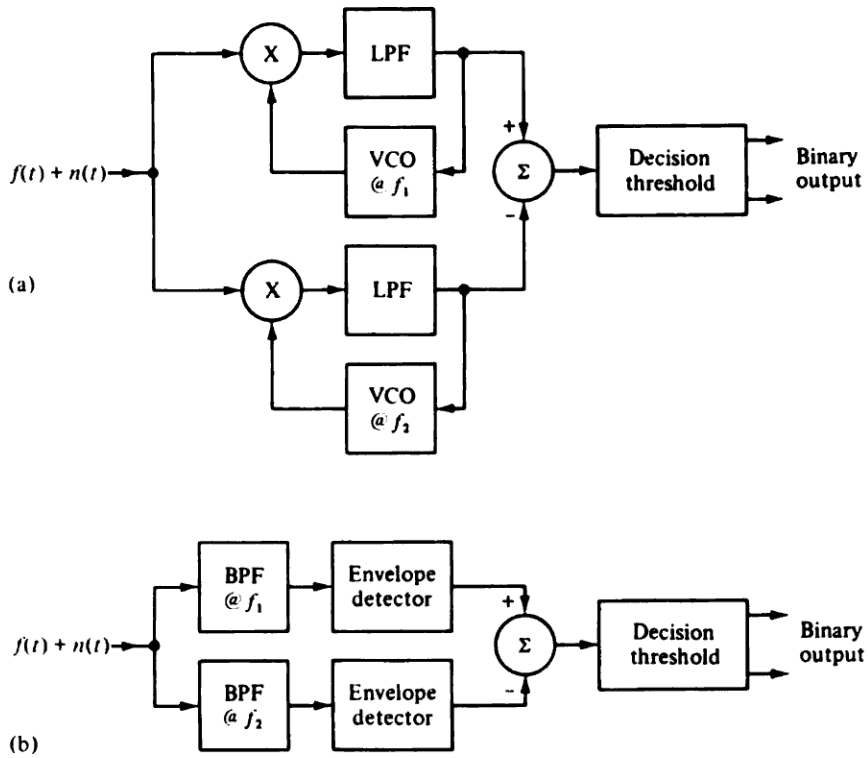


Figure 10.6 Examples of (a) coherent and (b) noncoherent FSK detection systems.

shown in Fig. 10.6(a). The frequency range of each loop is restricted and the low-pass filter is narrow enough so that the voltage-controlled oscillators (VCO) will not change frequency appreciably during a pause. A split-phase or Manchester code is well suited to this type of operation. Often one phase-locked loop is used and adjusted so that it will follow the input frequency.

A popular way to implement the FSK receiver is to use only the magnitude response of the matched filters, yielding the noncoherent FSK receiver shown in Fig. 10.6(b). The analysis of the envelope detector is omitted here, except for the result[†]

$$P_e = \frac{1}{2} \exp \frac{-E}{2\eta}. \tag{10.12}$$

Noncoherent FSK results in a S/N penalty of no more than 1 dB over that for orthogonal FSK with coherent detection for error rates of interest. It is a popular choice in operating systems. With this approach, the frequency spacing—to prevent significant overlap of the passbands of the two filters—must be at least $2\Delta f T \geq 1$, where $2\Delta f$ is the frequency difference between the two frequencies used and T is the symbol duration. Another possible approach is the use of a discriminator to convert the

[†] A. B. Carlson, *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*, Third ed. New York: McGraw-Hill, 1986, p. 538.

frequency variations to amplitude variations, followed by envelope detection. This second approach removes the constraint inherent in the first on $2\Delta fT$, but yields slightly poorer performance. Still another approach is the use of a zero-crossing detector, which also yields a slightly poorer performance.

The bandwidth of FSK transmission depends on the frequency separation used. For convenience, we refer the signaling frequencies ω_1, ω_2, to a center frequency ω_c so that the modulated signal can be written as

$$\phi(t) = A \sin \left\{ \int_0^t [\omega_c + (\Delta\omega)p(t)]\,dt \right\}, \tag{10.13}$$

where $2\omega_c = \omega_2 + \omega_1$, $2\Delta\omega = \omega_2 - \omega_1$, and $p(t)$ is a binary switching function with possible states of ± 1 over $(0, T)$. The instantaneous frequency of $\phi(t)$ is simply the bracketed term in Eq. (10.13). If T is long compared with the reciprocal frequency difference $(\omega_2 - \omega_1)^{-1}$, then the spectral density of $\phi(t)$ tends toward spectral groupings near $\omega = \omega_1$ and near $\omega = \omega_2$. For example, in the case of alternating ones and zeros, $p(t)$ is a symmetrical square wave. The spectrum for this case can be found using the Fourier series methods described in Problem 6.3.9, and results are shown in Fig. 10.7 for $2\Delta\omega T \gg 1$. As $2\Delta\omega T$ approaches zero, of course, the spectral density of $\phi(t)$ approaches that of a sinusoid at a frequency $\omega = \omega_c$.

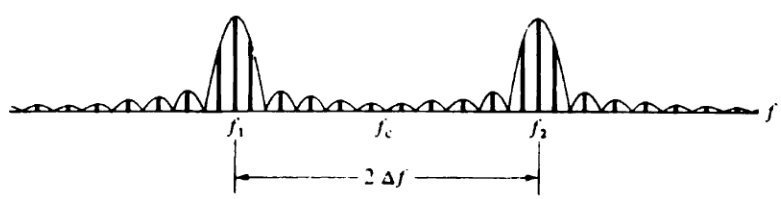


Figure 10.7 Magnitude spectrum of a periodic FSK waveform for $2\Delta\omega T \gg 1$ (positive frequencies).

What is the optimum frequency separation for the detection of FSK? Assuming that $\omega_c \gg \Delta\omega$ and ignoring possible transient effects, the low-pass difference in synchronous detector outputs (cf. Fig. 10.4) is proportional to

$$\int_0^T \{\sin[(\omega_c + \Delta\omega)t] - \sin[(\omega_c - \Delta\omega)t]\}\{\sin[(\omega_c + \Delta\omega)t]\}\,dt$$
$$= \frac{T}{2}\left[1 - \frac{\sin(2\Delta\omega T)}{(2\Delta\omega T)} + \frac{\sin(2\omega_c T)}{(2\omega_c T)}\right]. \tag{10.14}$$

For $\omega_c T \gg 2\pi$, Eq. (10.14) can be rewritten and then incorporated into Eq. (10.11) as

$$P_e = Q(\sqrt{[1 - \mathrm{Sa}\,(2\Delta\omega T)]E/\eta}\,). \tag{10.15}$$

The factor within the brackets in Eq. (10.15) is graphed in Fig. 10.8. From these results, we observe that $2\Delta fT = n/2$, n an integer, if $\phi_1(t), \phi_2(t)$ are orthogonal over $(0, T)$. Thus the minimum frequency separation is one-half cycle per bit interval be-

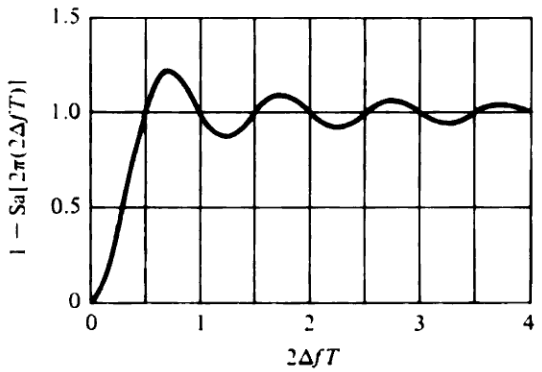


Figure 10.8 Signal detection gain in binary FSK.

tween the two signaling frequencies ω_1, ω_2 in order for $\phi_1(t), \phi_2(t)$ to be orthogonal over $(0, T)$. Some additional detection gain can be realized (see Fig. 10.8) if $2\Delta fT \approx 3/4$; i.e., the frequency separation $2\Delta fT$ for minimum probability of error requires about three-fourths of a cycle difference between ω_1 and ω_2 over $(0, T)$. Using this value, we find that the factor within the brackets yields $[1 + 2/(3\pi)] = 1.21$, or

$$P_\epsilon = Q\left(\sqrt{1.21\frac{E}{\eta}}\right). \tag{10.16}$$

Coherent FSK systems may take advantage of this extra 0.8 dB in effective S/N, but the orthogonality condition $2\Delta fT = n/2$ (where n is an integer) is usually assumed for other systems, giving the result expressed in Eq. (10.11). When possible, FSK systems are designed to try to use a combination of frequency separation and signaling interval to place the matched filter operating point on a local maximum of the graph shown in Fig. 10.8 (i.e., $2\Delta fT = 3/4, 7/4, 11/4$, etc.).

Derivation of the FSK power spectral density for random waveform inputs is rather involved.[†] However, some general trends in the results are as follows. For low values of $2\Delta fT$, the power spectral density has a single peak centered at the carrier frequency f_c and decreases smoothly to either side of the peak. As $2\Delta fT$ is increased, the central peak in the power spectral density decreases and peaks begin to develop near the deviation frequencies $f_c \pm \Delta f$. For still larger values of $2\Delta fT$, the FSK power spectral density tends toward two separately identifiable spectral groupings centered at $f_c \pm \Delta f$, as illustrated by the line spectrum in Fig. 10.7. The power spectral density is continuous for random binary waveform inputs, with the exception that it will contain impulses if $2\Delta fT = m$, with m an integer. For the particular choice $2\Delta fT = 1$, 50% of the total power in the FSK waveform is in line components at the two transmitted frequencies, and this choice is generally avoided to prevent possible interchannel interference. Systems intended primarily for the more inexpensive noncoherent receivers use $2\Delta fT > 1$, whereas FSK systems intended primarily for

[†] See, for example, R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communications*. New York: McGraw-Hill, 1968, Chapter 8.

coherent detection often use choices in the range $\frac{1}{2} < 2\Delta fT < 1$ to gain some S/N advantage and minimize the bandwidth required.

It might prove tempting to try to approximate the required bandwidth for FSK, using the approximations discovered in frequency modulation. Recall that in analog modulation the bandwidth of the modulated signal is not less than the double-sided bandwidth of the modulating signal and the bandwidth of a frequency-modulated signal is equal to or greater than the bandwidth of an amplitude-modulated signal. Neither of these conclusions is necessarily valid in digital modulation, however; and therefore some caution must be used in approximating the required bandwidth for FSK. For $2\Delta fT > 1$, a rough approximation of the bandwidth can be obtained using Carson's rule with $f_m = 1/T$. For $2\Delta fT < 1$, however, the required bandwidth, although it will always be greater than $2\Delta f$, may be less than the two-sided bandwidth of the modulating signal. These are predetection bandwidths. A postdetection bandwidth can be narrowed to a bandwidth on the order of $1/T$ (as assumed in the matched-filter receiver).

Results of some averaged spectrum analyzer measurements for binary FSK with equiprobable 1's and 0's are shown in Fig. 10.9. Note that Carson's rule for bandwidth works reasonably well for a 90% power containment bandwidth, except for low values of ΔfT, although the bounds are not as well defined as they were in sinusoidal FM. Thus, if $\Delta f \approx 1/T$ then an approximation to the FSK bandwidth for 90% power containment is $B \approx 4/T$.

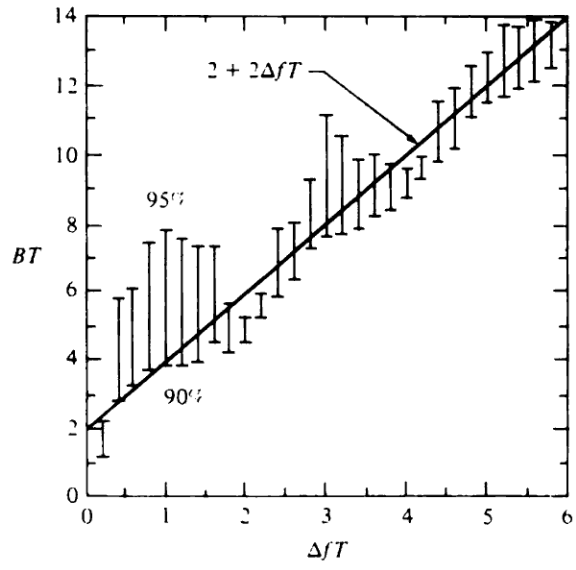


Figure 10.9 FSK bandwidth approximation for 90–95% power containment.

Because it is relatively efficient in terms of peak power requirements and also relatively simple to implement, FSK is almost universally used for low-speed modems. For binary signaling up to 300 bps over commercial telephone channels, the usual choice for the transmit frequencies is 1070, 1270 Hz and 2025, 2225 Hz. Use

of two-frequency pairs permits the use of full-duplex operation over one telephone channel. For modems operating at 1200 bps over commercial telephone channels, the transmit frequencies are 1200, 2200 Hz, and such modems can only transmit in the half-duplex mode. FSK modems intended for data transmissions up to 1800 bps are sometimes used, but they generally require conditioned telephone lines.

EXAMPLE 10.2.1

NRZ binary data is transmitted at 300 bps through a telephone channel using FSK with transmit frequencies of 2025, 2225 Hz.

a) Assuming an 800-Hz bandwidth centered on the carrier, calculate the minimum probability of error if the average signal-to-noise ratio is 8 dB;

b) Repeat for $S/N = 7$ dB.

Solution a) Here we have $f_c = 2125$ Hz, $2\Delta f = 200$ Hz, and $T = 1/300$ sec. Because $\omega_c T \gg 1$ and $\omega_c \gg \Delta\omega$, we can use Eq. (10.15); the bracketed term in this equation gives

$$1 - \mathrm{Sa}\left(2\pi\frac{200}{300}\right) = 1.21\,.$$

Also, we have

$$\frac{S}{800\eta} = 10^{0.8},$$

so that

$$P_\epsilon = Q\left(\sqrt{1.21\frac{ST}{\eta}}\right),$$

$$P_\epsilon = Q(4.51) = 3.3 \times 10^{-6}.$$

b) Changing S/N to $10^{0.7}$, we find the calculation above gives

$$P_\epsilon = Q(4.02) = 3.0 \times 10^{-5}.$$

Comparing the results of (a) and (b) shows that in the assumed range the probability of error changes by about one order of magnitude for a 1-dB change in S/N.

Using two oscillators and switching between them in response to the binary input is easy to visualize, but more popular in operating systems is the application of a polar binary waveform to the input of a voltage-controlled oscillator (VCO), causing the frequency of the oscillator output to vary in response to the input. At the switching times, both methods have sharp, random-magnitude phase transitions, which can

be controlled by filtering and amplitude-limiting. An alternative in the latter method is to band-limit the input to the VCO.

Note that FSK is not true frequency modulation and, in fact, even has more similarities to AM (e.g., see Fig. 10.3). Therefore it does not provide the wideband noise reduction usually associated with FM. Any noise reduction attained arises from the PCM encoding, not from the frequency-shift keying.

10.3 PHASE-SHIFT KEYING (PSK)

Even though a symmetrical distribution about zero was achieved for the FSK operating characteristic, we were not able to obtain the superior probability-of-error performance of the polar-baseband system. It is instructive to pause and to consider the overall binary PCM detection problem to gain some insight into why this is true.

The optimum receiver for binary PCM need only make a decision on the difference between two given alternatives based on an observation over a finite time interval. We therefore return to investigate a matched filter for the detection of the difference between two signals. First, we let

$$g(t) = f_1(t) - f_2(t) \qquad 0 < t \leq T, \tag{10.17}$$

where $f_1(t)$ and $f_2(t)$ are the two signals chosen to convey the binary information. Using Eq. (7.50), we find that the corresponding peak signal-to-noise ratio at the output of the matched filter is

$$\frac{|g_o(T)|^2}{\overline{n_o^2(t)}} = \frac{2E}{\eta} = \frac{1}{\pi\eta}\int_{-\infty}^{\infty} |G(\omega)|^2\, d\omega\,. \tag{10.18}$$

From Parseval's theorem, we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} |G(\omega)|^2\, d\omega = \int_{-\infty}^{\infty} |g(t)|^2\, dt\,. \tag{10.19}$$

Using Eqs. (10.17) and (10.19) in Eq. (10.18) and restricting our attention to the real-valued case, we obtain

$$\begin{aligned}\frac{g_o^2(T)}{\overline{n_o^2(t)}} &= \frac{2}{\eta}\int_0^T [f_1(t) - f_2(t)]^2\, dt \\ &= \frac{2}{\eta}\left[\int_0^T f_1^2(t)\, dt + \int_0^T f_2^2(t)\, dt - 2\int_0^T f_1(t) f_2(t)\, dt\right].\end{aligned} \tag{10.20}$$

The first two integrals in Eq. (10.20) represent the energy in $f_1(t)$ and $f_2(t)$. We now constrain these energies to be equal so that

$$\int_0^T f_1^2(t)\, dt = \int_0^T f_2^2(t)\, dt \triangleq E\,. \tag{10.21}$$

From Eq. (10.20), the maximum peak signal-to-noise ratio for a given signal energy is then obtained for the condition

$$f_2(t) = -f_1(t). \tag{10.22}$$

The optimal class of signals for which Eqs. (10.21) and (10.22) are valid is called *antipodal*. The two signals denoting the two possible information symbols have exactly the same shape but opposite polarity. For linear time-invariant channels corrupted only by additive white gaussian noise, antipodal signals are optimal in the sense of requiring minimum E/η for a specified probability of error.

Recall that Eq. (10.22) does not hold for FSK [Eq. (10.21) does hold] although it did hold for the polar-baseband case. Hence we did not obtain the optimal performance from FSK. On the other hand, it will soon be evident that both Eqs. (10.21) and (10.22) do hold for the case of phase-shift keying (PSK).

In phase-shift keying, the phase of the carrier signal is switched between two (or more) values in response to the PCM code. For binary PCM, a 180° phase shift is a convenient choice because it simplifies the modulator design and is often used. This particular choice is known as *phase-reversal keying* (PRK). The PRK waveform can be written as

$$\phi_1(t) = A \sin \omega_c t, \qquad \phi_2(t) = -A \sin \omega_c t. \tag{10.23}$$

Note that Eqs. (10.21) and (10.22) are satisfied for this type of modulation. A PRK waveform is shown in Fig. 10.10 (also cf. Fig. 7.35).

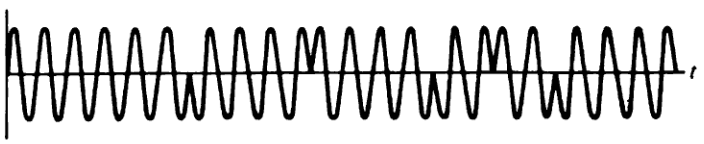


Figure 10.10 A PRK waveform.

Because our choice of signal design obeys Eq. (10.22) we require only one reference function in the correlation detector, as shown in Fig. 10.11. Also, because our choice of signal design obeys Eq. (10.21), the decision threshold is at zero.

Referring to the correlation detector shown in Fig. 10.11, if $[\phi_1(t) + n(t)]$ is present at the input, the output at $t = T$ is $y(T) = E + n_o(T)$. On the other hand, if $[\phi_2(t) + n(t)]$ is present at the input, the output at $t = T$ is $y(T) = -E + n_o(T)$. The variance of the noise is [cf. Eq. (7.54)] $\overline{n_o^2(T)} = \eta E/2$ and the corresponding probability density functions are shown in Fig. 10.12.

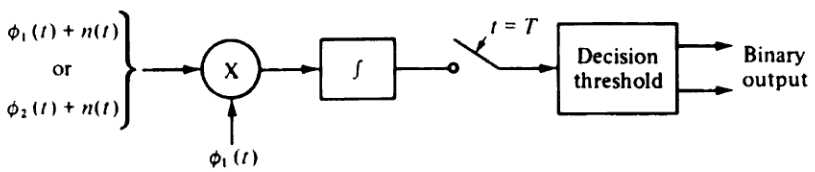


Figure 10.11 A correlation detector for antipodal signals.

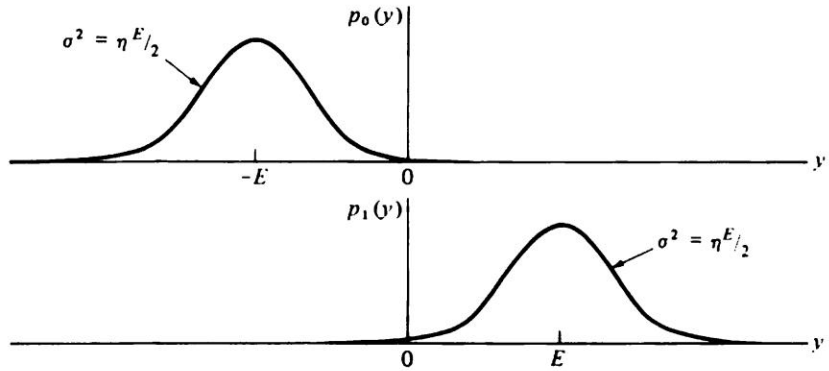


Figure 10.12 Probability density functions for PRK.

As is evident from Fig. 10.12, the optimum-decision threshold is at zero. Thus, the net probability of error (for equiprobable 1's and 0's) is

$$P_\epsilon = \int_0^\infty \frac{1}{\sqrt{\pi\eta E}} e^{-(y+E)^2/(\eta E)}\, dy,$$

$$P_\epsilon = Q\left(\sqrt{\frac{2E}{\eta}}\right). \tag{10.24}$$

Comparison of Eq. (10.24) with Eq. (9.51) shows that PRK has the same probability-of-error performance as polar-baseband systems. Comparison of Eq. (10.24) with Eq. (10.11) shows that to attain a given error probability the average power required for FSK and ASK is twice that required for PRK.

A more general signal representation for binary PSK (BPSK) is

$$\phi(t) = A \cos\left[\omega_c t + \Delta\theta\, p(t)\right], \tag{10.25}$$

where $\Delta\theta$ is the peak phase deviation and $p(t)$ is a binary switching function with possible states of ± 1. Also, it is convenient to define a modulation index, m, for BPSK as

$$m = \cos \Delta\theta, \tag{10.26}$$

where $0 \le m \le 1$. Expanding Eq. (10.25) using a trigonometric identity, yields

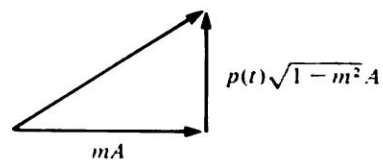
$$\phi(t) = A \cos \omega_c t \cos\left[p(t) \cos^{-1} m\right] - A \sin \omega_c t \sin\left[p(t) \cos^{-1} m\right].$$

Using the fact that $\cos(\pm\cos^{-1} m) = m$, and $\sin(\pm\cos^{-1} m) = \pm\sqrt{1 - m^2}$, we have

$$\phi(t) = mA \cos \omega_c t - p(t)\sqrt{1 - m^2}\, A \sin \omega_c t. \tag{10.27}$$

The first term in Eq. (10.27) is the carrier component, and the second term is the modulation component. A phasor representation of this result is shown in Fig. 10.13.

Figure 10.13 Phasor diagram for general PSK modulation.



The average power in the carrier component of the BPSK waveform is $m^2A^2/2$ and the power in the modulation component is $(1 - m^2)A^2/2$. Therefore the carrier component has the fraction m^2 of the total power in the modulated signal. It follows that the carrier component is zero in a PRK waveform (i.e., one for which $\Delta\theta = \pi/2$) (see also Drill Problem 6.5.1).

To determine the probability of error for BPSK, we use the correlation detector (see Fig. 10.11). Using Eq. (10.25), we find the two possible signal outputs at $t = T$ are

$$\int_0^T A^2 \cos(\omega_c t \pm \Delta\theta) \sin \omega_c t \, dt = \pm\tfrac{1}{2}A^2T \sin \Delta\theta . \tag{10.28}$$

The probability-of-error computation follows in the same manner as for the PRK case [cf. Eq. (10.24)], except that the probability density functions are centered at $\pm E \sin \Delta\theta$:

$$P_\epsilon = \int_0^\infty \frac{1}{\sqrt{\pi\eta E}} e^{-(y+E \sin \Delta\theta)^2/(\eta E)} dy ,$$

$$P_\epsilon = Q(\sqrt{(2E \sin^2 \Delta\theta)/\eta}) ,$$

$$P_\epsilon = Q(\sqrt{2E(1 - m^2)/\eta}) . \tag{10.29}$$

Thus the effect of allocating the fraction m^2 of the total transmitted power to the carrier component is to degrade P_ϵ by an equivalent S/N loss of $10 \log_{10} (1 - m^2)$ dB.

The superior performance of PSK comes with the disadvantage of a need for synchronous detection because the information is in the phase. An advantage in retaining a carrier component in the PSK waveform is that it can be used for receiver synchronization. For example, a phase-locked loop (PLL) can be used to demodulate BPSK if a sufficient carrier component is present. This is done, however, at the expense of a degradation in P_ϵ unless more transmitter power is allocated. A reduction of $\Delta\theta$ from 90° to 63° would result in a 1-dB S/N penalty for fixed P_ϵ while providing $m^2 = 21\%$ of the total power in the carrier component for carrier synchronization.

The power spectral density of PRK is centered around ω_c and has an identical shape to that of the double-sideband modulating spectral density. For NRZ modulation, this power spectral density has the $(\sin x/x)^2$ shape [cf. Sec. 9.9]

$$S_\phi(\omega) = \tfrac{1}{2}A^2T \, \mathrm{Sa}^2[(\omega + \omega_c)T/2] + \tfrac{1}{2}A^2T \, \mathrm{Sa}^2[(\omega - \omega_c)T/2] . \tag{10.30}$$

Even though the power spectral density of the randomly modulated PRK signal is highest around the carrier, there is no discrete spectral line (or impulse) at the carrier frequency. Thus PRK is really a double-sideband, suppressed-carrier modulation technique. In fact, the binary phase modulator can be implemented simply as a balanced mixer with a polar binary input.

For BPSK with $\Delta\theta < \pi/2$, there is a carrier component [cf. Eq. (10.27)], and therefore the spectral density has a discrete spectral line (impulse) at the carrier frequency. In this case the spectral density is analogous to that for double-sideband with carrier, although the carrier component need not be large compared to the sidebands. The theoretical bandwidth efficiency of BPSK systems is 1 bps/Hz. For NRZ modulation, the first-nulls bandwidth is $2/T$, which is commonly used in operating systems designs.

Because the particular choice of $\Delta\theta = \pi/2$ is so widely used in commercial communication systems, it is common to use the abbreviation BPSK to refer to this case. Therefore we will use BPSK to refer to the PRK case unless specified otherwise.

EXAMPLE 10.3.1

Using BPSK (PRK) modulation, the Geostationary Orbiting Experimental Satellite (GOES) series of satellites transmit quantized, meteorological cloud-picture data at a rate of 1.75 Mbps. Assuming $\eta = 1.26 \times 10^{-20}$ W/Hz (corresponding to a net receiving system noise temperature of 229 K) and total path and system losses including antenna gains of 144 dB (cf. Table 4.1), calculate the minimum satellite transmitter power required for $P_e = 10^{-7}$.

Solution Using Eq. (10.24), we require

$$Q\left(\sqrt{\frac{2ST}{\eta}}\right) = 10^{-7}$$

or

$$\frac{2ST}{\eta} = (5.2)^2.$$

Multiplying by the expected losses, we have

$$2S = (5.2)^2(1.26 \times 10^{-20})(1.75 \times 10^6)(10^{14.4})$$

$$S = 75 \text{ W}.$$

Note that the system bandwidth does not enter explicitly into this computation. However, we must remember that our results assume the use of a matched-filter receiver. The bandwidth of such a receiver is matched to that of the signal. If desired, of course, the probability-of-error results for BPSK systems can be rewritten in terms of S/N by noting that $E = ST$ and $N = \eta B$ so that

$$\frac{E}{\eta} = \left(\frac{S}{N}\right)BT,$$

where T is the unit bit duration and B is the receiver noise bandwidth. In most practical systems the receiver noise bandwidth is larger than the double-sided Nyquist bandwidth, so that $BT > 1$. (A factor of 1.2 to 1.4 is often used in satellite systems designs.)

DRILL PROBLEM 10.3.1

Suppose a phase error is present in the correlation detector for a PRK system so that the reference carrier signal is $A \cos(\omega_c t + \psi)$, with ψ the phase error. (a) Derive an

expression for P_ϵ corresponding to Eq. (10.24) for this case. (b) Estimate the phase error ψ required to reduce P_ϵ from 10^{-5} to 10^{-4}.

ANSWER: (a) $Q[\sqrt{(2E \cos^2 \psi)/\eta}\,]$; (b) 29.4°.

Several methods have been proposed for generating a reference carrier signal from a received PRK waveform. One possible solution is to first square the input PRK waveform. The phase of the resulting double-frequency term is determined regardless whether the phase is 0 or $\pm\pi$ radians. A frequency divider is used to obtain the desired carrier reference, as shown in Fig. 10.14. The frequency divider may be mechanized using a phase-locked loop and a divider in the feedback loop. The squaring loop tends to increase the noise near the desired double-frequency carrier component so that a very narrow BPF is required.† For this reason, phase-locked loops (PLL) or crystal-controlled phase-locked loops are often used.

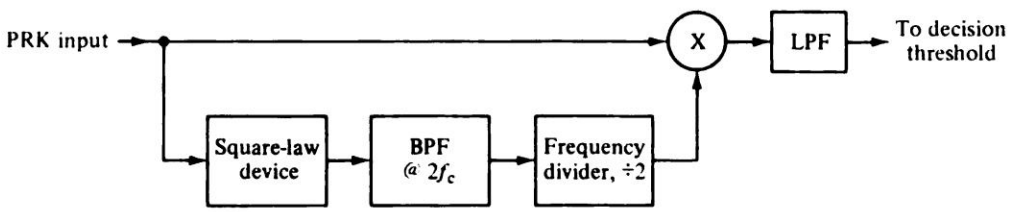


Figure 10.14 A PRK detection system.

Choice of NRZ modulation is a popular one for PRK systems. If NRZ is used, however, a symbol timing recovery system must be used to regenerate the symbol clock frequency. A symbol timing recovery circuit can be constructed in a manner similar to that described for carrier recovery. The first step in the symbol timing recovery loop is applying the bit stream to a nonlinear circuit (e.g., square-law), as shown in Fig. 10.15. A discrete spectral line is present at the symbol rate at the output of the nonlinear circuit and can be filtered with a BPF or PLL. The recovered

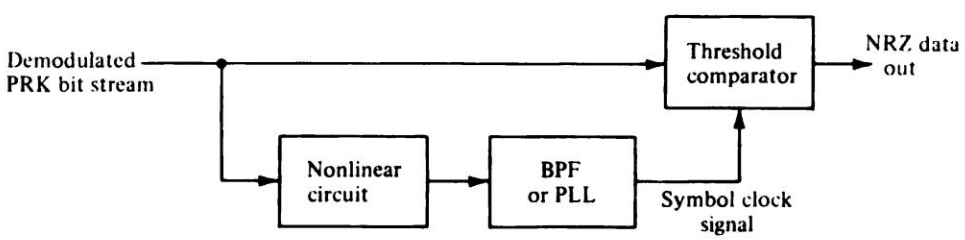


Figure 10.15 A symbol timing recovery system for NRZ modulation.

† The reader will profit from sketching the power spectral density of a carrier in the presence of band-limited white noise and then taking a convolution to see how the noise increases near the double-frequency carrier term.

symbol clock signal with an integrate-and-dump threshold circuit can be used to demodulate the bit stream.

Unfortunately, it is impossible to filter out all of the undesired noise power about the spectral line of the desired symbol clock reference. Those undesired spectral components close to the desired clock frequency are a frequent cause of timing jitter in the recovered symbol clock signal. Timing jitter results in performance degradation because the sampling is not done at the optimum times (i.e., at the maximum eye opening of the demodulated bit stream). Timing jitter tends to accumulate from the regenerative repeaters and is a fairly serious problem in long-haul systems, often placing an upper bound on the total distance of transmission. Circuits called "dejitterizer" circuits can be used to help maintain as low a timing jitter as possible. As a rule of thumb, the maximum permissible timing jitter should not exceed 30% of the symbol duration.[†]

Another method of carrier recovery for PRK, called a *Costas loop*, uses both in-phase and quadrature phase detectors to keep a VCO centered at the suppressed-carrier frequency. A Costas loop is shown in Fig. 10.16. For an input of the form $p(t) \cos(\omega_0 t + \theta)$, where $p(t) = \pm 1$, the loop will track θ while remaining insensitive to the sign of $p(t)$.

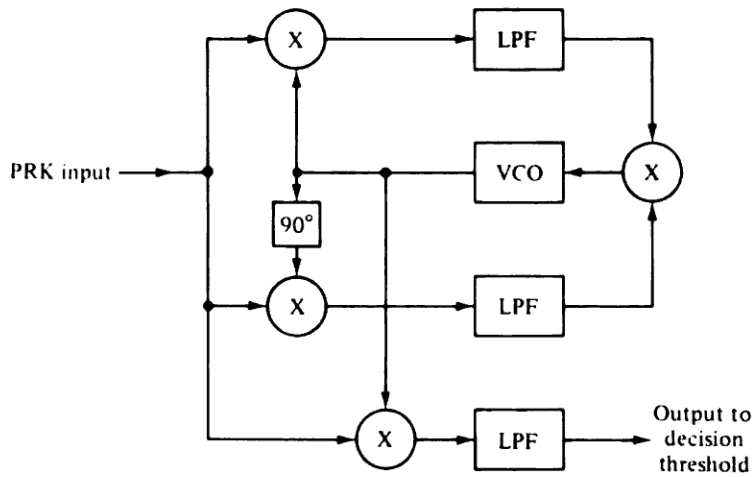


Figure 10.16 The Costas loop.

Both of these methods have a phase ambiguity of $\pm\pi$ radians which can be solved by sending a known preamble code to establish the identity of the initial state. An alternative is to use a modified form of PSK called *differentially encoded PSK* (DE-PSK). With DE-PSK, the information is conveyed via transitions in carrier phase (e.g., no transition for a space and a 180° transition for a mark). Because a bit

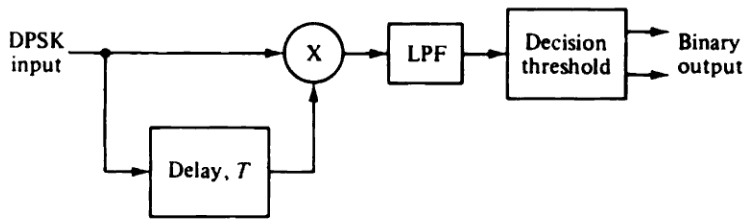
[†] K. Feher, *Digital Modulation Techniques in an Interference Environment*, EMC Encyclopedia, Vol. 9. Gainesville, Va.: Don White Consultants, 1977, Chapter 4.

decision error on the current bit will induce another error on the subsequent bit, the performance of DE-PSK is slightly inferior to that of coherent PSK.[†] The methods above can also be extended to the detection of multilevel PSK.

Another way to get around the synchronization problem is to use a modification of PSK known as *differential PSK* (DPSK). In DPSK, the information is encoded using the differences between bits in two successive bit intervals, as illustrated in Fig. 10.17(a). A differential binary sequence is generated from the input binary message at the transmitter. This sequence has one extra starting digit, which is arbitrary, and which is assumed to be a one here. Succeeding digits in the differential encoding are determined by the rule that there is no change in the output state if a 1 is present. There is a change in output state if a 0 is present.

PCM code input		1	0	1	1	0	1	0	0	1
Differential code	1	1	0	0	0	1	1	0	1	1
Phase	0	0	π	π	π	0	0	π	0	0

(a)



(b)

Figure 10.17 DPSK and its detection.[‡]

A system for the detection of DPSK is shown in Fig. 10.17(b). The phase of the previous digit serves as the reference signal. If the phases are the same, a plus output results; if they differ, a minus output results. Checking back in Fig. 10.17(a), we see that this exactly decodes into the original message. A disadvantage of DPSK is that the signaling speed is fixed by the delay used. Also, because a bit determination is made on the basis of the signal received in two successive bit intervals, there is a tendency for bit errors to occur in pairs. The probability of error of DPSK is[§]

$$P_\epsilon = \frac{1}{2} \exp \frac{-E}{\eta}. \tag{10.31}$$

When compared to PSK systems, those systems using DPSK suffer a signal power penalty of 1 dB (or less) for $P_\epsilon < 10^{-4}$.

[†] See, for example, W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*. Englewood Cliffs, N.J.: Prentice Hall, 1973, p. 252.

[‡] Multiplication here is defined as an equivalence (EXCLUSIVE-NOR logic) operation.

[§] Ibid., p. 248.

10.4 COMPARISON OF BINARY DIGITAL MODULATION SYSTEMS

The net probability-of-error performances of ASK, FSK, and PSK systems, as presented in the preceding sections, are plotted in Fig. 10.18 over a range of error rates of interest in typical operating systems. These results are graphed versus the bit energy E divided by η, the (one-sided) noise power spectral density. Thus the same peak power is assumed for all systems in this comparison. (If the comparison were made in terms of average power requirements, then the ASK schemes would be displaced to the left by 3 dB in the graph and would require about the same amount of power as the FSK schemes.)

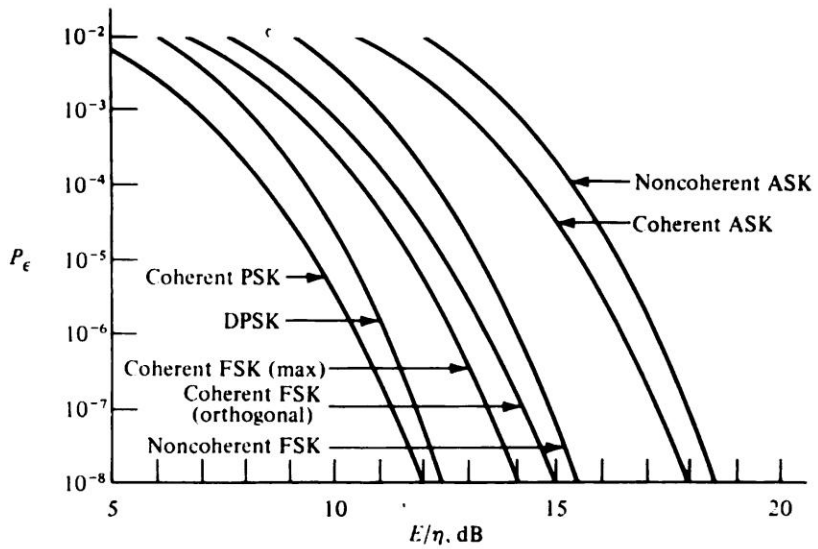


Figure 10.18 Error probabilities for binary digital modulation systems.

Figure 10.18 shows that coherent PSK signaling requires the least amount of power of any binary digital modulation method. Requiring the next least amount of power is DPSK, followed by coherent FSK, noncoherent FSK, coherent ASK, and noncoherent ASK. Whereas differences between some plots are only on the order of 1–2 dB, recall that, in the ranges of practical interest, a change in signal power of 1 dB will result in a corresponding approximate order-of-magnitude change in P_ϵ.

The transmitters for ASK systems are very easy to build and have an advantage in that there is no power transmitted when there is no data being sent (if OOK is used). Such systems find some applications in short-range miniature telemetry systems. Receivers for noncoherent ASK systems are easy to build. The difference in performance between coherent and noncoherent detection is slight compared to the increase in complexity required so that coherent detection of OOK is generally not used. A disadvantage of ASK is that the decision threshold in the receiver must be adjusted with changes in received signal levels. These adjustments are normally made with an automatic gain control.

FSK systems, in contrast to ASK, operate symmetrically about a zero decision-threshold level regardless of carrier signal strength. Besides a possible increase in required frequency stability, there is little difference in the complexity of FSK and PSK transmitters over those for ASK. Receiver complexity depends primarily upon whether a coherent or noncoherent modulation method is used. Noncoherent FSK is relatively easy to instrument and is a popular choice for low-to-medium data-transmission rates such as teletype. FSK transmissions intended for noncoherent demodulation require more bandwidth for a given bit rate than either ASK or PSK. Bandwidths of FSK transmissions intended for coherent demodulation can be made as small as desired by controlling Δf, but cases for $2\Delta fT < \frac{1}{2}$ demand an S/N penalty. Bandwidths of FSK transmissions intended for coherent demodulation are typically about equal to or slightly greater than those required for ASK or PSK.[†]

As we have noted, PSK systems are superior to both ASK and FSK systems in that they require less transmitted power for a given error probability. However, synchronous detection is required, and carrier recovery systems are more difficult (and therefore more expensive) to build. DPSK systems are often a good compromise that sacrifices some error performance but permits a more economical receiver. The three most widely used digital modulation methods for communication systems are PSK, DPSK, and noncoherent FSK.

Not one of the digital modulation methods we have described is particularly efficient in terms of the bandwidth used. Although we discuss this topic in greater detail in later sections of this chapter, we note here that one method we have not mentioned is the use of baseband signal shaping with a linear modulation (such as VSB). This method provides a means of modulation that is more conserving of bandwidth than any of the preceding methods.

The final choice of which type of digital modulation is chosen depends on trade-offs between performance, costs, bandwidth, etc. In addition, propagation distortions (e.g., delay distortion), fading, nonwhite-noise spectra, nongaussian noise (e.g., interference), etc., may affect these choices.

10.5 DIRECT-SEQUENCING SPREAD SPECTRUM SYSTEMS

In some systems we choose to modulate an already modulated signal a second time to gain certain advantages in transmission. An example of this is the case in which several BPSK signals are combined using frequency division multiplexing. In this method the adjacent frequency bands are band-limited and then placed side by side in frequency prior to transmission to gain advantages in transmitting several messages through one communication system. This procedure is discussed in Chapter 5.

[†] A problem in attempting to be more specific than this is that the definition of bandwidth used makes a significant difference; see, for example, F. Amoroso, "The Bandwidth of Digital Data Signals," *IEEE Communications Magazine*, vol. 18, no. 6 (November 1980): 13–24.

Interference between various messages in this method is minimized by the combination of band-limiting and choice of the offset frequencies in the frequency division multiplexing.

A second possibility is time division multiplexing, in which several signals are sampled and interleaved in an orderly way using a common frequency band. Interference between various messages is minimized by use of adequate sampling and filtering to avoid energy spillover into adjacent time slots. The multiplexed signal is then band-limited and may be modulated a second time using a high-frequency carrier.

In contrast to these methods, there is an increasing interest in systems which instead purposely spread out their frequency content to such an extent that each waveform interferes in a barely noticeable way with any other signal operating in the same frequency band. One application of such communication systems is in situations in which the sender does not want the presence of a transmission to be detected by others; e.g., in a military communications system. The goal in such a system is to purposely spread the spectrum out to decrease the transmitted spectral density to an extent that it is well below the thermal noise level of any unfriendly receiver. Also, the transmission may encounter intentional interference ("jamming"), and spread spectrum systems can be used as a protection against such jamming sources.

In commercial communication systems an application of spread spectrum methods is the transmission of two or more unsynchronized signals within a common spectral assignment with only minimal interference. This type of spectral multiplexing is illustrated in Fig. 10.19. This approach is of interest, for example, in a satellite communication system in which a number of users have access to the system and all share a common frequency band. Because the key to the performance of such a system in detection of one signal in the presence of minimum interference from other signals sharing the band is in the signal coding, this type of multiplexing system is called *code division multiple access* (CDMA). CDMA is an important method for transmission of digital data using satellite communication systems.

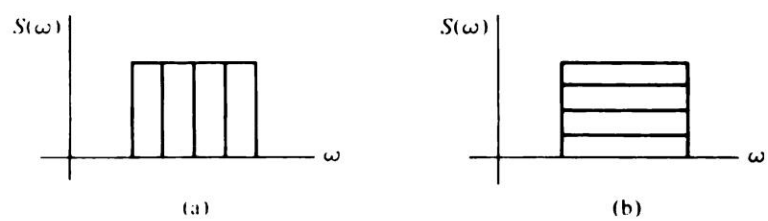


Figure 10.19 Spectra for transmission of several channels of information using (a) frequency division multiplexing, and (b) spread spectrum multiplexing.

In order to classify a system as a spread spectrum system, we require that the system's transmitted signal energy occupy a bandwidth larger than the information bit rate (usually, much larger) and relatively independent of the information bit rate. Demodulation in such systems is accomplished, at least in part, by correlation of the received signal with a replica of the signal used in the transmitter to spread the band-

width. Thus, even though the use of wideband FM also results in a large transmission bandwidth it is not usually classified as spread spectrum.

One method of spreading the spectrum of a modulated signal is to directly modulate that signal a second time using a wideband spreading waveform. Communication systems using direct modulation of a modulated carrier by a wideband spreading signal for purposes of bandwidth spreading are called *direct-sequence* (DS) *spread spectrum* systems. The simplest form of DS spread spectrum employs BPSK as the spreading modulation. For the common choice of $\Delta\theta = \pi/2$, an expression for the BPSK waveform is:

$$\phi(t) = Ap(t) \cos \omega_c t, \tag{10.32}$$

where $p(t)$ is a binary switching function with possible states ± 1 and governed by the modulating (input) signal. The resulting DS spread spectrum signal is

$$\chi(t) = Ac(t)p(t) \cos \omega_c t, \tag{10.33}$$

where $c(t)$ is the spreading modulation waveform. A popular choice for $c(t)$ is that of a pseudorandom noise (PN) binary sequence having values ± 1 (cf. Section 7.13). The block diagram of a system to accomplish this is shown in Fig. 10.20(a).

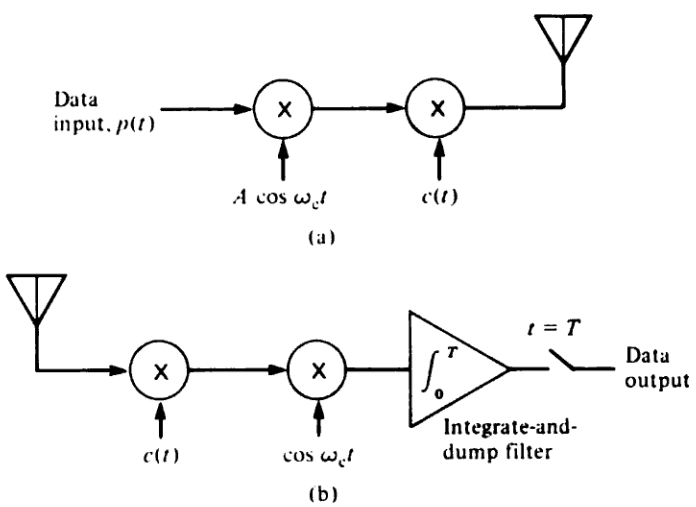


Figure 10.20 DS-BPSK spread spectrum (a) modulation and (b) demodulation systems.

Demodulation of this DS spread spectrum signal requires multiplication by $c(t)$ and by $\cos \omega_c t$, an integrate-and-dump filter, and a binary decision threshold. This is illustrated in Fig. 10.20(b). Phase synchronization between transmitter and receiver is assumed, not only for the BPSK waveform but also for the spreading waveform. Proper synchronization and multiplication of the spreading waveform, as well as the receiver's knowledge of this waveform, is called *despreading*, and is a critical function in spread-spectrum systems.

As discussed in Chapter 7, the PN sequence can be generated readily using feedback shift registers and modulo-2 adders. (We assume use of the NRZ format.) Although the resulting code is deterministic, it has the characteristics of a random binary waveform until the sequence begins to repeat and displays its repetitive pattern. We assume that the sequence length before repetition is very long, compared with other time constants in the system. (Note that it is definitely an advantage to stress this point in the system design if making unwanted detection difficult is one of the objectives.) In this case we may assume that the sequence is random; that is, there is zero correlation between the value of a particular bit and the value of other bits in the sequence. Within these limitations, Example 8.13.2 gives the power spectral density of the spreading waveform $c(t)$ as:

$$S_{ch}(\omega) = T_{ch}\,\mathrm{Sa}^2(\omega T_{ch}/2)\,, \tag{10.34}$$

where T_{ch} is the bit pulse duration in the spreading waveform.

The power spectral density of the data-modulated BPSK signal is [cf. Eq. (10.30)]:

$$S_\phi(\omega) = \tfrac{1}{2}A^2T_b\{\mathrm{Sa}^2[(\omega + \omega_c)T_b/2] + \mathrm{Sa}^2[(\omega - \omega_c)T_b/2]\}\,, \tag{10.35}$$

where T_b is the data bit pulse duration. If statistical independence between the data-modulated carrier and the spreading waveform is assumed, the autocorrelation function of their product equals the product of the autocorrelation functions. Therefore the power spectral density of their product is the convolution of the power spectral densities:

$$S_\chi(\omega) = \frac{1}{2\pi}S_\phi(\omega) \circledast S_{ch}(\omega)\,. \tag{10.36}$$

The bit rate $f_{ch} = 1/T_{ch}$ of the spreading waveform $c(t)$ is usually chosen to be much larger than the data bit rate $f_b = 1/T_b$. To distinguish between the data bit rate and the spreading waveform bit rate, the rate f_{ch} is often called the *chip* rate. As a consequence of the convolution operation, the bandwidth of $S_\chi(\omega)$ is equal to the sum of the bandwidths of $S_\phi(\omega)$ and $S_{ch}(\omega)$. In particular, if $f_{ch} \gg f_b$ then the spreading waveform is the primary control on the net bandwidth of the spread spectrum signal $\chi(t)$.

The data modulation in a DS spread spectrum system does not have to be BPSK. However, it is common to use the same type of digital phase modulation for the data and for the spreading code. When BPSK is used for both modulations, one phase modulator (mixer) can be eliminated, and the binary data and the spreading code can be combined using a modulo-2 sum of the data and the spreading code.

As shown in Fig. 10.20(b), the receiver in the DS spread spectrum system multiplies the incoming signal with the spreading waveform $c(t)$ and with the carrier $\cos \omega_c t$. It is necessary not only to synchronize the receiver carrier frequency oscillator to that of the transmitter but also to synchronize the spreading waveform generator in the receiver to that used in the transmitter.

The procedure for receiver synchronization of the spreading waveform is generally divided into at least two modes: (1) an acquisition mode, and (2) a track (lock) mode. Both modes require that the receiver has (or can generate) a replica of the correct PN

sequence. A commonly used acquisition method is that of the *stepped serial search*. In this method the output of a PN generator in the receiver is multiplied with the incoming bit stream. The resulting signal is applied to a bandpass filter (centered at ω_c), then to an envelope detector and a threshold detector that is preset to trigger if a "good" match occurs. In the absence of such a match, the timing of the PN sequence in the receiver is advanced and the process is repeated. Once an acceptable match is discovered (usually an allowance is made for a few erroneous bits in the sequence to account for possible noise effects), the receiver switches into the lock (track) mode. A third (intermediate) mode may be used to verify the match a second or third time before the system switches to the track (lock) mode.

Once acquisition has been accomplished, the receiver must be able to maintain the correct timing alignment of the received PN sequence and the generated replica to within a fraction of the chip interval. A circuit often used for this purpose is the *delay-locked loop* (DLL) shown in Fig. 10.21. Use of this circuit assumes initial synchronization within one-half chip interval.

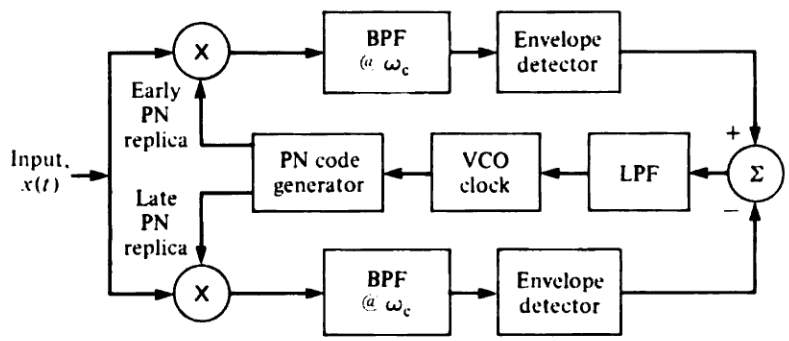


Figure 10.21 Delay-locked loop (DLL) for synchronization of two PN code sequences.

In the DLL, the received signal is multiplied by an early (advanced by one-half chip interval) and a late (retarded by one-half chip interval) replica of the spreading waveform. The resulting early and late products are passed through bandpass filters and envelope detectors, then subtracted to give an error signal. Because $f_c \gg f_b$, the bandpass filter acts as an integrator and the signals at the outputs of the envelope detectors are the magnitudes of the autocorrelation functions of the PN sequence. These two signals are subtracted to form the error signal. The low-pass portion of this error signal should be zero if the early and late replicas are correctly centered about the incoming PN sequence. A correction voltage is obtained by low-pass filtering (to remove harmonic components) and then used to drive a voltage-controlled oscillator (VCO). This oscillator governs the chip rate of the PN sequence generator, which changes the relative timing of the multiplied signals to drive the error signal toward zero. In this manner the replica PN sequence generator tracks the received PN sequence, provided that the initial timing error is relatively small (i.e., less than plus or minus one-half chip interval). This is illustrated in Example 10.5.1.

EXAMPLE 10.5.1

Sketch the low-pass error signal of a DLL for a seven-bit PN sequence.

Solution The autocorrelation function of a seven-bit PN sequence is shown in Fig. 7.52. Advancing and retarding this sequence by one-half a chip interval and then subtracting gives the error signal shown in Fig. 10.22. Note that an unambiguous error signal is generated only if $|\tau| < T_{ch}/2$.

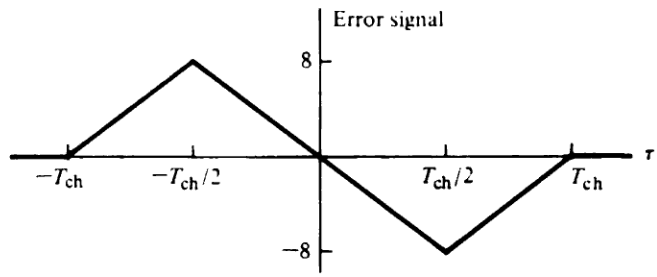


Figure 10.22 Error correction signal for the delay-locked loop of Example 10.5.1.

A spread spectrum system can offer advantages in rejection of various jamming signals. The performance of a DS spread spectrum system in the presence of a jamming signal may be estimated from the average probability of error. Consider, for example, the case in which the intentional interference is a single sinusoidal tone having an average power P_J. The jammer's best strategy is to place the jamming tone directly in the center of the transmission bandwidth. Letting P be the average power of the DS spread spectrum signal and neglecting other sources of additive noise, the sum of the desired signal and the undesired signal at the input of the spread spectrum receiver is

$$v_i(t) = \sqrt{2P}\, p(t)c(t) \cos \omega_c t + \sqrt{2P_J} \cos (\omega_c t + \theta) . \tag{10.37}$$

The second term in this equation is that one contributed by the jamming source. The angle θ is the phase of the jamming signal relative to the data-modulated carrier signal. We assume that the jammer has no knowledge of the phase reference of the data-modulated carrier, and therefore it is reasonable to assume that θ is a random variable with a uniform probability density function.

Because $c^2(t) = 1$, the low-pass portion of the signal $v_o(t)$ at the input to the integrate-and-dump filter in the receiver is

$$v_o(t) = \sqrt{P}\, p(t) + \sqrt{P_J}\, c(t) \cos \theta . \tag{10.38}$$

The first term on the right-hand side of Eq. (10.38) shows that the bandwidth of the data-modulated carrier has been reduced from the very wide bandwidth of $c(t)$ to that of the data $p(t)$. Meanwhile the interfering signal, which originally had a spectral density consisting of an impulse at $\omega = \pm\omega_c$ now appears as a wide-bandwidth spread-spectrum signal. Using Eq. (10.34) and the fact that $\overline{\cos^2\theta} = \frac{1}{2}$ under the assumptions above, the spectral density of the interfering signal is

$$S_J(\omega) = \tfrac{1}{2}P_J T_{ch}\,\mathrm{Sa}^2(\omega T_{ch}/2)\,. \tag{10.39}$$

The zero crossings of this power spectral density are at $\omega = \pm 1/T_{ch}$. For $T_{ch} \ll T_b$ (i.e., $f_{ch} \gg f_b$) the bandwidth of this power spectral density is very wide compared with that of the data-modulated carrier. As an approximation, we assume that it is flat across the bandwidth of interest; that is,

$$S_J(\omega) \approx \tfrac{1}{2}P_J T_{ch} \quad \text{for } |\omega| \ll 2\pi/T_{ch}\,. \tag{10.40}$$

We return to a consideration of Eq. (10.38), and note that the first term is the same as it would be if the system were not a spread spectrum system. If the interfering signal were white gaussian (e.g., thermal) noise having a power spectral density $\eta/2$, the probability of a bit error would be [cf. Eq. (10.24)]:

$$P_\epsilon = Q\left(\sqrt{\frac{E}{\eta/2}}\right). \tag{10.41}$$

As an approximation (it will be, in fact, an upper bound) we can continue to use this result here by replacing $\eta/2$ with $S_J(\omega)$ from Eq. (10.40). Letting $E = PT_b$ in Eq. (10.41) and using $f_{ch} = 1/T_{ch}$ and $f_b = 1/T_b$, we get

$$P_\epsilon \approx Q\left(\sqrt{\frac{2Pf_{ch}}{P_J f_b}}\right). \tag{10.42}$$

Thus the effective jamming power of the interfering source can be offset by making the ratio f_{ch}/f_b as large as possible. This ratio (f_{ch}/f_b) is called the *processing gain* of the spread spectrum system. It gives a simple and convenient measure of how well the system can suppress interfering signals.

DRILL PROBLEM 10.5.1

Retain the power spectral density $\eta/2$ to represent thermal noise. What is the probability of a bit error of a DS-PSK spread spectrum system in the presence of a tone jamming signal?

ANSWER: $P_\epsilon \approx Q\left(\sqrt{\dfrac{2E}{\eta + P_J/f_{ch}}}\right)$.

DRILL PROBLEM 10.5.2

The power spectrum of an unknown DS-PSK signal is observed to have spectral lines with 10 kHz spacings, and the first-nulls bandwidth is 20.7 MHz. Estimate the number of shift registers being used for the PN sequence, assuming that it is a maximal-length sequence.

ANSWER: 10.

10.6 QUADRATURE AM (QAM) AND QUATERNARY PSK (QPSK)

In the binary digital modulation systems we have considered so far, only one of two possible signals can be transmitted during each signaling interval. Such digital modulation systems have a theoretical bandwidth efficiency of 1 bps/Hz. In many applications, a transmission system is more cost effective if, within a given bandwidth, more bits per second can be transmitted. This leads to a consideration of M-ary digital modulation methods in which one of M possible signals are transmitted during each signaling interval.

Our first approach to increasing spectral efficiency is to use the principle of quadrature multiplexing (cf. Example 5.1.1) in which two modulated signals are combined in phase quadrature. A basic system to accomplish this is shown in Fig. 10.23. This is called *quadrature AM* (QAM).

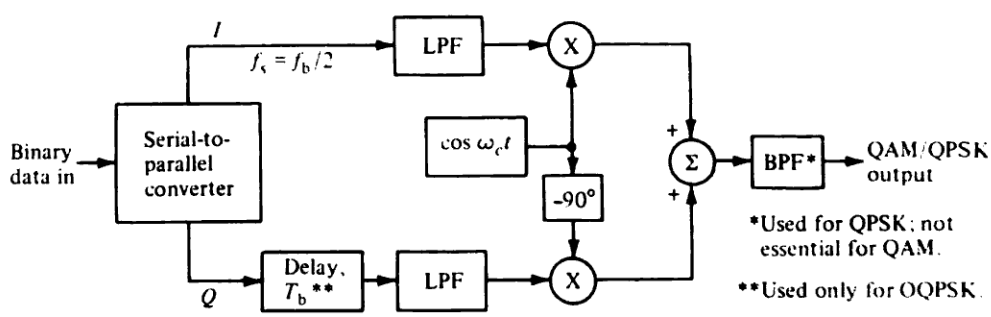


Figure 10.23 Block diagram of a QAM/QPSK modulator.

In the QAM system of Fig. 10.23, the serial-to-parallel converter accepts the incoming binary data stream at a bit rate of $f_b = 1/T_b$ bps, and supplies two parallel binary data streams at $f_s = 1/T_s$ bps, where $T_s = 2T_b$. Here f_s is the symbol rate or *baud*, and f_b is the overall system bit rate. Following the serial-to-parallel conversion, low-pass filters are used to restrict the bandwidth and provide the desired spectral shaping. The I and Q data signals are modulated using DSB-SC modulation; the I signal, using the in-phase carrier reference; and the Q signal, using the quadrature

carrier reference. The modulated I and Q signals are added to form the resultant QAM signal.

Because the QAM signal is the result of the linear addition of two quadrature DSB-SC signals, its spectral density is that of the DSB-SC signals. Thus the QAM power spectral density for a random NRZ input data stream is in the $(\sin x/x)^2$ form given by (for positive frequencies)

$$S_x(\omega) = C\left[\frac{\sin(\omega - \omega_c)T_s/2}{(\omega - \omega_c)T_s/2}\right]^2, \tag{10.43}$$

where C is a constant proportional to the average transmitted QAM power. Because $f_s = f_b/2$, the bandwidth efficiency of QAM is 2 bps/Hz.

A consideration of the phasor diagrams, as shown in Fig. 10.24, reveals that QAM can be viewed as phase modulation if the I and Q signals have identical magnitudes. Such systems are popularly known as *quadrature PSK* (QPSK) systems. As indicated in Fig. 10.24(d), Gray coding is used so that adjacent signal states differ by only one bit.

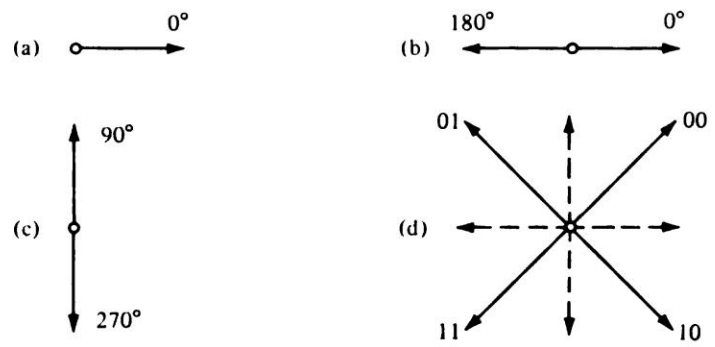


Figure 10.24 Phasor diagrams of QAM/QPSK modulation: (a) unmodulated carrier; (b) in-phase (I) modulated signal; (c) quadrature (Q) modulated signal; (d) composite QAM/QPSK signal (Gray code values indicated).

In QPSK, all of the information is conveyed by the phase, and a constant envelope is desirable. The low-pass filters of the modulator diagram in Fig. 10.23 are used merely to restrict the bandwidth of the data stream to less than one-half the carrier frequency. The major spectral shaping in QPSK systems is accomplished by a bandpass filter after the summation of the I and Q signals. The relative attenuations or gains in the I and Q circuits are kept as equal as possible. Thus a difference between quaternary QAM and QPSK is that the QAM systems employ premodulation (low-pass) filtering for spectral shaping, whereas QPSK systems use postmodulation (bandpass) filtering and attempt to maintain a constant envelope in the modulated waveform. Theoretically, both quaternary QAM and QPSK systems have identical power spectral densities and probabilities of error. As a result of these similarities and

the relative popularity of QPSK systems, the term QPSK is frequently used also for QAM systems.†

In QPSK, one of four possible signal waveforms is transmitted during each symbol interval T_s. These waveforms are as follows:

$$s_1(t) = A \cos \omega_c t,$$

$$s_2(t) = -A \sin \omega_c t,$$

$$s_3(t) = -A \cos \omega_c t,$$

$$s_4(t) = A \sin \omega_c t.$$

These waveforms correspond to phase shifts of 0°, 90°, 180°, and 270°, as shown in the phasor diagram, Fig. 10.24. Because the transmitted signal for a QPSK system can be viewed as the sum of two BPSK signals in quadrature, it is reasonable to expect that demodulation and detection would involve two BPSK receivers, or correlators, in parallel, one for each quadrature carrier. The block diagram of such a system is shown in Fig. 10.25.

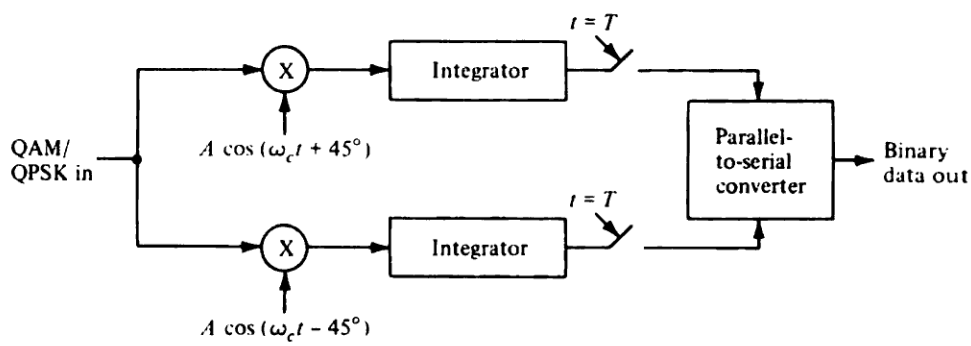


Figure 10.25 Simplified block diagram of a QAM/QPSK demodulator.

Referring to Fig. 10.25, we find that the possible outputs of the first correlator in the absence of noise are

$$y_1(T_s) = \int_0^{T_s} [\pm A \cos \omega_c t][A \cos (\omega_c t + \pi/4)]\, dt, \tag{10.44}$$

$$y_1(T_s) = \frac{\pm A^2 T_s}{2\sqrt{2}} = \frac{\pm E_s}{\sqrt{2}}, \tag{10.45}$$

where E_s is the energy per symbol. In a similar manner, the possible outputs of the second correlator are

$$y_2(T_s) = \frac{\pm E_s}{\sqrt{2}}.$$

† The term *unbalanced QPSK* is sometimes used to describe the special case in which the in-phase and quadrature data rates or powers (or both) are unequal.

The probability of error in the output of the first correlator when additive white gaussian noise is present is calculated in the same manner as that for PRK [cf. Eq. (10.24)], except that the probability density functions are centered at $\pm E_s/\sqrt{2}$, or

$$P_{\epsilon_1} = \int_0^\infty \frac{1}{\sqrt{\pi\eta E}} e^{-(y+E_s/\sqrt{2})^2/\eta E_s} dy,$$

$$P_{\epsilon_1} = Q\left(\sqrt{\frac{E_s}{\eta}}\right). \tag{10.46}$$

Similarly, the probability of error for the output of the second correlator is

$$P_{\epsilon_2} = Q\left(\sqrt{\frac{E_s}{\eta}}\right). \tag{10.47}$$

The probability that the QPSK receiver will correctly identify the transmitted signal is equal to the product of the probabilities that both correlators have yielded correct results, or[†]

$$P_c = (1 - P_{\epsilon_1})(1 - P_{\epsilon_2}). \tag{10.48}$$

Because $P_{\epsilon_1} = P_{\epsilon_2}$, we can write

$$P_c = 1 - 2P_{\epsilon_1} + P_{\epsilon_1}^2, \tag{10.49}$$

$$P_c \approx 1 - 2P_{\epsilon_1} \qquad \text{for } P_{\epsilon_1} \ll 1. \tag{10.50}$$

The probability of error of the QPSK system is

$$P_\epsilon = 1 - P_c \tag{10.51}$$

$$P_\epsilon \approx 2P_{\epsilon_1} = 2Q\left(\sqrt{\frac{E_s}{\eta}}\right). \tag{10.52}$$

Comparing Eq. (10.52) with Eq. (10.24), we conclude that the probability of error for QPSK is greater than that for BPSK (PRK). If, however, we account for the fact that twice as much data is being transmitted, the probability of error performance of both systems is the same. (Note that E_s in Eq. (10.52) is the energy per symbol; it could be rephrased in terms of energy per bit by using $T_s = 2T_b$.) Because some uncertainty can arise in the use of E for both energy per symbol and energy per bit, systems calculations often refer specifically to the *bit error rate*, BER. This is illustrated in Example 10.6.1.

An objective in QPSK signaling is to maintain a constant waveform envelope, letting the information be conveyed by the phase. Phase transitions in each frequency mixer (cf. Fig. 10.23) occur at the symbol interval period and are smoothed by the bandpass filter. In practice, the realizability constraints of the filter cause envelope

[†] This simplification assumes that the noise at the output of the two correlators is uncorrelated. This assumption is valid if the period of integration T extends over one or more complete cycles of the carrier waveform.

fluctuations during the filtering of these phase transitions. Envelope fluctuations are not of major concern in linear channels. However, many QPSK systems designed for satellite and terrestrial use operate with nonlinear output stages. The nonlinear amplifier reduces the envelope fluctuations at the expense of spreading the spectrum, which defeats the purpose of the bandpass filter and can cause unacceptable interference in adjacent frequency bands.

To decrease envelope fluctuations after filtering, some QPSK systems use a one-bit delay (T_b) in the quadrature data stream before modulation, as shown in Fig. 10.23. With this delay, the phase transitions at the two frequency mixers are separated by $T_b = T_s/2$ seconds. Such systems are called offset-keyed QPSK (or OQPSK) systems.† Consideration of a phasor diagram (e.g., see Fig. 10.24) shows that the maximum phase transition at any point in the OQPSK waveform is $\pm 90°$ (although these transitions can occur twice as often as in QPSK waveforms). As a result of the reduced magnitude of the phase transitions, the composite OQPSK signal has smaller envelope fluctuations following the bandpass filter. Another approach is to derive the carrier frequency from a multiple of the symbol timing clock so that the phase transition times can be controlled. This approach offers minimal envelope fluctuations and spectral spreading at some increase in equipment complexity.

It is convenient in summarizing our results at this point, and to prepare for the next section, to set up a more general signal representation for quadrature modulation. In general, the signal waveform for quadrature amplitude modulation can be written as:‡

$$\phi_{\text{QAM}}(t) = A_1 m_1(t) \cos(\omega_c t + \gamma_0) - A_2 m_2(t) \sin(\omega_c t + \gamma_0), \qquad (10.53)$$

where γ_0 is an arbitrary (constant) phase shift. The constants A_1, A_2 permit differing in-phase and quadrature carrier magnitudes, and the amplitude modulation terms $m_1(t), m_2(t)$ permit a (slow) variation in amplitude with time over the symbol interval $(0, T_s)$. A phasor diagram of this QAM signal representation is shown in Fig. 10.26.

In terms of envelope and phase (cf. Fig. 10.26), we can rewrite Eq. (10.53) as

$$\phi_{\text{QAM}}(t) = r(t) \cos[\omega_c t + \gamma(t)], \qquad (10.54)$$

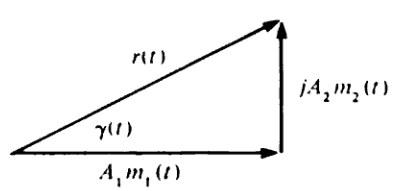


Figure 10.26 Phasor diagram of a QAM waveform.

† OQPSK is also sometimes referred to as staggered QPSK.

‡ The sign of the quadrature term has been chosen here to yield a positive angle in the phasor diagram.

where

$$r(t) = \sqrt{A_1^2 m_1^2(t) + A_2^2 m_2^2(t)}\,, \tag{10.55}$$

$$\gamma(t) = \tan^{-1}\left[\frac{A_2 m_2(t)}{A_1 m_1(t)}\right] + \gamma_0\,. \tag{10.56}$$

For a constant envelope, we require that

$$\left\{\begin{array}{c} A_1 = A_2 \\ \sqrt{m_1^2(t) + m_2^2(t)} = \text{constant} \end{array}\right\}. \tag{10.57}$$

For QPSK and OQPSK systems, this requirement is met by letting $A_1 = A_2$ and by letting $m_1(t), m_2(t)$ be binary switching functions with possible states of ± 1. This results in a phase-modulated signal. If the possible sign changes for $m_1(t), m_2(t)$ are aligned then the resulting modulated signal is QPSK; if they are staggered the modulated signal is OQPSK.

Carrier recovery circuits must be used in QPSK demodulators. Like that for BPSK, one method quadruples the incoming QPSK waveform and then divides by four to derive the carrier (see Problem 10.6.4). A second method is the Costas loop shown in Fig. 10.27. A third method, often called a decision feedback loop, also uses both in-phase and quadrature feedback loops like the Costas loop. However, the bit decisions are made using, for example, such devices as integrate-and-dump filters before the feedback. Systems that use this third method are capable of good carrier recovery, but their acquisition time may be relatively long. In a fourth method, the modulated signal is first demodulated, then remodulated, and finally phase-compared to a delayed replica of the modulated signal. This method offers an advantage in much shorter acquisition times. In all of the feedback-type methods of carrier recovery, an error signal is generated that can be used to correct a VCO in phase and frequency, and then serve as the coherent reference. Other variations of these methods are described in the literature.†

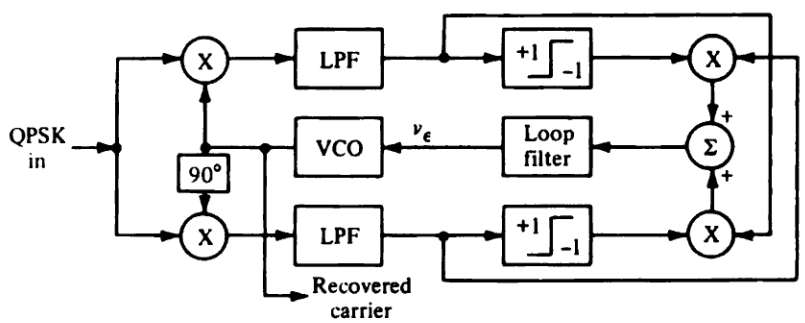


Figure 10.27 Costas loop for QPSK carrier recovery.

† See, for example, W. C. Lindsey, *Synchronization Systems in Communications and Control*. Englewood Cliffs, N.J.: Prentice Hall, 1972; see also J. J. Stiffler, *Theory of Synchronous Communications*. Englewood Cliffs, N.J.: Prentice Hall, 1971.

In summary, QAM and QPSK systems have bandwidth efficiencies up to 2 bps/Hz. They offer attractive compromises between increased data rates with good performance characteristics (acceptable power versus bandwidth trade-offs) and relative ease of implementation. They are widely used in practice for medium-speed data transmission.

EXAMPLE 10.6.1

A given QPSK system is designed to transmit binary data at 44 Mbps in the presence of additive white gaussian noise with one-sided power spectral density $\eta = 1.67 \times 10^{-20}$ W/Hz. Assume Gray coding, a transmitter average power of 9 dB, and an expected net system loss of 120 dB.

a) Compute the bit error rate (BER) for this system.

b) Compare this BER with that of an equivalent BPSK (PRK) system that handles the same data rate.

Solution a) Using Eqs. (9.130) and (10.52), we find the bit error rate for the QPSK system using Gray coding is

$$\begin{aligned}\text{BER} &= Q\left(\sqrt{\frac{ST_s}{\eta}}\right)\\ &= Q\left(\sqrt{\frac{7.94 \times 10^{-12}}{(1.67 \times 10^{-20})(22 \times 10^{6})}}\right)\\ &= Q(4.65) = 1.66 \times 10^{-6}.\end{aligned}$$

b) Using Eq. (10.24), the bit error rate for the BPSK system is

$$\begin{aligned}\text{BER} &= Q\left(\sqrt{\frac{2ST_b}{\eta}}\right)\\ &= Q\left(\sqrt{\frac{2(7.94 \times 10^{-12})}{(1.67 \times 10^{-20})(44 \times 10^{6})}}\right)\\ &= Q(4.65) = 1.66 \times 10^{-6}.\end{aligned}$$

We conclude that the BER for both systems is identical if Gray coding is used for the QPSK system (the usual case). The minimum theoretical bandwidth of the QPSK system is 22 MHz, whereas that for the BPSK system is 44 MHz.

10.7 CONTINUOUS-PHASE FSK (CPFSK) AND MINIMUM-SHIFT KEYING (MSK)†

The primary objective of spectrally efficient modulation methods is to maximize the bandwidth efficiency as measured in bps/Hz. In the preceding section, for example, we found QPSK offers a theoretical bandwidth efficiency of 2 bps/Hz. A second objective is to minimize those spectral components that fall outside the minimum (Nyquist) bandwidth. In a linear channel, this can be accomplished by filtering. In nonlinear channels, OQPSK can be used to reduce the out-of-band spectral components after filtering. Away from the carrier frequency, the power spectral densities of QPSK and OQPSK decrease at a rate proportional to ω^{-2}. With filtering, the rate of fall-off can be increased; but then the waveform envelope may not be constant, and any nonlinearities in the channel tend to restore the initial spectral fall-off characteristics.

In this section, we investigate a modulation method that offers the bandwidth efficiency of QPSK and OQPSK, but whose power spectral density decreases more rapidly beyond the minimum bandwidth. This method is a version of frequency-shift keying discussed in Section 10.2.

Recall that in the transmission of information, the phase of FSK was not exploited, other than to provide possible synchronization of the receiver to the transmitter. Now we show how the phase can be utilized fully to decrease the out-of-band spectral content. To do this, we control the phase in such a way as to avoid any discontinuities in the modulated signal while using the frequency shift to convey the information. This type of digital modulation is called *continuous phase FSK* (CPFSK). It turns out that the power spectral density of CPFSK decreases at a rate proportional to at least ω^{-4} away from the carrier frequency. In addition, CPFSK offers the possibility of a theoretical bandwidth efficiency of 2 bps/Hz. These advantages are gained at the expense of some increase in equipment complexity for modulation and detection.

We define the CPFSK waveform by

$$\phi(t) = A \cos [\omega_c t + \gamma(t)] \tag{10.58}$$

where the phase with respect to carrier, $\gamma(t)$, is a continuous function of time. Rewriting Eq. (10.58) in terms of the frequency shift from carrier, $\Delta\omega$, we have

$$\phi(t) = A \cos [\omega_c t \pm \Delta\omega t + \gamma(0)], \qquad 0 < t \leq T_b, \tag{10.59}$$

† S. Pasupathy, "Minimum Shift Keying: A Spectrally Efficient Modulation," *IEEE Communications Soc. Magazine*, vol. 19 (July 1979): 14–22; see also S. A. Gronemeyer and A. L. McBride, "MSK and Offset QPSK Modulation," *IEEE Transactions on Communications*, COM-24 (August 1976): 809–820.

where

$$\omega_c = \frac{\omega_1 + \omega_2}{2}, \tag{10.60}$$

$$\Delta\omega = \frac{\omega_2 - \omega_1}{2}. \tag{10.61}$$

The two frequencies ω_1, ω_2 are used to represent the binary symbols 0, 1. Comparing Eqs. (10.58) and (10.59), we find that the phase is a linear function of time in the interval $0 < t \leq T_b$, or

$$\gamma(t) = \pm\Delta\omega t + \gamma(0), \qquad 0 < t \leq T_b. \tag{10.62}$$

The initial phase, $\gamma(0)$, depends on the past history of the modulation process and is to be chosen in such a way as to avoid any discontinuities.

In Section 10.2 we stated that the condition for orthogonal keying is that $2\Delta\omega T_b = n\pi$. Since our goal here is spectral efficiency, we choose the minimum condition

$$\Delta\omega T_b = \frac{\pi}{2} \tag{10.63}$$

or

$$2\Delta f T_b = \frac{1}{2}. \tag{10.64}$$

Because this is the minimum frequency spacing between ω_1 and ω_2 that allows the two FSK signal waveforms to be orthogonal to each other, this particular choice of CPFSK is called *minimum shift keying* (MSK). MSK is the only specific case of CPFSK that we shall consider further in this section. Note that the criterion of Eq. (10.64) requires a frequency separation between f_1 and f_2 such that there is a one-half cycle difference in one bit interval.

Using Eqs. (10.62), (10.63), we find that Eq. (10.58) can be rewritten for MSK as

$$\phi_{\text{MSK}}(t) = A \cos\left[\omega_c t \pm \frac{\pi}{2T_b} t + \gamma(0)\right], \qquad 0 \leq t < T_b. \tag{10.65}$$

The instantaneous frequency of this signal is

$$\omega_i = \omega_c \pm \frac{\pi}{2T_b},$$

or

$$f_i = f_c \pm \frac{1}{4T_b}. \tag{10.66}$$

Thus the requirement for minimum-shift keying is maintained over $T_s = 2T_b$, while also maintaining a requirement for phase continuity. Choosing $\gamma(0) = 0$ for conve-

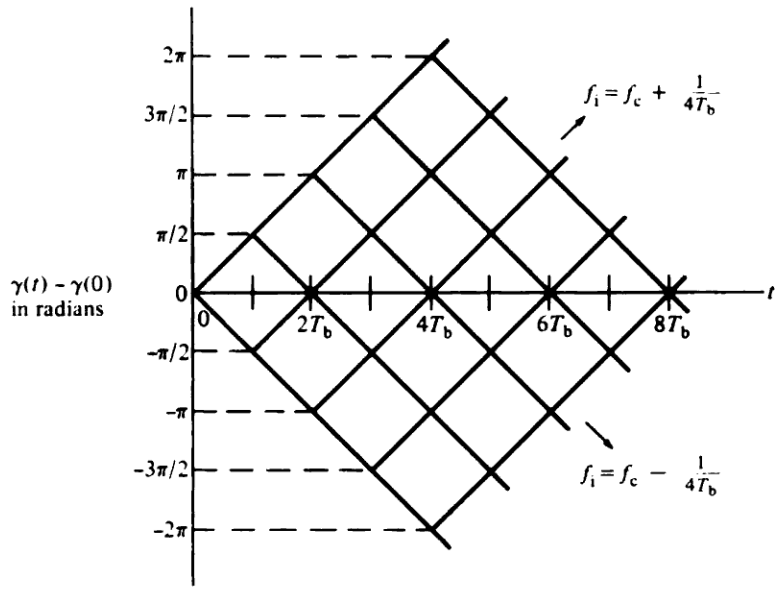


Figure 10.28 Excess phase trellis for MSK.

nience here [recall that $\gamma(0)$ depends on the past history of the modulation process], the possible values of $\gamma(t)$ for $t > 0$ are shown as a phase trellis diagram in Fig. 10.28 for several successive bit intervals. Each path from left to right through the phase trellis corresponds to a specific binary input sequence. Figure 10.28 illustrates that over each bit interval the phase of the MSK waveform is advanced or retarded exactly 90°, depending upon whether the data for that interval is 0 or 1, with respect to the carrier phase in the preceding bit interval. Because all phase shifts are modulo 2π, the phase can take on only the two values $\pm\pi/2$ at odd multiples of T_b and only the values 0, π at even multiples of T_b.

Next we rewrite Eq. (10.65) to extend it over more than one bit interval

$$\gamma(t) = \frac{\pi t}{2T_b} p_k + \gamma_k, \tag{10.67}$$

where

p_k = a binary switching function with possible states of ± 1 representing the binary input data; $k = 1, 2, \ldots,$

$\gamma_k = \gamma[(k-1)T_b]$ = excess phase at the beginning of the kth bit interval required to maintain a continuous waveform.

In addition, we can write the recursive phase constraint (cf. Fig. 10.28)

$$\gamma_k = \gamma_{k-1} + \frac{\pi}{2} p_k. \tag{10.68}$$

Using Eq. (10.67) in Eq. (10.58), we find that the MSK waveform can be expressed as

$$\phi_{\text{MSK}}(t) = A \cos\left[\omega_c t + \frac{\pi t}{2T_b} p_k + \gamma_k\right]. \tag{10.69}$$

Using a trigonometric identity, we can express Eq. (10.69) in terms of in-phase and quadrature components as

$$\phi_{\text{MSK}}(t) = A\left[\cos\left(\frac{\pi t}{2T_b} p_k + \gamma_k\right) \cos \omega_c t - \sin\left(\frac{\pi t}{2T_b} p_k + \gamma_k\right) \sin \omega_c t\right]. \tag{10.70}$$

As in QPSK signaling, we group the input binary stream into data pairs (e.g., assigning all even-indexed sample values to the in-phase component and odd-indexed sample values to the quadrature component). In doing this, we note that for successive values of either of the component values, the excess phase γ_k will always increment by 0, π, modulo 2π. The phases assigned to one component will differ from the other by $\pm\pi/2$ [see Eq. (10.68)] which, except for sign, can be taken care of by a delay of T_b. Therefore our analogy is more closely related to OQPSK signaling. The sign of the needed $\pm\pi/2$ shift is to be chosen to maintain phase continuity.

Using a trigonometric identity and noting that $\sin \gamma_k = 0$ in both the in-phase and quadrature components, we get

$$\phi_{\text{MSK}}(t) = A\left\{\cos \gamma_k \cos \frac{\pi t}{2T_b} \cos \omega_c t - p_k \cos \gamma_k \sin \frac{\pi t}{2T_b} \sin \omega_c t\right\}. \tag{10.71}$$

Consider this representation of the MSK waveform as being composed of two quadrature data channels, $a_I(t)$ and $a_Q(t)$, of an OQPSK signaling system

$$\phi_{\text{MSK}}(t) = A\left\{a_I(t) \cos \frac{\pi t}{2T_b} \cos \omega_c t - a_Q(t) \sin \frac{\pi t}{2T_b} \sin \omega_c t\right\}. \tag{10.72}$$

Now MSK can be thought of as a special case of OQPSK with sinusoidal pulse weighting (rather than rectangular weighting). A pictorial interpretation showing similarities and differences in OQPSK and MSK modulation is shown in Fig. 10.29.

Comparing Eqs. (10.71) and (10.72), we have

$$a_I(t) = \cos \gamma_k ,$$

$$a_Q(t) = p_k \cos \gamma_k ,$$

and recalling that $\gamma_k = 0, \pi$, modulo 2π, we have

$$p_k = a_I(t) a_Q(t). \tag{10.73}$$

The quadrature data channels and sinusoidal weightings can be sketched using Eq. (10.73) and the requirement of phase continuity.

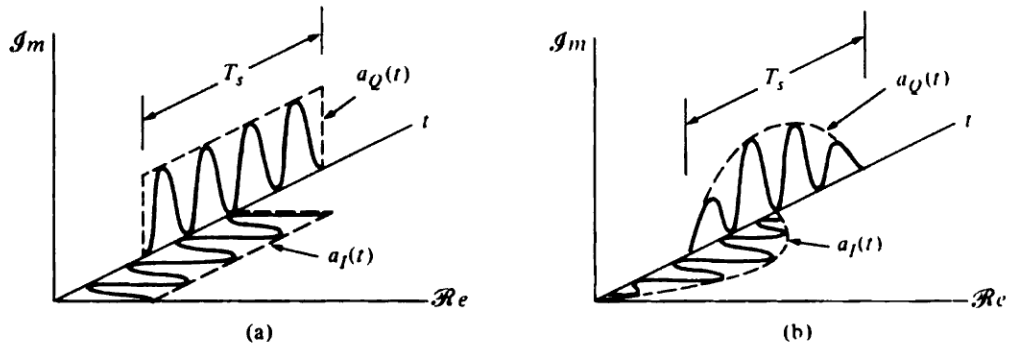


Figure 10.29 An interpretation of (a) OQPSK and (b) MSK in terms of in-phase and quadrature signal waveforms.

Various components of the MSK signal are illustrated in Fig. 10.30 for an input binary sequence 1 0 0 1 0 0 1. The even-indexed sample values in Fig. 10.30(a) are indicated by ± 1, held constant over two bit periods ($T_s = 2T_b$), and weighted by cos ($\pi t/2T_b$). The odd-indexed sample values, offset by one bit period with respect to the even-indexed sample values, held for two bit periods and weighted by sin ($\pi t/2T_b$), are shown in Fig. 10.30(c). The modulated in-phase and quadrature carrier terms are

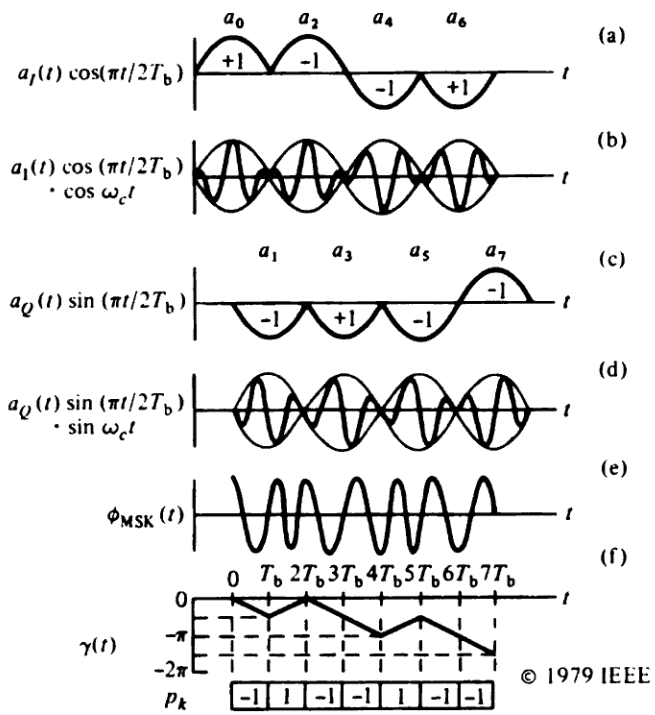


Figure 10.30 MSK waveforms.†

† © 1979 IEEE. Reprinted, with permission, from "Minimum Shift Keying: A Spectrally Efficient Modulation," by S. Pasupathy from *IEEE Communications Magazine*, vol. 19, no. 4 (July 1979): 18.

shown in Fig. 10.30(b), (d), respectively. Subtracting these two waveforms [see Eq. (10.72)] yields the MSK waveform shown in Fig. 10.30(e). The excess phase trellis diagram for this MSK waveform is shown in Fig. 10.30(f). Also, the binary values of the p_k are indicated [these can be found by using Eq. (10.73)].

We return to the use of an excess phase trellis diagram (cf. Fig. 10.28) to illustrate a more specific example of MSK generation. Suppose that not only $\gamma(0) = 0$, but also that the previous value of p_k is $+1$. Assuming that the input data bits $a_I(t)$ and $a_Q(t)$ are ± 1 and staggered, we use the rule in Eq. (10.73) to generate the p_k for each interval of overlap in the parallel bit streams. The resulting trellis diagram, with the 0, 1 binary choices corresponding to the ± 1 of the $a_I(t)$ and $a_Q(t)$ labeled on the diagram, is shown in Fig. 10.31. The path taken in the example in Fig. 10.30 is indicated by the dark lines in Fig. 10.31.

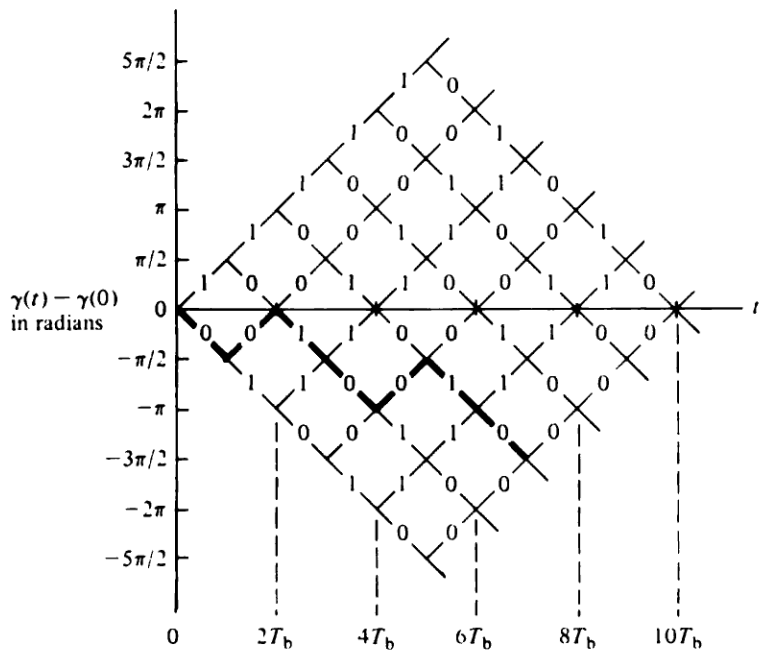


Figure 10.31 Trellis diagram for the example in Fig. 10.30.

Because MSK signaling is essentially equivalent to OQPSK, the probability-of-error performance of MSK using matched-filter detection in the presence of additive white gaussian noise is the same as that for QPSK and OQPSK [see Eq. (10.52)]. From this we conclude that detection of MSK as two orthogonal binary channels provides a 3-dB E/η advantage over detection of orthogonal FSK. Note that the integration time for the matched-filter detection for each of the quadrature signals is $2T_b$ seconds. The bandwidth efficiency of MSK is 2 bps/Hz, the same as for QPSK and OQPSK signaling.

Next we turn our attention to the power spectral density of MSK waveforms. The power spectral density of a random binary waveform with possible levels $\pm A$, deter-

ministic pulse shape $x(t)$ over $(0, T_b)$, and uniformly distributed starting time over $(0, T_b)$ is shown in Example 8.13.2 to be given by

$$S_x(\omega) = \frac{A^2|X(\omega)|^2}{T_b}, \tag{10.74}$$

where $X(\omega)$ is the Fourier transform of $x(t)$. An extension of these methods to in-phase and quadrature components $x(t)$, $y(t)$

$$z(t) = x(t) + jy(t) \tag{10.75}$$

over $(0, T_s)$ with random amplitudes A_1, A_2, respectively, gives the result:

$$S_z(\omega) = \frac{A_1^2|X(\omega)|^2 + A_2^2|Y(\omega)|^2}{T_s}. \tag{10.76}$$

These are complex-valued low-pass signal representations; the power spectral density for the corresponding real bandpass signal

$$\phi(t) = \mathscr{R}e\{z(t)e^{j\omega_c t}\} \tag{10.77}$$

is given by

$$S_\phi(\omega) = \tfrac{1}{2}S_z(\omega - \omega_c) + \tfrac{1}{2}S_z(\omega + \omega_c). \tag{10.78}$$

For BPSK, we have $x(t) = \text{rect}\,(t/T_b)$, and use of Eq. (10.74) gives

$$S_x(\omega) = A^2T_b\,\text{Sa}^2(\omega T_b), \qquad \text{BPSK}. \tag{10.79}$$

Normalizing to the same average power for QPSK and OQPSK as in BPSK, we let $A_1 = A_2 = A/\sqrt{2}$ and $x(t) = \text{rect}\,(t/T_s)$. For QPSK, $y(t) = x(t)$, and for OQPSK, $y(t) = x(t - T_s/2)$; in either case, use of Eq. (10.76) gives

$$S_z(\omega) = 2A^2T_b\,\text{Sa}^2(\omega T_b), \qquad \text{QPSK}. \tag{10.80}$$

For MSK, we have $A_1 = A_2 = A$ and $x(t) = \cos\,(\pi t/T_s)\,\text{rect}\,(t/T_s)$, which yields

$$S_z(\omega) = \frac{16A^2T_b}{\pi^2}\left[\frac{\cos\,(\omega T_b)}{1 - (2\omega T_b/\pi)^2}\right]^2, \qquad \text{MSK}. \tag{10.81}$$

Graphs of Eqs. (10.80), (10.81) are shown in Fig. 10.32. From these graphs, we observe that MSK has a wider main lobe (the first null is at $0.75/T_b$) than QPSK and OQPSK (the first null is at $0.50/T_b$), but the power spectral density of MSK has lower sidelobes than QPSK and OQPSK at frequencies more removed from the carrier frequency. These observations are supported in the comparison of various measures of bandwidth listed in Table 10.1.

The modulation/demodulation circuits for MSK are not a great deal more difficult to implement than those for OQPSK. A block diagram of a MSK modulator is shown in Fig. 10.33. The matched-filter MSK receiver is similar to that used for OQPSK with the exception that the reference signals for the synchronous detection of the in-phase and quadrature components are $\cos\,(\pi t/2T_b)\cos\,\omega_c t$, and $\sin\,(\pi t/2T_b)\sin\,\omega_c t$, respectively. A carrier recovery system for MSK is described in Problem 10.7.5.

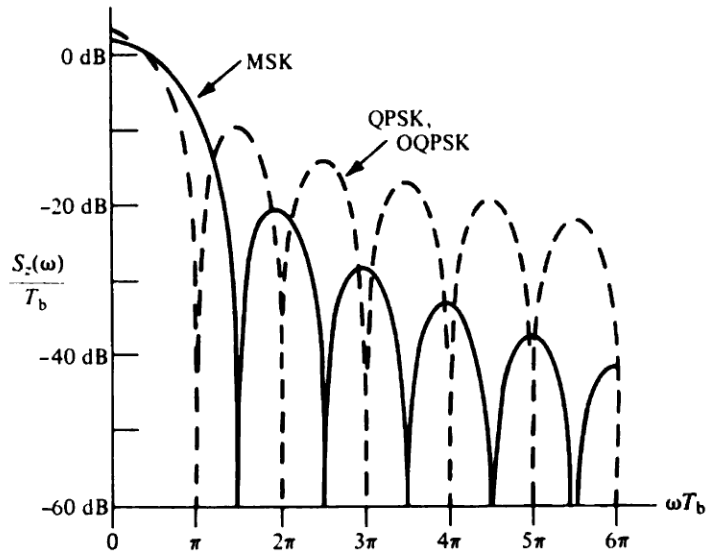


Figure 10.32 Power spectral densities of QPSK, OQPSK, and MSK waveforms.

TABLE 10.1 Bandwidth Comparison of MSK and QPSK Signaling

	Bandwidths		
Modulation Type	−3 dB	−50 dB	99% Power Containment
QPSK, OQPSK	$0.44/T_b$	$100/T_b$	$10.3/T_b$
MSK	$0.59/T_b$	$8.18/T_b$	$1.17/T_b$

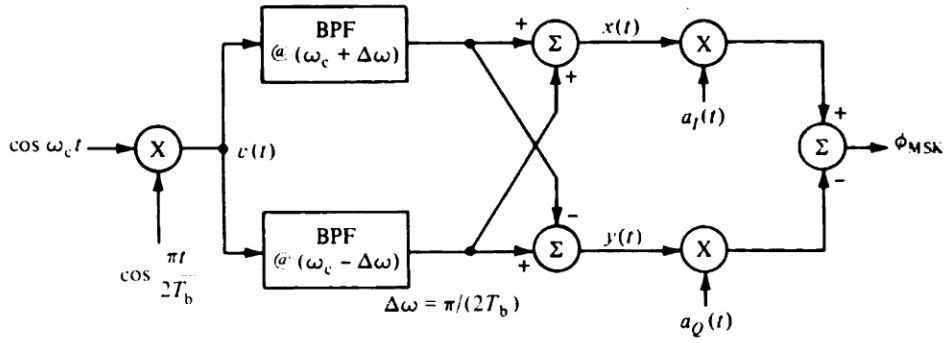


Figure 10.33 A modulator for MSK.

The preceding discussion has focused attention on a parallel modulator structure for generation of MSK. This followed naturally from a discussion of OQPSK. It is interesting to observe that the generation of MSK can be accomplished also using a serial modulator approach. To aid in understanding this approach, first consider the response $g_1(t)$ of a linear time-invariant filter with rectangular impulse response $h(t)$

over $(0, T_b)$ if the input $x(t)$ is also rectangular with duration $(0, T_b)$. As is well known from Chapter 3, the resulting output is triangular over $(0, 2T_b)$ and, aside from the delay effects, is given by the symmetrical form:

$$g_1(t) = \text{rect}\,(t/T_b) \circledast \text{rect}\,(t/T_b) = T_b\Lambda(t/T_b)\,. \tag{10.82}$$

Noting that MSK requires a half-sinusoid weighting over $(0, 2T_b)$ in place of the triangular weighting, we decide to introduce a frequency shift ω_1 in the impulse response; that is,

$$g_2(t) = \text{rect}\,(t/T_b) \circledast e^{j\omega_1 t}\,\text{rect}\,(t/T_b)\,, \tag{10.83}$$

which gives

$$g_2(t) = \frac{2}{\omega_1} e^{j\omega_1 t/2} \sin\,[\omega_1(t + T_b)/2], \qquad -T_b \le t < T_b\,. \tag{10.84}$$

For the desired half-sinusoid weighting for MSK, we choose $\omega_1 T_b = \pi$; using this particular choice of frequency, the filter response $g_2(t)$ is

$$g_2(t) = \frac{2T_b}{\pi} e^{j(\pi/2T_b)t} \cos\left(\frac{\pi}{2T_b}\right)t, \qquad -T_b \le t < T_b\,. \tag{10.85}$$

Observing this result, we see that we have generated the desired half-sinusoid weighting for MSK, but we have an undesired frequency shift (upward) in the result. We seek to delete this shift by heterodyning the input $f(t)$ (downward) by $\omega_1/2$ while keeping the net frequency difference between $f(t)$ and $h(t)$ fixed:

$$g_3(t) = e^{-j\omega_1 t/2}\,\text{rect}\,(t/T_b) \circledast e^{j\omega_1 t/2}\,\text{rect}\,(t/T_b)\,, \tag{10.86}$$

which gives

$$g_3(t) = \frac{2}{\omega_1} \sin\,[\omega_1(t + T_b)/2], \qquad -T_b \le t < T_b\,. \tag{10.87}$$

Again, choosing $\omega_1 T_b = \pi$, we obtain

$$g_3(t) = \frac{2T_b}{\pi} \cos\left(\frac{\pi}{2T_b}\right), \qquad -T_b \le t < T_b\,. \tag{10.88}$$

This accomplishes our objective of the appropriate serial operations to generate the desired MSK weightings for an input data sequence. A block diagram of the serial MSK modulator is shown in Fig. 10.34.

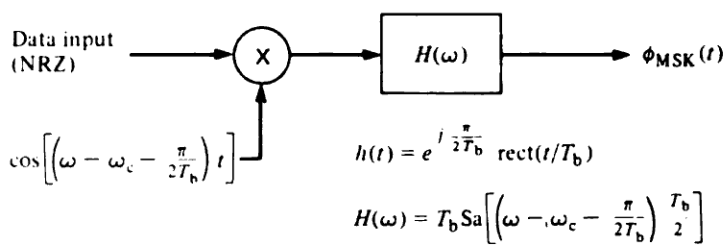


Figure 10.34 A serial modulator for generation of MSK.

A serial demodulator for MSK consists of a bandpass matched filter and a synchronous detector, with the appropriate frequency shifts, and it essentially performs the reverse operations of the serial MSK modulator shown in Fig. 10.34.† The bandpass matched filter must sample at the correct sampling instants for zero intersymbol interference because the time response of the MSK modulator to an input bit of duration T_b lasts for a time $2T_b$. However, the frequency shift has been chosen to yield orthogonality at the correct sampling points. From a theoretical perspective, the performances of the serial and the parallel approach to MSK are identical. Serial MSK modulation and demodulation systems have some advantage in high data rate systems because they do not require a carefully balanced parallel modulator structure for good performance.

In summary, MSK can be considered either as an OQPSK signal with sinusoidal pulse weighting or as a CPFSK signal with a frequency separation ($2\,\Delta f$) equal to one-half the bit rate. The MSK waveform has a continuous phase at the bit transition times and has advantages of a constant envelope and a power spectral density that decreases at a rate proportional to ω^{-4} away from the carrier frequency. The rate of spectral fall-off can be improved by using other weightings in the in-phase and quadrature components, for example, $\mathrm{Sa}(2\pi t/T_b)$, but only at the expense of a wider main lobe in the power spectral density.

CPFSK signaling may be generalized to include other choices of $\Delta\omega$ and the use of longer integration times before a decision is made in detection. However, potential improvements in performance are small and gained at the expense of added system complexity. In general, MSK offers a compromise for digital systems in which both bandwidth conservation and the use of amplitude-saturating (nonlinear) transmitters are important requirements.

EXAMPLE 10.7.1

Show that the system illustrated in Fig. 10.33 is capable of generating an MSK waveform.

Solution At the output of the first frequency mixer, we have the DSB-SC signal

$$v(t) = \tfrac{1}{2}\cos(\omega_c - \Delta\omega)t + \tfrac{1}{2}\cos(\omega_c + \Delta\omega)t,$$

where $\Delta\omega = \pi/(2T_b)$. Each BPF (assume ideal) attenuates one sideband, and the filter outputs are combined to give

$$\begin{aligned} x(t) &= \tfrac{1}{2}\cos(\omega_c - \Delta\omega)t + \tfrac{1}{2}\cos(\omega_c + \Delta\omega)t, \\ &= \cos\Delta\omega t\,\cos\omega_c t; \\ y(t) &= \tfrac{1}{2}\cos(\omega_c - \Delta\omega)t - \tfrac{1}{2}\cos(\omega_c + \Delta\omega)t, \\ &= \sin\Delta\omega t\,\sin\omega_c t. \end{aligned}$$

† See, e.g., R. E. Ziemer and C. R. Ryan, "Minimum-Shift Keyed Modem Implementations for High Data Rates," *IEEE Communications Magazine*, vol. 21 (October 1983): 28; additional references are cited in this paper.

Modulating by the data streams $a_I(t)$ and $a_Q(t)$, respectively, we obtain the desired MSK signal [cf. Eq. (10.72)],

$$\phi_{\text{MSK}}(t) = a_I(t) \cos \Delta\omega t \cos \omega_c t - a_Q(t) \sin \Delta\omega t \sin \omega_c t .$$

DRILL PROBLEM 10.7.1

What relationship should be maintained between the carrier frequency f_c and the bit rate f_b in a MSK system if the bit transitions are to occur at the zero crossings of the carrier?

ANSWER: $f_c = nf_b/4$, n an integer.

10.8 ★ M-ARY ORTHOGONAL FSK

Two basic resources in communication systems are transmitted power and channel bandwidth. In terms of channel capacity, they are related by the Hartley–Shannon theorem. One of these resources may be more precious than the other in a particular situation, and therefore many channels can be classified as either primarily power-limited or band-limited. Preceding sections have emphasized spectrally efficient methods of digital modulation that find application in band-limited channels. In this section we turn our attention to M-ary orthogonal signaling, which may be more useful in power-limited channels to conserve transmitted power at the expense of added bandwidth and equipment complexity.

In a series of M-ary signals, we transmit each one in the interval $(0, T_s)$, and we are interested in uniquely identifying each symbol at the receiver. To do so, we make use of a set of M orthogonal signals, $\phi_n(t)$. These signals are all taken to be of equal energy E_s and satisfy the orthogonality condition [cf. Eq. (2.21)]

$$\int_0^{T_s} \phi_n(t)\phi_m(t)\,dt = \begin{cases} 0 & n \neq m \\ E_s & n = m \end{cases} . \tag{10.89}$$

Although many different sets of orthogonal functions could be chosen, we are interested particularly in frequency-shift keying and therefore choose the set of sinusoids

$$\phi_n(t) = A \cos \omega_n t \qquad 0 < t \leq T_s . \tag{10.90}$$

Using Eq. (10.90) in Eq. (10.89), we require a minimum frequency separation between the signals $\phi_n(t)$ of

$$(\omega_m - \omega_n)_{\min} = \frac{\pi}{T_s} . \tag{10.91}$$

Using this result, we find the minimum net bandwidth needed in signaling using the $\phi_n(t)$ is

$$B_{\min} \approx \frac{M}{2T_s}, \tag{10.92}$$

where

$$T_s = T_b \log_2 M. \tag{10.93}$$

The optimum receiver for this orthogonal signal set consists of a bank of M matched filters as shown in Fig. 10.35. (It is interesting to compare this with Figs. 2.5 and 7.49.) At the times $t = kT_s$, we find which matched-filter output is the largest, and that output symbol is taken as the correct one for a symbol interval. In the presence of noise, some errors in choice will occur. If the noise is additive white gaussian noise with zero mean and with power spectral density $S_n(\omega) = \eta/2$, we can show that the probability of being correct is[†]

$$P_c = \frac{1}{\pi^{M/2}} \int_{-\infty}^{\infty} e^{-z^2} \left(\int_{-\infty}^{z+\sqrt{E_s/\eta}} e^{-y^2} dy \right)^{M-1} dz. \tag{10.94}$$

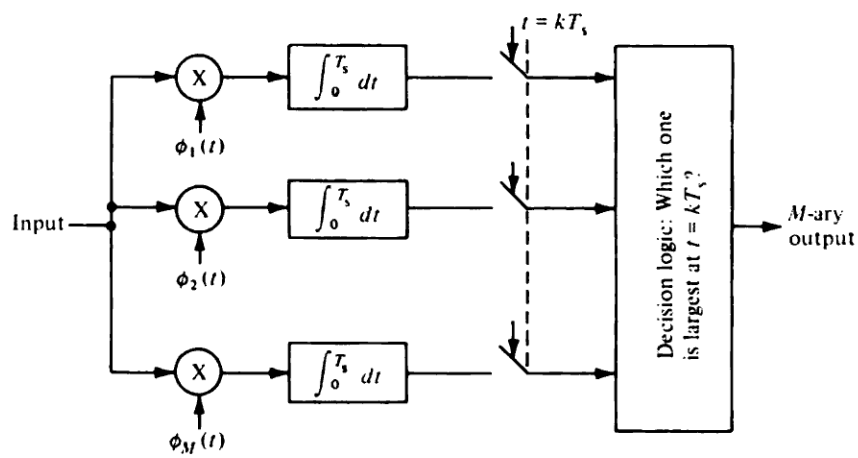


Figure 10.35 Optimum receiver for M-ary orthogonal signaling.

Then the probability of error can be found from

$$P_\epsilon = 1 - P_c. \tag{10.95}$$

The integration in Eq. (10.94) must be evaluated numerically in terms of E_s/η.[‡] Some results are graphed in Fig. 10.36 versus $E_s/(\eta \log_2 M)$.[§] Using both $E_s = ST_s$, where S is the average signal power ($A^2/2$), and Eq. (10.93), we can write

$$\frac{E_s}{\eta \log_2 M} = \frac{ST_b}{\eta}. \tag{10.96}$$

[†] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*. Englewood Cliffs, N.J.: Prentice Hall, 1973, Chapter 5.

[‡] Tables of P_ϵ are given in W. C. Lindsey and M. K. Simon, op. cit.

[§] Note that this could be expressed in terms of an equivalent energy per bit, E_b, by letting $E_b = E_s/\log_2 M$.

Also, some values of $E_s/(\eta \log_2 M)$ for a given probability of error are given in Table 10.2.

From the results of Table 10.2 and Fig. 10.36, we conclude that the average power requirements for fixed η, T_b, and P_ϵ in M-ary orthogonal signaling decrease as M is increased. However, the required bandwidth increases in proportion to M [Eq. (10.92)], and larger values of M tend to increase the complexity of the transmitting and receiving equipment. Therefore M-ary orthogonal signaling conserves power at the expense of bandwidth and equipment complexity.

TABLE 10.2 S/N Requirements of M-ary Orthogonal Signaling for Fixed Error Rates

		$E_s/(\eta \log_2 M)$	
$\log_2 M$	M	$P_\epsilon = 10^{-4}$	$P_\epsilon = 10^{-5}$
1	2	13.8	18.2
2	4	7.94	10.1
3	8	5.82	7.29
4	16	4.72	5.83
5	32	4.04	4.94
6	64	3.58	4.34
7	128	3.25	3.91
8	256	3.00	3.58
9	512	2.81	3.33
10	1024	2.65	3.12
∞	∞	0.69	0.69

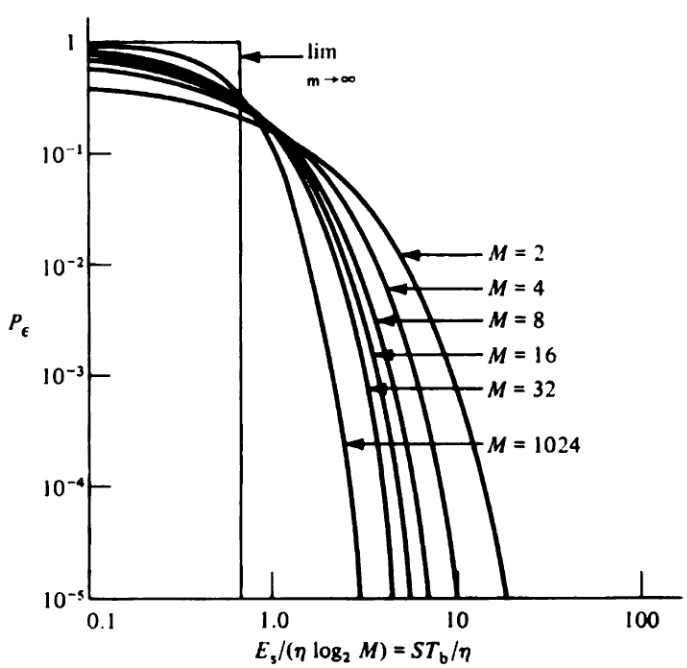


Figure 10.36 Probability of error for M-ary orthogonal signaling.

A more surprising result is that as $M \to \infty$, the M-ary signaling method is capable of transmitting information error-free ($P_\epsilon \to 0$) if $E_s/(\eta \log_2 M) > 1/(\log_2 e)$. In contrast, the probability of error reduces to a certainty ($P_\epsilon \to 1$) if

$$\frac{E_s}{\eta \log_2 M} < \frac{1}{\log_2 e};$$

in this condition, the system fails completely in any attempted transmission of information. Therefore the limiting bit rate f_b as $M \to \infty$ is, using Eq. (10.96),

$$\lim_{M\to\infty} f_b = \frac{S}{\eta} \log_2 e. \tag{10.97}$$

Because the bandwidth is proportional to M, Eq. (10.97) is exactly what was obtained in Example 9.2.3 for the channel capacity of an ideally coded system operating in a channel with no bandwidth limitation. Thus we have succeeded in demonstrating a signaling method which, at least in theory, is capable of approaching the limit of the Shannon–Hartley theory!

Finally, we note that the basic principles of orthogonal signaling are applicable to a much wider choice of orthogonal signals than sinusoids. Our particular interest here is in the sinusoids as an example of FSK modulation. In a broader sense, however, M-ary orthogonal signaling is a method of encoding information and can include more generalized sets of orthogonal signals (see, for example, those in Problem 2.5.3).

EXAMPLE 10.8.1

Determine the savings in required transmitter power, assuming 16-tone orthogonal FSK in place of QPSK signaling in Example 10.6.1 if both systems are operated at $P_\epsilon = 10^{-5}$.

Solution The average power required in the $M = 16$ FSK system is (using Table 10.2)

$$\frac{S_1 T_b}{\eta} = 5.83.$$

The average power of the QPSK system is

$$\frac{2S_2 T_b}{\eta} = 18.2.$$

Taking the ratio, we get

$$\frac{S_1}{S_2} = 2\frac{5.83}{18.2} = 0.641 \ (-1.93 \text{ dB}).$$

The minimum bandwidth of the FSK system is, using Eqs. (10.92) and (10.93), $B \approx 2f_b = 88$ MHz. Also, the 16-tone FSK system uses 16 matched filters (correla-

tors) whereas the QPSK system uses four. Therefore the orthogonal FSK signaling uses about 2 dB less average power than QPSK at the expense of added bandwidth and receiver complexity.

DRILL PROBLEM 10.8.1

What is the maximum bandwidth efficiency of M-ary orthogonal FSK signaling?

ANSWER: $\dfrac{2 \log_2 M}{M}$ bps/Hz.

10.9 FREQUENCY HOPPING (FH) SPREAD SPECTRUM SYSTEMS

An alternative to DS spread spectrum, discussed in Section 10.5, is that of *frequency hopping* (FH) *spread spectrum*. In FH spread spectrum the spectral density of a data-modulated carrier is widened by changing the frequency of the carrier at a rate called the hopping rate. Although other types of modulation may be used, FH spread spectrum systems often use frequency-shift keying for the data modulation. If binary FSK is used, the FH spread spectrum signal can be written as [cf. Eq. (10.13)]:

$$\chi(t) = A \sin \left\{ \int_0^t [\omega_c + (\Delta\omega)p(t)]\, dt \right\},$$

where $\Delta\omega$ is the frequency shift from carrier and $p(t)$ is a binary switching function with possible states ± 1. The carrier frequency, ω_c, in the FH signal is a constant for an interval T_h and then changes to another (preselected) frequency for the next interval of time; the rate at which the changes are made is called the frequency hopping rate, f_h. A block diagram of a modulation system to generate a FH spread spectrum signal is shown in Fig. 10.37(a). The frequency synthesizer produces a constant-amplitude sinusoid whose frequency is determined by the digital code supplied to the synthesizer. It is an essential component of FH spread spectrum systems.

In the receiver the frequency hopping is removed by mixing the FH signal with a local-oscillator signal that is hopping synchronously with the transmitted signal. A block diagram of a FH receiver is shown in Fig. 10.37(b). The incoming signal is mixed with the frequency hopping carrier to produce a data-modulated signal at the intermediate frequency. This intermediate frequency signal is then demodulated using the methods described in preceding sections. As a result of difficulties in building completely coherent frequency synthesizers (and in maintaining phase coherence), most FH spread spectrum systems use noncoherent or differentially coherent modulation.

An advantage of FH spread spectrum is that it allows separate control of the hopping (chip) rate and the bandwidth. Let each carrier frequency be chosen from a set

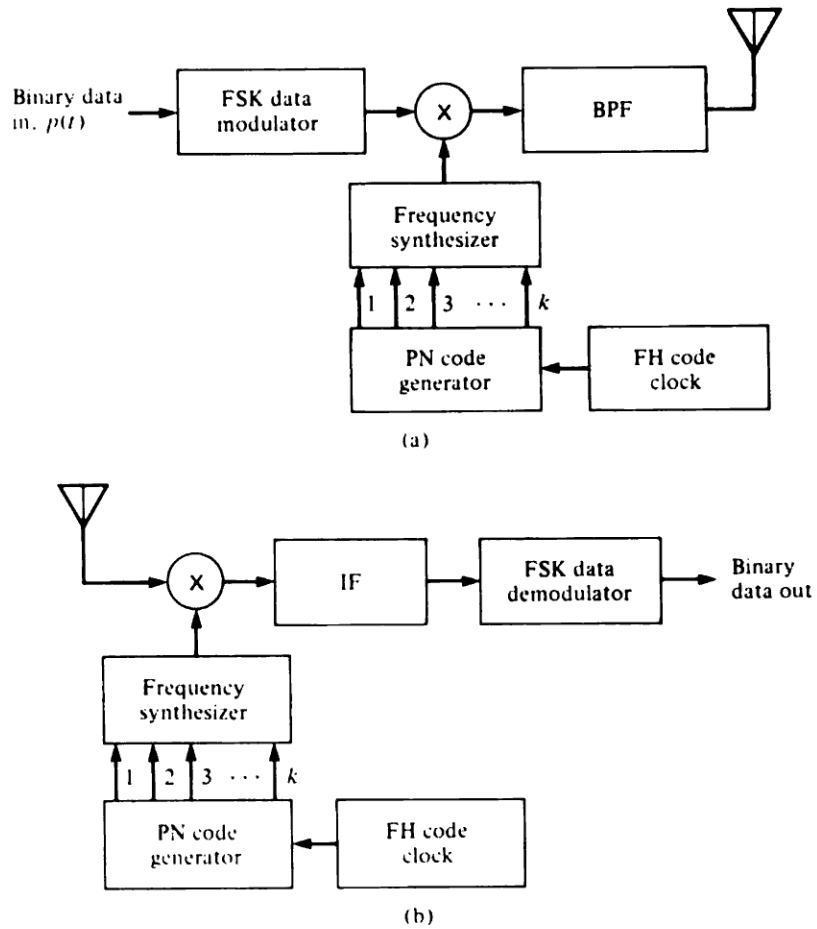


Figure 10.37 (a) Modulation and (b) demodulation of FH spread spectrum.

of 2^k frequencies, each spaced by the bandwidth of the data modulation spectrum. Thus, the bandwidth spreading factor is 2^k times the bandwidth needed for the data modulation without spread spectrum. However, it is not necessary that this be the case. The hopping rate is chosen independently from the bandwidth consideration. A "slow-hop" system is one in which the frequency hopping rate, f_h, is equal to or less than the symbol rate; i.e., $f_h \leq f_b$ for binary FSK. (If M-ary FSK were used, then this would be $f_h \leq f_b/\log_2 M$.) The processing gain of a spread spectrum system is the ratio of the spread spectrum bandwidth to the minimum bandwidth required for the data. Processing gain is a measure of the relative immunity of the system to jamming. Because the energy in the FH spread spectrum system above is distributed over 2^k spectral bands instead of over one modulation spectral band, the approximate processing gain of the spread spectrum system is 2^k. In other words, a jammer would have to place uniform energy across all 2^k bands to cause the receiver to have the same performance as without spread spectrum.

Without coding, FH spread spectrum systems are susceptible to partial-band jamming. The jammer transmits a signal whose bandwidth is relatively small compared

to that of the spread spectrum system and is concentrated in a portion of the band being used. As a result, whenever the FH spread spectrum system chooses that portion of the band for transmission the error rate on the detected data becomes very high. To lessen this effect, the frequency hopping rate can be increased to that of several (or more) frequency hops for each data bit duration. These "fast-hop" systems use frequency hop rates of $f_h > f_b$ for binary FSK ($f_h > f_b \log_2 M$ for M-ary FSK) and thus provide some frequency diversity on each transmitted symbol. However, this may not be sufficient to obtain an acceptable error rate, and error-correcting codes in the data can be used to provide an increased margin against the partial-band jamming.

A problem area in the use of spread spectrum systems is that of synchronization. As in DS spread spectrum, we divide the spreading code synchronization into two modes: (1) acquisition, and (2) tracking. Acquisition for a FSK-FH system is similar to the serial search scheme used for DS spread spectrum, except that the PN sequence generator controls a frequency synthesizer to generate the hopping frequencies instead of directly multiplying by the PN code. A block diagram of an acquisition system for the FH signal is shown in Fig. 10.38. The intermediate frequency (IF) filter of the receiver has a bandwidth on the order of $2f_h$. The envelope detector provides a relative measure of the match between frequency hopping codes, and the threshold detector is set to trigger when that match is acceptable. Lacking an acceptable match, the PN sequence generator is advanced and the process is repeated. A variation, called "camp-and-wait," is to completely stop the sequencing of the PN generator until an initial match is found and then begin the sequence. When an adequate match is found, the system signals that synchronization has been achieved and that tracking can begin. An option is to use a "verify" mode as a third (and intermediate) mode to verify the match a preset number of times before proceeding to the tracking mode.

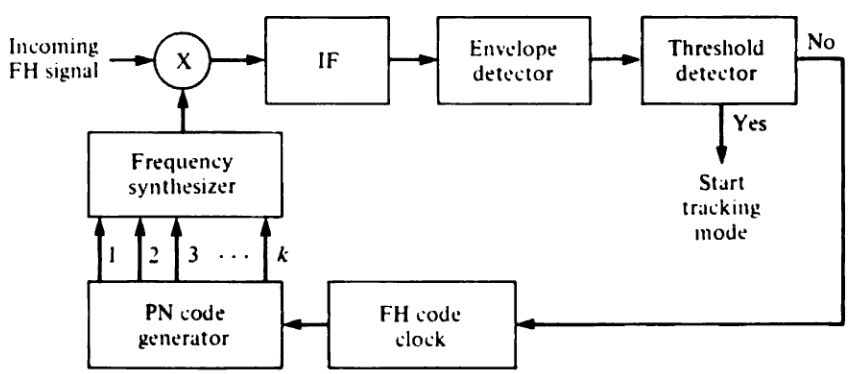


Figure 10.38 Serial-search acquisition for FH spread spectrum.

Once acquisition has been achieved, the synchronization system is activated to track the incoming FH signal. Synchronization for FH spread spectrum can be accomplished by substituting a phase-locked loop (PLL) for the frequency hopping code clock used in the acquisition system. The IF (bandpass) filter output is approximately zero when the two signals to the input mixer are at differing frequencies, and is a maximum when the two frequencies are the same. The resulting output of the

threshold detector is a 0, 1 binary waveform when there is a timing difference between the input FH code and the frequency synthesizer, 0 for that portion of a cycle when the two frequencies differ, and 1 when they are matched. This is illustrated in Fig. 10.39. The PLL locks to this binary waveform, and the VCO output of the PLL is used to drive the PN generator toward closer alignment.

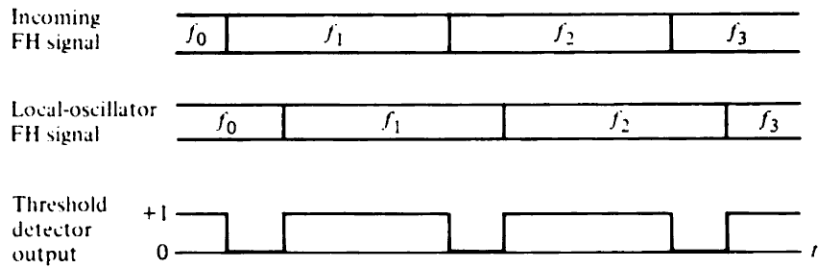


Figure 10.39 Waveforms for FH synchronization (tracking).

Both DS and FH spread spectrum systems are being developed for use in modern secure communication systems. At present, the former tend to be more applicable to commercial satellite systems using a random access approach while the latter are more popular in military systems. Because the accuracy with which the PN codes can be aligned is inversely proportional to the chip duration, these methods also find applications in radar, sonar, and navigation systems for distance ranging.

A third method of spectrum spreading is to employ both DS and FH spreading techniques in a hybrid system. Such hybrid systems offer possibilities of extremely wide spectrum spreading.

EXAMPLE 10.9.1

A FH spread spectrum system uses MSK for the data modulation. Frequency hopping is governed by a maximal-length PN sequence generated by a 10-stage feedback shift register. The register clock rate is 2 kHz and the data rate is 20 kbps. The set of frequencies synthesized in response to the PN code have frequency spacings of 10 kHz.

a) What is the frequency hopping rate?

b) What is the minimum required bandwidth of the system?

Solution a) $f_h = 2$ kHz; because the data modulation rate is 20 kHz, this is a slow-hop system.

b) $B = [(2^{10} - 2)(10) + 20 + 2]$ kHz, or, $B = 10.24$ MHz.

Thus the FH spreading is the dominant effect in determining the bandwidth.

10.10 M-ARY PSK

There is an increasing interest in multi-level modulation methods for digital communication systems that are required to handle high data rates within fixed bandwidth constraints. In this section we investigate the use of M-ary PSK modulation for this purpose. Other methods will be investigated in a later section.

A convenient set of signals for M-ary PSK is the set

$$\phi_i(t) = A \cos(\omega_c t + \theta_i), \qquad 0 < t \le T_s, \tag{10.98}$$

where the M possible phase angles θ_i are chosen as

$$\theta_i = 0, \frac{2\pi}{M}, \frac{4\pi}{M}, \cdots, \frac{2(M-1)\pi}{M}. \tag{10.99}$$

The one-sided power spectral density of M-ary PSK, for a random binary input with equiprobable 1's and 0's at a bit rate $f_b = 1/T_b$, is

$$S_\phi(\omega) = A^2 T_s \operatorname{Sa}^2\left[(\omega - \omega_c)\frac{T_s}{2}\right] \tag{10.100}$$

where T_s is the unit symbol duration given by

$$T_s = T_b \log_2 M. \tag{10.101}$$

Because the minimum required bandwidth is $B = f_s$, the potential bandwidth efficiency of M-ary PSK is

$$\frac{f_b}{B} = \log_2 M \text{ bps/Hz}. \tag{10.102}$$

For example, $M = 8$ PSK signaling systems can transmit data at bandwidth efficiencies up to 3 bps/Hz.

A phasor diagram for M-ary PSK, $M = 8$, is shown in Fig. 10.40(a). All signals in the set have equal energy E_s over a symbol interval $(0, T_s)$, and each signal is demodulated correctly at the receiver if the phase is within $\pm\pi/M$ radians of the correct phase θ_i at the sample time. Thus the receiver decision thresholds are centered in phase angle between the θ_i, as shown in Fig. 10.40(a). No information is contained in the signal amplitude, and the input amplitude is assumed to be limited to a fixed level. All possible signal states (if a noise-free case is assumed) lie at equidistant spacings in a circular pattern centered at the origin, as illustrated for the $M = 8$ case in Fig. 10.40(b). The type of signal-state diagram shown in Fig. 10.40(b) is called a *signal constellation*. Signal constellations for M-ary PSK have circular symmetry, and the usual choice of M is a power of 2.

In the presence of noise, a calculation of the probability of error involves a computation that the received phase lies outside the angular segment $(-\pi/M) \le \theta_i <$

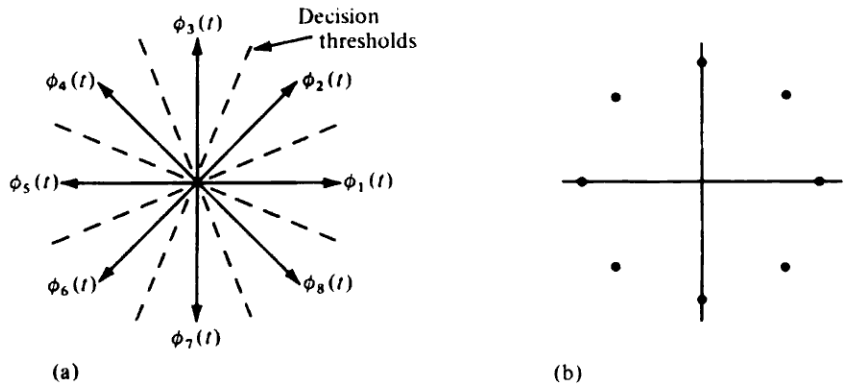


Figure 10.40 Diagrams for M-ary PSK, $M = 8$: (a) phasor diagram; (b) signal state constellation.

(π/M) at the sample time. This is rather difficult to compute exactly.[†] Some numerical results versus $E_s/(\eta \log_2 M)$[‡] are graphed in Fig. 10.41. Also, some values of $E_s/(\eta \log_2 M)$ for a given probability of error are given in Table 10.3. For $P_\epsilon < 10^{-3}$, the probability of error of M-ary PSK can be approximated by[§]

$$P_\epsilon \approx 2Q\left(\sqrt{\frac{2E_s}{\eta}\sin^2\frac{\pi}{M}}\right), \qquad M > 2. \tag{10.103}$$

This approximation improves for fixed M, as E_s/η increases. It should be noted that these results are for the *symbol* probability of error.

For large M, $\sin^2(\pi/M) \approx \pi^2/M^2$; to offset this for a fixed P_ϵ, E_s/η must increase by a factor of 4 (6 dB) for a twofold increase in M. The bit probability of error, P_{be}, can be found by modifying the preceding results according to methods described in Section 9.12. In particular, if a Gray code is used (the usual case),

$$P_{be} \cong P_\epsilon/\log_2 M \tag{10.104}$$

if P_ϵ is acceptably small [see Eq. (9.130)].

Values shown in Table 10.3 indicate that QPSK ($M = 4$) offers a good trade-off between power and bandwidth, requiring very modest increases (0.3–0.4 dB) in transmitted power for a potential doubling of bandwidth efficiency over that of coherent PSK ($M = 2$). For this reason, QPSK is widely used in practice for medium data rate transmissions in band-limited channels, even in some situations in which transmitted power is also a major concern (e.g., high data rate satellite transmissions). For

[†] See, for example, S. Stein and J. J. Jones, *Modern Communication Principles*. New York: McGraw-Hill, 1967, Chapter 14.

[‡] Tables of P_ϵ are given in W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*. Englewood Cliffs, N.J.: Prentice Hall, 1973, p. 232.

[§] W. C. Lindsey and M. K. Simon, op. cit.; see also J. R. Pierce and E. C. Posner, *Introduction to Communication Science and Systems*. New York: Plenum Press, 1980, p. 238.

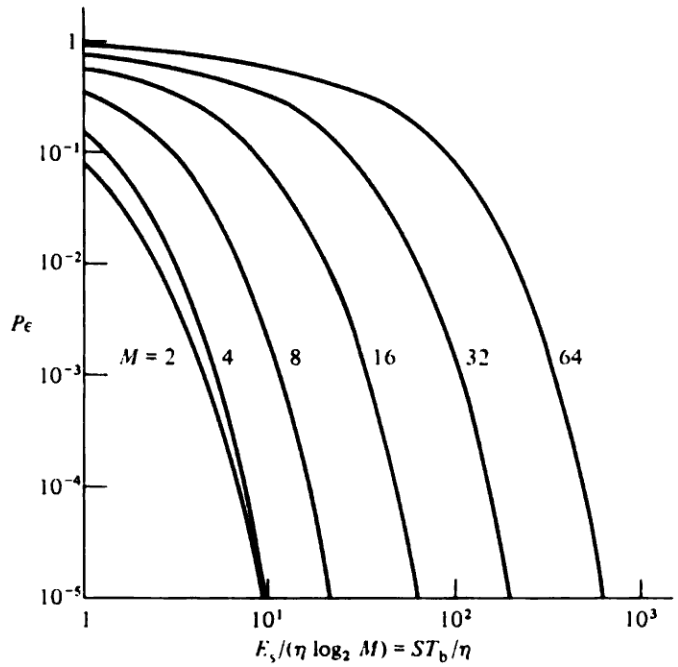


Figure 10.41 Probability of error of M-ary PSK.

TABLE 10.3 S/N Requirements of M-ary PSK for Fixed Error Rates

		$E_s/(\eta \log_2 M)$	
$\log_2 M$	M	$P_\epsilon = 10^{-4}$	$P_\epsilon = 10^{-5}$
1	2	6.92 (8.40 dB)	9.10 (9.59 dB)
2	4	7.57 (8.79 dB)	9.75 (9.89 dB)
3	8	17.2 (12.4 dB)	22.2 (13.5 dB)
4	16	49.7 (17.0 dB)	64.1 (18.1 dB)
5	32	158.0 (22.0 dB)	203.0 (23.1 dB)
6	64	523.0 (27.2 dB)	673.0 (28.3 dB)

higher data rate transmissions in band-limited channels, the choice $M = 8$ is often used. While this choice offers a potential bandwidth efficiency of 3 bps/Hz, it exacts a S/N penalty of almost 4 dB (see Table 10.3) over that required for coherent PSK ($M = 2$) at error rates between 10^{-4} and 10^{-5}. Choices of $M > 8$ are seldom used in M-ary PSK signaling as a result of the excessive additional power requirements.

When compared to BPSK signaling, the use of M-ary PSK requires more complex equipment for signal generation and detection, including that for carrier recovery. The requirement for recovery of the carrier for coherent detection can be alleviated by using a comparison between the phase of two successive symbols. This is the same concept as that described in Section 10.3 for $M = 2$ DPSK signaling and is

called *M-ary differential PSK*. A derivation of the probability of error for M-ary differential PSK is rather involved; an approximation for large signal-to-noise ratios is[†]

$$P_\epsilon \approx 2Q\left(\sqrt{\frac{2E_s}{\eta}\sin^2\frac{\pi}{\sqrt{2}M}}\right). \tag{10.105}$$

A comparison of Eqs. (10.103) and (10.105) shows that, within the approximations, differential detection increases the power requirements by a factor of approximately

$$\Gamma = \frac{\sin^2 \pi/M}{\sin^2 (\pi/\sqrt{2}M)}. \tag{10.106}$$

Using this approximation and $M = 4$, we find the increase in power requirement for differential detection is about 2.5 dB. In some applications, savings in equipment complexity may be worth this moderate increase in power. Note that Eq. (10.106) approaches 3 dB for $M \geq 8$.

EXAMPLE 10.10.1

Compute the bit error rate (BER) of a $M = 8$ PSK system operating under the conditions described in Example 10.6.1.

Solution Using Eqs. (10.103), (10.104), we have

$$P_{be} = \frac{2}{\log_2 M}Q\left(\sqrt{\frac{2ST_s}{\eta}\sin^2\frac{\pi}{8}}\right)$$

$$P_{be} = \frac{2}{3}Q\left(\sqrt{\frac{2(7.94 \times 10^{-12})\sin^2(\pi/8)}{(1.67 \times 10^{-20})(14.67 \times 10^6)}}\right)$$

$$P_{be} = \frac{2}{3}Q(3.08) = 6.91 \times 10^{-4}.$$

DRILL PROBLEM 10.10.1

How many decibels must the transmitter power be increased in the $M = 8$ PSK system in Example 10.10.1 to give the same BER as the QPSK system described in Example 10.6.1?

ANSWER: 3.4 dB.

[†] W. C. Lindsey and M. K. Simon, *op. cit.*, p. 248; tables of P_ϵ using numerical integration procedures are also given.

10.11 AMPLITUDE-PHASE KEYING (APK)

The need for increases in data transmission rates in band-limited channels leads us to a consideration of digital communication systems in which both amplitude and phase modulation are used. This general category of digital modulation is called *amplitude-phase keying* (APK). APK offers the advantage of less required power than PSK for a given probability of error and alphabet size M, but at the cost of increased equipment complexity and a sensitivity to possible nonlinearities in the channel.

Two possible signal-state constellations for 16-ary APK systems are shown in Fig. 10.42. Note that each involves a combination of amplitude and phase modulation. The constellation shown in Fig. 10.42(a) uses three amplitudes, two with four phases and one with eight phases. In contrast, the constellation shown in Fig. 10.42(b) uses four amplitudes and four phases. Both are examples of signal constellations used in 9600 bps modem designs for conditioned telephone channels over the frequency range 300–3000 Hz with 1650 Hz used as the carrier frequency.[†] The constellation shown in Fig. 10.42(a) can be generated by inserting two-to-four-level converters in each of the I and Q signal paths before the balanced modulators and low-pass filters in a QPSK modulator (see Fig. 10.23). Because the two-to-four-level converters reduce the symbol rate by a factor of two and the serial-to-parallel converter reduces the symbol rate by a factor of two, it follows that the potential bandwidth efficiency of such a modulator is 4 bps/Hz.

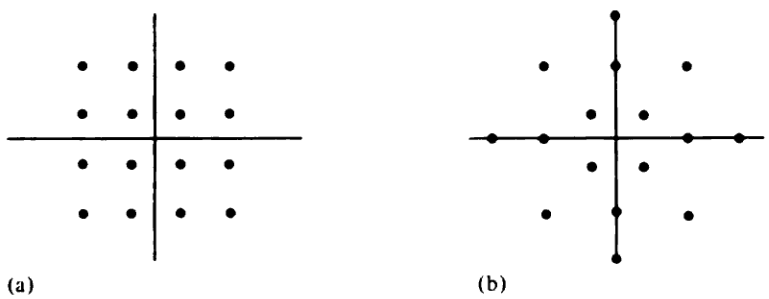


Figure 10.42 Two examples of signal-state constellations for 16-ary APK signaling.

In general, APK can be generated using binary-to-$L = \sqrt{M}$-level converters in each of the I and Q signal paths in a QPSK modulator. In this sense, APK can be considered also to be M-ary QAM. The block diagram of a corresponding M-ary APK demodulator using in-phase and quadrature detection is shown in Fig. 10.43. The L-level A/D flash encoders each consist of $(L - 1)$ comparators, which are set at the various threshold levels. The comparator outputs are sampled at the correct

[†] See, for example, E. R. Kretzmer, "The Evolution of Techniques for Data Communication over Voiceband Channels," *IEEE Communications Society Magazine*, vol. 16 (January 1978): 10–14.

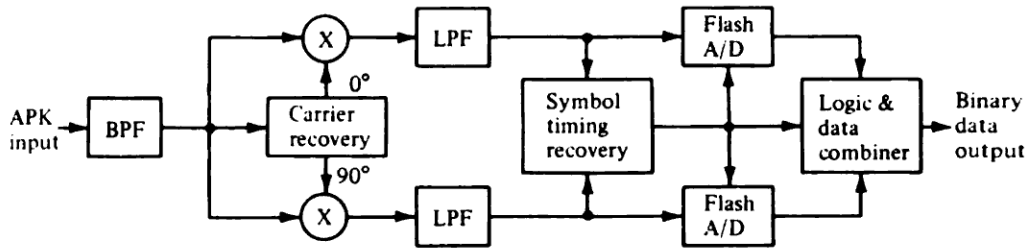


Figure 10.43 Block diagram of an APK demodulator.

sampling time to determine the logic states of the comparators. These states are applied to a logic circuit that determines the output bit stream of the I and Q signal paths.

In band-limited channels with phase distortion and multipath distortion, the demodulator often is preceded by an adaptive equalizer operating at an intermediate frequency (IF). Operation of these equalizers is similar to those described in Section 9.11 for baseband operation, with the exception that equalization is made for both the I and Q signal components. Modern high-speed modems intended for band-limited channels with distortion (e.g., a telephone channel) often consist of an M-ary APK modulator/demodulator with adaptive equalization.

Another possible means of generation, which corresponds to the signal-state constellation shown in Fig. 10.42(a), is to first use a serial-to-parallel converter to divide the incoming data stream into two bit streams, each at half the rate. Then each of these bit streams can be applied to a QPSK modulator, as shown in Fig. 10.44(a). The outputs of the two QPSK modulators are added to yield a 16-ary APK system, as demonstrated by the phasor addition in Fig. 10.44(b). The theory of the corresponding demodulator is more complex and is described in the literature.†

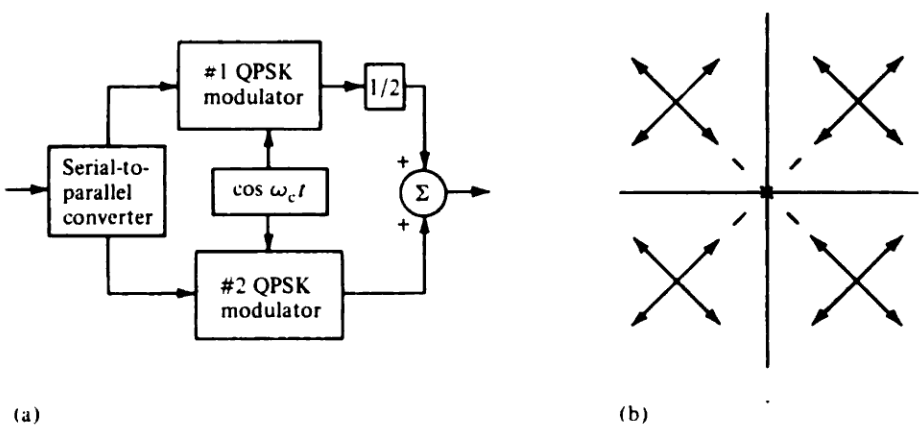


Figure 10.44 Generation of 16-ary APK using the superposed QPSK method: (a) modulator diagram; (b) superimposed phasor diagrams.

† See K. Miyauchi, S. Saki, H. Ishio, "New Technique for Generating and Detecting Multilevel Signal Formats," *IEEE Transactions on Communications*, COM-24 (February 1976): 263–267.

The symbol probability-of-error performance of several PSK and APK systems is shown in Fig. 10.45. Note that 8-ary APK systems can require about 1 dB less power than 8-ary PSK for the same probability of error, while 16-ary APK systems can require about 3.5 dB less power than 16-ary PSK. Therefore a definite advantage develops for use of APK for $M > 8$. The bandwidth efficiency of M-ary APK systems is the same as that of M-ary PSK systems.

Another interesting approach in APK is to make use of the partial-response signaling techniques, which are described in Section 9.10, in both the I and Q signal paths of a QPSK (QAM) modulator. This technique is called *quadrature partial response* (QPR) signaling. Recall that a partial-response system generates an ℓ-level baseband signal data stream x_k using a linear weighted superposition of the binary input. For the specific case of duobinary (also known as Class 1) or modified duobinary (also

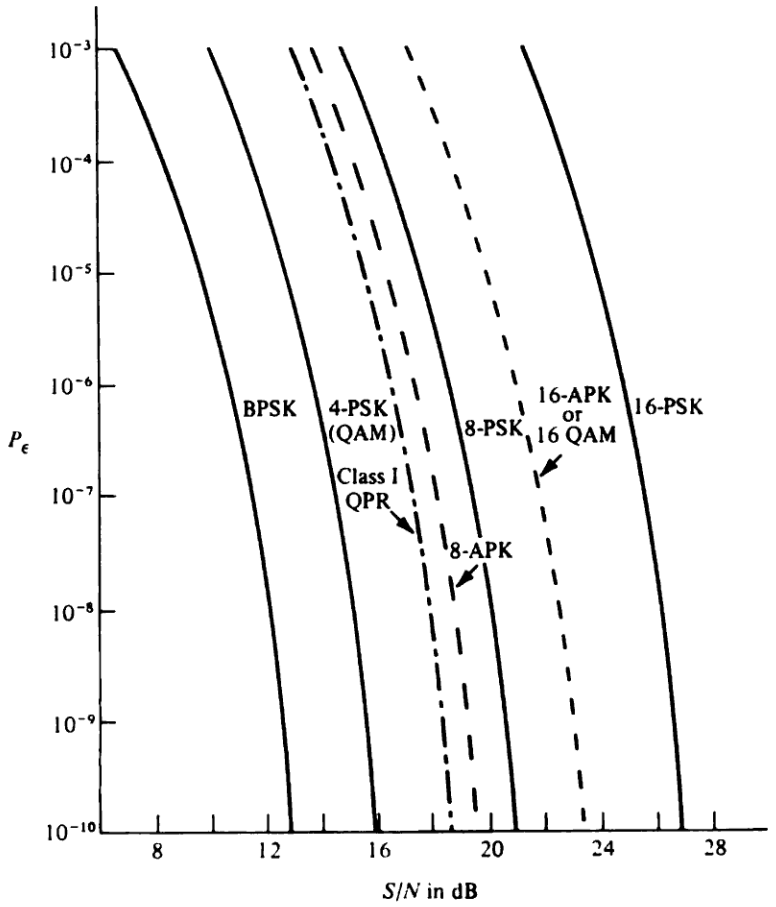


Figure 10.45 Probability-of-error performance of M-ary PSK, QAM, QPR, and M-ary APK coherent systems. The mean-square S/N is specified in the double-sided Nyquist bandwidth that equals the symbol-rate bandwidth.†

† From K. Feher, *Digital Communications: Microwave Applications*, © 1981, p. 71. Reprinted by permission of Prentice Hall, Inc., Englewood Cliffs, New Jersey.

known as Class 4) signaling, three possible output levels result. Thus if both the I and Q signal paths of a QPSK (QAM) modulator use duobinary signaling, the resulting output constellation will have nine signal states, as shown in Fig. 10.46. An advantage of QPR systems is that they have some S/N advantage over comparable PSK systems (see Fig. 10.45) and that they are relatively easy to implement.

In general, problems of carrier recovery for M-ary APK systems are similar to those for M-ary PSK systems. An exception is that the $\times M$ frequency multiplier loop does not lend itself to applications in which the signal-state constellation does not have circular symmetry. One alternative is to use a Costas loop (see Fig. 10.27) with $L = \sqrt{M}$-level converters in each of the I and Q signal paths. Other methods are described in the literature.

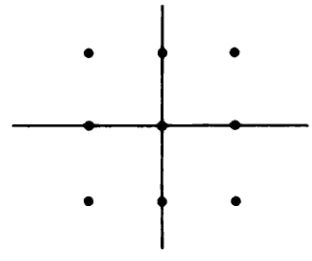


Figure 10.46 Signal-state constellation of a duobinary QPR modulator.

Although M-ary PSK and FSK are very popular and widely used digital modulation methods for current data transmission systems, the APK and QPR methods have been receiving much attention recently. The ever-increasing requirements for data transmission through band-limited channels are expected to make these—and versions of them—popular design choices in new operating systems.

10.12 COMPARISON OF DIGITAL MODULATION SYSTEMS

Choice of digital modulation methods is dependent primarily upon bandwidth efficiency (in bps/Hz), error performance (P_ϵ versus S/N) and equipment complexity (cost). Required transmitted power and equipment complexity generally increase with improved bandwidth efficiency, as emphasized in the preceding sections. Results for bandwidth efficiencies and error performances were given in preceding sections, and representative digital modulation methods are ranked according to their inherent equipment complexity in Fig. 10.47. However, a specific modulation choice may be dependent on other factors, and it is the purpose of this section to delineate some of these other factors.

Not all channels are linear and one factor that will influence the choice of digital modulation method is whether there may be saturation-type nonlinearities present (e.g., those resulting from use of an amplifier operated with a nonlinear gain characteristic). FSK, BPSK, DPSK, OQPSK, and MSK systems are notably good in main-

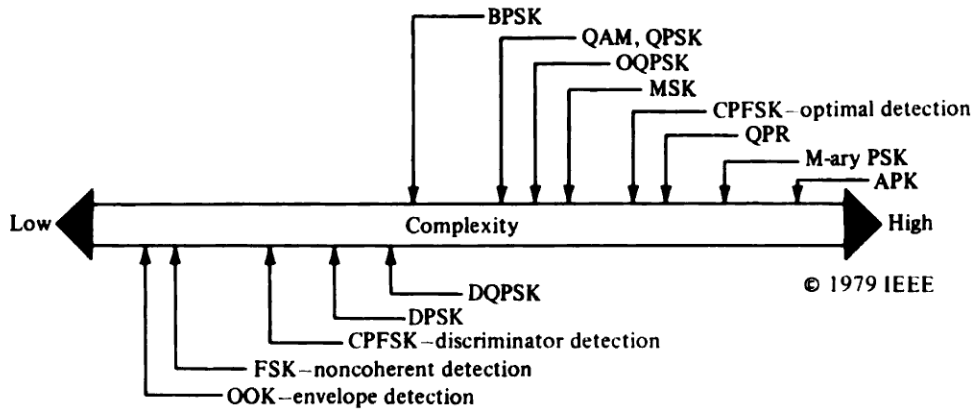


Figure 10.47 Relative complexity of representative modulation schemes.†

taining their performance in such channels, whereas M-ary PSK and M-ary APK systems are poor choices in this respect.

Another type of channel impairment is *delay distortion*. If the magnitude response of a linear time-invariant filter does not vary with frequency, then the slope of the phase-frequency characteristic can be identified with time delay (cf. Section 3.12). For nonideal filters, it is convenient to define the delay as the slope. Any departures of the derivative of the phase-frequency characteristic from this constant are defined as delay distortion. (Appendix E presents some delay-distortion data for commercial telephone channels.) OOK and BPSK maintain good performance in the presence of linear delay distortion, whereas the error performances of QPR and of coherent biorthogonal signaling methods (i.e., QAM, QPSK, OQPSK, MSK) are degraded significantly. For quadratic-type delay distortion, FSK is a good choice, whereas DQPSK suffers severe degradation.

In the presence of fading, the error performances of OOK and the coherent biorthogonal systems do not degrade as rapidly as the others, whereas FSK and the differentially coherent systems are poorer than average. BPSK, DPSK, FSK, VSB, and the coherent biorthogonal systems have above-average tolerance to interference from other signals, whereas OOK and M-ary PSK systems are below average in this respect. BPSK, VSB, and the coherent biorthogonal systems have above-average tolerance to ISI, whereas M-ary PSK is below average. For some of the newer methods (e.g., M-ary APK), not as much data have been published to date and, therefore, not many comments can be made with respect to them.

Finally, we note that the performance characteristics in the preceding sections dealt with idealized cases. Examples of error performances and signaling speeds of some systems reported in the literature are given in Table 10.4. Observe that deratings of 1.0–1.1 dB in E_b/η and 80–90% in signaling speed (bandwidth efficiency) are typical.

† © 1979 IEEE. Reprinted, with permission, from "A Comparison of Modulation Techniques for Digital Radio," J. D. Oetting, *IEEE Transactions on Communications*, COM-27 (December 1979): 1757.

TABLE 10.4 Error Performance and Signaling Speeds of Representative Digital Modulation Methods, 10^{-4} BER†

Modulation Type	Ideal System E_b/η (dB)	Band-limited System E_b/η (dB)	Band-limited System B.W. Eff. (bps/Hz)
BPSK	8.4	9.4	0.8
DE-PSK	8.9	9.9	0.8
DPSK	9.3	10.6	0.8
OOK-coherent detection[a]	11.4	12.5	0.8
MSK	8.4	9.4	1.9
QAM	8.4	9.5	1.7
QPSK	8.4	9.9	1.9
MSK-differential encoding	9.4	10.4	1.9
DQPSK	10.7	11.8	1.8
8-ary PSK	11.8	12.8	2.6
QPR	10.7	11.7	2.25
16-ary APK	12.4	13.4	3.1
16-ary PSK	16.2	17.2	2.9

[a] Average E_b/η is used here so that the values shown for OOK are 3 dB below the peak values indicated in Fig. 10.23.

10.13 ★ REPRESENTATION OF DIGITAL WAVEFORMS

The previous sections have described several specific methods of transmitting digital waveforms. Now we wish to address the more general problems of waveform design and subsequent detection of digital signals.

In digital communication systems there is a finite collection of distinguishable waveforms from which one is chosen at periodic intervals and transmitted. A convenient approach in the analysis of such systems is that of a signal space and vector space—a concept introduced in Chapter 2. In particular, we wish to express a given signal in terms of a set of orthonormal functions. A motivation behind this is that there is a countable set of possible waveforms, and if the receiver is set to recognize one member of the set, then it will not respond to the other members in the set. We use an orthonormal set so that the energy in the waveform will not depend on which member of the set is used.

Let the interval for transmitting each digital waveform be $(0, T)$. The orthonormal basis functions are then [cf. Eq. (2.21)]‡

$$\int_0^T \phi_i(t)\phi_j(t)\,dt = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \tag{10.107}$$

For convenience, we restrict our attention to the case of a set of two orthonormal functions; extensions to a larger orthonormal set are straightforward.

† J. D. Oetting, op. cit., p. 1755.

‡ For simplicity in notation, we use only real-valued functions here.

The signal waveforms $s_m(t)$ may be represented by the orthonormal set [cf. Eq. (2.22)],

$$s_m(t) = a_{m1}\phi_1(t) + a_{m2}\phi_2(t), \tag{10.108}$$

where the coefficients a_{mi} can be found from [cf. Eq. (2.26)]

$$a_{mi} = \int_0^T s_m(t)\phi_i(t)\,dt. \tag{10.109}$$

Thus we can represent the two signals $s_1(t)$, $s_2(t)$ by

$$\begin{aligned} s_1(t) &= a_{11}\phi_1(t) + a_{12}\phi_2(t), \\ s_2(t) &= a_{21}\phi_1(t) + a_{22}\phi_2(t). \end{aligned} \tag{10.110}$$

The energy in each waveform is [using Eqs. (10.107) and (10.108)]

$$\begin{aligned} \int_0^T s_1^2(t)\,dt &= a_{11}^2 + a_{12}^2, \\ \int_0^T s_2^2(t)\,dt &= a_{21}^2 + a_{22}^2. \end{aligned} \tag{10.111}$$

In a similar way, the energy in the difference between signals is

$$\int_0^T [s_1(t) - s_2(t)]^2\,dt = (a_{11} - a_{21})^2 + (a_{12} - a_{22})^2. \tag{10.112}$$

These concepts can be extended to waveforms constructed using more than two orthonormal basis functions.

A geometric interpretation can be given to these results by considering the orthonormal set $\phi_n(t)$ as an orthonormal set of basis vectors, $\boldsymbol{\phi}_n$. A given vector in this vector space can be expressed in terms of the basis vectors [cf. Eq. (2.18)] and the coefficients can be found in terms of the dot product between the given vector and the basis vector [cf. Eq. (2.19)]. In addition, Eqs. (10.111) and (10.112) show that distance (or length—called the "norm" of the space) in the vector space is analogous to the square root of the signal energy in the signal space.

The received signals have additive random noise present and this noise may also be represented in terms of the set of orthonormal basis functions. Because the noise is random, the coefficients in the orthonormal series representation of the noise are random. The coefficients are uncorrelated with each other. If the noise is white with zero mean value, it can be shown that the mean-square value of each coefficient is equal to the mean-square value of the noise.[†] Only those noise coefficients which use the same orthonormal functions as the signal waveforms are important—the remaining ones are noise-only terms and rejected by the receiver.

A geometric representation in two dimensions of received signal-plus-noise is shown in Fig. 10.48. Because the noise coefficients n_1, n_2 are uncorrelated and have

[†] See, for example, H. L. Van Trees, *Detection, Estimation and Linear Modulation Theory, Part I*. New York: John Wiley & Sons, 1968, Chapter 3.

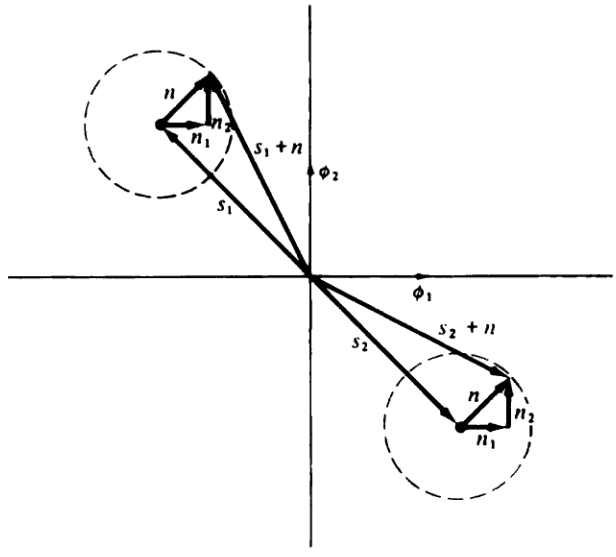


Figure 10.48 Geometric representation of signal-plus-noise.

equal variances, the effect of the noise is to spread out the possible positions of the received vectors over a region around the noiseless signal point. The location of the resultant can be described statistically; the dashed circles in Fig. 10.48 are used to outline equiprobable loci.

The noise distribution does not depend on the signal. Therefore the transmitted signal states should be located as far apart as possible. If two waveforms are used (as shown in Fig. 10.48), the best placement within a fixed energy constraint is such that they are collinear and symmetrically opposite the origin; that is, $s_1(t) = -s_2(t)$. This type of waveform choice is called *antipodal* signaling. In contrast, an orthogonal waveform system provides maximum receiver output for the desired waveform while giving zero output for the others in the set. Thus antipodal signaling provides a lower error rate for a given average power.

We reconsider the ASK, FSK, BPSK, and QPSK digital waveforms which were used in previous sections but now in a more general context. Geometric representations of these are shown in Fig. 10.49 for a fixed average power (assuming equiprobable ones and zeros) and fixed noise power. The ASK waveforms are obtained by using one-sided collinear vectors (OOK is shown). The FSK waveforms are chosen to be orthogonal over $(0, T)$ while the BPSK waveforms are an example of antipodal signaling.

From Fig. 10.49, it is easy to see the relative probability-of-error advantage of BPSK over that of ASK and FSK. Besides the physical intuition it lends, the geometric representation also has the advantage that it can easily be extended to higher-dimensional waveforms.[†]

[†] For further reading on this topic, see E. Arthurs and H. Dym, "On the Optimum Detection of Digital Signals in the Presence of White Gaussian Noise—A Geometric Interpretation and a Study of Three Basic Data Transmission Systems," *IRE Transactions on Communication Systems*, CS-10 (December 1962): 336–372.

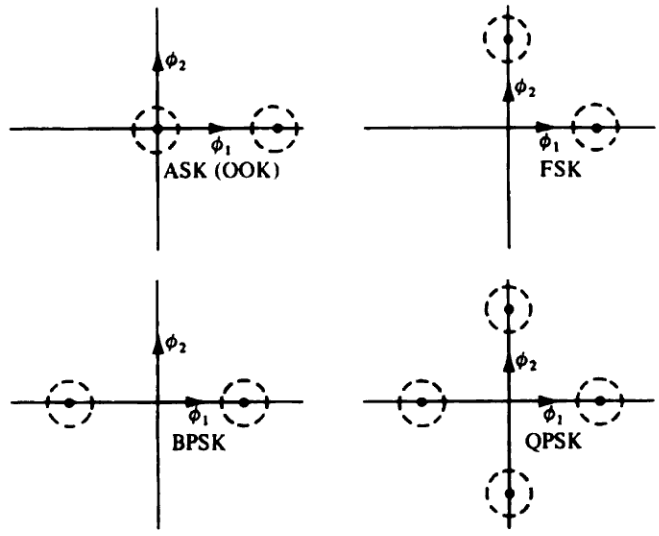


Figure 10.49 Geometric representation of ASK, FSK, BPSK, and QPSK.

In QPSK, the two carriers used (i.e., $\cos \omega_c t$ and $\sin \omega_c t$) are orthogonal and, therefore, the two bit streams in the I and Q signal paths can be demodulated independently. For a given E/η, then, the error probabilities of coherently detected BPSK and QPSK are identical. Moreover, because staggering the bit streams does not alter the orthogonality of the carriers, OQPSK has the same error performance as BPSK and QPSK.

MSK also uses the modulation of two orthogonal carriers to achieve antipodal signaling of two binary data streams. It differs from OQPSK in that it uses weighting functions—$\cos(\pi t/2T_b)$ and $\sin(\pi t/2T_b)$—in the I and Q signal paths, but with no sacrifice in signal energy. Therefore MSK has the same error performance as BPSK, QPSK, and OQPSK when matched filtering is used to recover the data. However, if MSK were coherently detected as a FSK signal (with a bit decision time of T_b second), then the performance of MSK would be poorer than that of BPSK by 3 dB. The reason for this is that the distance between signal states has been reduced $1/\sqrt{2}$ of that for antipodal signaling.

In M-ary PSK, all signal states must lie in a circular pattern centered on the origin. This leaves little room for signal design choices. In contrast, the M signal states in M-ary APK can be placed in any pattern subject to the constraint of maximizing the distance of each signal state from all of the others to keep the error rate as low as possible. This must be achieved within a fixed average power constraint. Other important parameters are the minimum phase angle differences between signal states (a measure of immunity to phase jitter/delay distortion) and the ratio of peak-to-average power (a measure of immunity against nonlinear distortion).

In the orthonormal signal representation, the distance from the origin is chosen to represent the square root of the signal energy. To proceed in a more quantitative manner, we also need to express the noise in terms of the orthonormal basis functions.

Following Eq. (10.108), we write

$$n(t) = \sum_i n_i \phi_i(t) \tag{10.113}$$

where the coefficients n_i are found from [cf. Eq. (10.109)]

$$n_i = \int_0^T n(t)\phi_i(t)\,dt\,. \tag{10.114}$$

A difference here is that $n(t)$ is random. We assume that $n(t)$ is an ergodic random process and that interchanges in order of operations are valid, and proceed to determine the correlation between series coefficients.

$$\begin{aligned} E\{n_i n_j\} &= E\left\{\int_0^T n(t)\phi_i(t)\,dt \int_0^T n(t)\phi_j(t)\,dt\right\} \\ &= E\left\{\int_0^T \int_0^T n(t)n(t_1)\phi_i(t)\phi_j(t_1)\,dt\,dt_1\right\} \\ &= \int_0^T \int_0^T E\{n(t)n(t_1)\}\phi_i(t)\phi_j(t_1)\,dt\,dt_1 \\ &= \int_0^T \int_0^T R_n(t - t_1)\phi_i(t)\phi_j(t_1)\,dt\,dt_1 \end{aligned} \tag{10.115}$$

In the particular case of interest in which the noise $n(t)$ has zero mean and a white power spectral density $S_n(\omega) = \eta/2$, we have $R_n(t - t_1) = (\eta/2)\delta(t - t_1)$; then Eq. (10.115) becomes

$$\begin{aligned} E\{n_i n_j\} &= \int_0^T \int_0^T (\eta/2)\delta(t - t_1)\phi_i(t)\phi_j(t_1)\,dt\,dt_1 \\ &= \frac{\eta}{2}\int_0^T \phi_i(t)\phi_j(t)\,dt \\ &= \begin{Bmatrix} \eta/2 & i = j \\ 0 & i \neq j \end{Bmatrix}. \end{aligned} \tag{10.116}$$

Thus the variance of the noise under these conditions is equal to $\eta/2$ in each coordinate.

EXAMPLE 10.13.1

Determine the net probability of error for BPSK (PRK) in additive white gaussian noise using a geometric approach and assuming equiprobable 1's and 0's.

Solution Using the diagram in Fig. 10.49 and observing that the signal states are at $\pm\sqrt{E_s}$, the decision boundary is a vertical line through the origin. Using symmetry and noting that the variance of the noise is given by $\sigma^2 = \eta/2$, the probability of error is

$$P_\epsilon = \frac{1}{\sqrt{\pi\eta}} \int_0^\infty e^{-(x+\sqrt{E_s})^2/\eta}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2E_s/\eta}}^\infty e^{-z^2/2}\, dz = Q(\sqrt{2E_s/\eta}).$$

This agrees with Eq. (10.24). Moreover, this approach gives some intuitive insights into waveform design and detection considerations.

10.14 ★ OPTIMUM DETECTION ALGORITHMS

In previous sections we have discussed the calculation of the probability of error for a given binary system. Now we wish to investigate the synthesis of an optimum signal processor in the receiver to yield a minimum probability of error. Again we shall restrict our attention to the binary case even though these methods can be extended to more general cases.

To simplify the problem, we return to a consideration of on-off keying (OOK) in which the presence of a pulse signifies a one and the absence of a pulse signifies a zero. Thus either a one or a zero exists over $(0, T)$ and the receiver must make a decision as to which one. We begin by assuming that there is only one sample available on which to base this decision.

Let us label the two possible choices as H_0 and H_1; when a 1 is sent we call it H_1, and when a 0 is sent we call it H_0. The binary transmission is in the presence of additive noise. For simplicity, we assume that the additive noise $n(t)$ is gaussian-distributed with zero mean. The two observed states over $(0, T)$ are [cf. Eqs. (9.25) and (9.26)]

$$y(t) = \begin{cases} s(t) + n(t) & \text{signal present,} \\ n(t) & \text{signal absent.} \end{cases} \tag{10.117}$$

At the end of the interval $(0, T)$ the receiver has the four following choices:

1. H_0 true, choose H_0;
2. H_0 true, choose H_1;
3. H_1 true, choose H_1;
4. H_1 true, choose H_0.

The first and third alternatives correspond to correct choices. In addition, two types of error may occur (cf. Fig. 9.7). In the second alternative the noise may be mistaken

for signal when the signal is not present (P_{ϵ_0}) and in the fourth alternative the signal may be missed when it is present (P_{ϵ_1}). The overall (net) probability of error is then [cf. Eq. (9.33)]

$$P_\epsilon = P_0 P_{\epsilon_0} + P_1 P_{\epsilon_1}, \tag{10.118}$$

where P_1 is the probability of occurrence of a 1 (or signal in this case) and $P_0 = 1 - P_1$.

Both P_{ϵ_0} and P_{ϵ_1} are conditional probabilities because they depend on the signal being present or absent. To derive them, we write

$$p(y|0) = p_n(y). \tag{10.119}$$

We assume that the noise is gaussian with zero mean; thus,

$$p_n(n) = \frac{1}{\sqrt{2\pi}\,\sigma_n} e^{-n^2/(2\sigma_n^2)}. \tag{10.120}$$

Using Eq. (10.120) in Eq. (10.119), we have

$$p(y|0) = \frac{1}{\sqrt{2\pi}\,\sigma_n} e^{-y^2/(2\sigma_n^2)}. \tag{10.121}$$

Let us assume that the signal is a constant, A, over $(0, T)$ so that when the signal is present, $y(t) = A + n(t)$. This simply shifts the mean value of the noise from 0 to A (see Fig. 10.50) and

$$\begin{aligned} p(y|s) &= p_n(y - A) \\ &= \frac{1}{\sqrt{2\pi}\,\sigma_n} e^{-(y-A)^2/(2\sigma_n^2)}. \end{aligned} \tag{10.122}$$

From Fig. 10.50 we can write

$$P_{\epsilon_1} = \int_{H_0} p(y|s)\,dy, \tag{10.123}$$

$$P_{\epsilon_0} = \int_{H_1} p(y|0)\,dy, \tag{10.124}$$

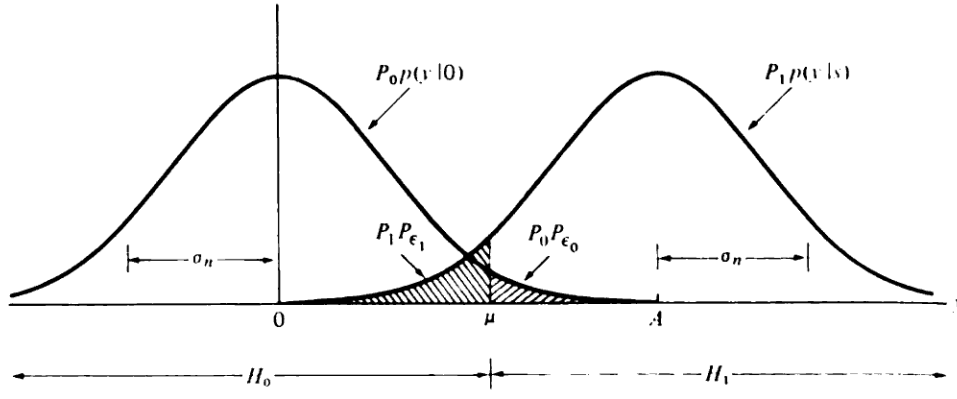


Figure 10.50 Decision regions for the binary case.

so that Eq. (10.118) becomes

$$P_\epsilon = P_1 \int_{H_0} p(y|s)\,dy + P_0 \int_{H_1} p(y|0)\,dy. \tag{10.125}$$

This is the equation we wish to minimize by choosing H_0, H_1 correctly.

We wish to minimize the probability of error by determining an algorithm (i.e., a rule) for deciding if a given observable, $y(t)$, should be interpreted as a 1 or a 0. Suppose we choose to develop a rule for H_1. We have

$$\int_{H_0+H_1} p(y|s)\,dy = 1, \tag{10.126}$$

so that Eq. (10.125) can be rewritten as

$$\begin{aligned} P_\epsilon &= P_1\left[1 - \int_{H_1} p(y|s)\,dy\right] + P_0 \int_{H_1} p(y|0)\,dy \\ &= P_1 + \int_{H_1} [P_0 p(y|0) - P_1 p(y|s)]\,dy. \end{aligned} \tag{10.127}$$

We wish to make P_ϵ as small as possible. Because P_0, P_1 and both probability density functions in Eq. (10.127) are nonnegative, the best we can do is to make

$$P_1 p(y|s) > P_0 p(y|0). \tag{10.128}$$

Thus those values of y satisfying the inequality in Eq. (10.128) should be assigned to the decision H_1. By following the same reasoning, we conclude that those values of y which reverse the inequality should be assigned to H_0.

Equation (10.128) can be rewritten as

$$\ell(y) = \frac{p(y|s)}{p(y|0)} > \frac{P_0}{P_1}, \tag{10.129}$$

where $\ell(y)$ is called the *likelihood ratio*. The signal processing algorithm for a minimum probability of error can now be specified as

$$\ell(y) = \frac{p(y|s)}{p(y|0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}. \tag{10.130}$$

We have assumed in this algorithm that the two types of errors are equally costly to system performance. This is normally true in binary communication systems. If desired, allowance for errors of unequal cost can easily be made in the algorithm.

Because probability density functions are nonnegative and the logarithm function is monotonic, it is often convenient to take the natural logarithm of both sides of Eq. (10.130) to give

$$\ln \ell(y) \underset{H_0}{\overset{H_1}{\gtrless}} \ln \frac{P_0}{P_1}. \tag{10.131}$$

The signal-processing algorithm specified in Eqs. (10.130) and (10.131) is called the *likelihood ratio test* (LRT).

EXAMPLE 10.14.1

(a) Determine the LRT for on-off keying in the presence of additive gaussian noise with zero mean. (b) If in addition the noise is white, specify the LRT for the output of a matched filter.

Solution a) Assume that the signal has an amplitude A when a 1 is sent and zero amplitude when a 0 is sent. The conditional probabilities are given by Eqs. (10.121) and (10.122). Using Eq. (10.131), the LRT is

$$\frac{(2Ay - A^2)}{(2\sigma_n^2)} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\left(\frac{P_0}{P_1}\right),$$

or

$$y \underset{H_0}{\overset{H_1}{\gtrless}} \frac{A}{2}\left[1 + \frac{2\sigma_n^2}{A^2}\ln\frac{P_0}{P_1}\right].$$

Thus, when $P_0 = P_1 = 0.5$, the signal processor sets the decision threshold at $A/2$, as we could expect [cf. Eq. (9.35)]. If P_1 differs from 0.5, the optimum decision threshold is shifted. The relative amount it is shifted is dependent on the natural logarithm of $(P_0/P_1)^2$ and inversely proportional to the peak signal-to-noise ratio.

b) The matched filter is a linear system so that the output noise is also gaussian-distributed with zero mean value and [cf. Eq. (9.48)] $\sigma_n^2 = E\eta/2$. The output amplitude with signal present is E so that the LRT of part (a) becomes

$$y \underset{H_0}{\overset{H_1}{\gtrless}} \frac{E}{2} + \frac{\eta}{2}\ln\frac{P_0}{P_1},$$

where E is the energy of the signal $s(t)$ over $(0, T)$ and $\eta/2$ is the (two-sided) power spectral density of the noise.

DRILL PROBLEM 10.14.1

Repeat Example 10.14.1 for the case of a polar binary signal with transmitted signals $-A/2$, $A/2$.

ANSWER: (a) $y \gtrless (\sigma_n^2/A)\ln(P_0/P_1)$; (b) $y \gtrless (\eta/2)\ln(P_0/P_1)$.

The likelihood ratio test (LRT) was developed above for the case of one sample upon which to base a decision. However, the concept of the LRT is much more general than this and we now look at the more general case in which more than one sample is available over $(0, T)$.

The sampling theorem states that we can obtain $2BT$ independent samples of a deterministic signal band-limited to B Hz over $(0, T)$. For band-limited white noise, the

noise samples are uncorrelated if they are spaced $1/(2B)$ seconds apart (see Drill Problem 8.13.1). For gaussian noise it follows that these noise samples are also independent. Therefore we shall assume that $N = 2BT$ independent samples of the observable, $y(t)$, are available.

Because the $N = 2BT$ samples are statistically independent, we can multiply their probability density functions to obtain the net probability density function. Thus the results obtained for the one-sample case are easily generalized to give a LRT for N samples:

$$\ell(y_1, y_2, \ldots, y_N) = \frac{\prod_{i=1}^{N} p(y_i|s)}{\prod_{i=1}^{N} p(y_i|0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1} \tag{10.132}$$

where Π indicates multiplication. Taking the logarithm of both sides of Eq. (10.132), we get

$$\ln \ell(y_1, y_2, \ldots, y_N) \underset{H_0}{\overset{H_1}{\gtrless}} \ln \frac{P_0}{P_1}. \tag{10.133}$$

Note that Eqs. (10.132) and (10.133) are generalized versions of Eqs. (10.130) and (10.131).

EXAMPLE 10.14.2

Assume that under H_1 the transmitter sends a constant level A and under H_0 it sends a zero level over $(0, T)$. Additive gaussian noise with zero mean and variance σ^2 is present and the observable signal-plus-noise is ideally band-limited to B Hz. Determine the LRT and the optimum signal processor.

Solution If we follow Eqs. (10.121) and (10.122) for the multiple-sample case, Eq. (10.132) is

$$\frac{\prod_{i=1}^{N} (1/\sqrt{2\pi}\,\sigma) e^{-(y_i - A)^2/(2\sigma^2)}}{\prod_{i=1}^{N} (1/\sqrt{2\pi}\,\sigma) e^{-y_i^2/(2\sigma^2)}} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}.$$

Canceling common terms, using summations in the arguments of the exponentials for the products, and taking the natural logarithm, we have

$$\frac{1}{N} \sum_{i=1}^{N} y_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{A}{2} + \frac{2\sigma^2}{NA} \ln \frac{P_0}{P_1},$$

where $N = 2BT$. Therefore the optimum signal processor performs an operation known as the *sample mean:* $(1/N)\sum_{i=1}^{N} y_i$ on the observable signal-plus-noise in $(0, T)$ and compares it to a specified level. This result agrees with that obtained in Example 10.14.1(a) for the case where $N = 1$.

DRILL PROBLEM 10.14.2

Repeat Example 10.14.2 for the case of a polar binary signal with transmitted signals $-A/2$, $A/2$.

ANSWER: $$\frac{1}{N}\sum_{i=1}^{N} y_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma^2}{NA}\ln\frac{P_0}{P_1}.$$

The likelihood ratio test (LRT) is a very important one because it specifies the optimum signal processor algorithm in terms of the conditional statistics of the observable quantity. As illustrated in Example 10.14.2, it can be used to specify the required receiver operations (which are not required to be linear) to achieve this optimum. The criterion for optimality is a minimum probability of error for equal costs in making errors.

The next logical step is to try to find an optimum processor to estimate a parameter in the observed signal after its detection. This problem is important not only for digital communication systems but also for analog systems. For example, we would like to know what is the optimum signal processor to recover the variation of phase (or frequency) in a received signal. It turns out that the likelihood ratio plays a key role in finding the optimum processor for many of these problems. But we shall have to leave these (and other) very interesting topics for future work in communication systems.

10.15 SUMMARY

Digital modulation methods are used for transmission of PCM signals through bandpass channels. Binary modulation methods used are amplitude-shift keying (ASK), frequency-shift keying (FSK), and phase-shift keying (PSK). In these methods, the amplitude, frequency, or phase of a sinusoid is switched in response to the PCM input.

Using matched-filter detection in additive white gaussian noise, PSK requires up to 3 dB less average signal power than ASK or FSK for a given error probability. Coherent detection is required for PSK, whereas either coherent or noncoherent detection can be used for ASK and FSK. Differential PSK (DPSK) can be used to gain most of the advantages of PSK and yet avoid the requirement for coherent detection. For a given error performance, FSK requires about 3 dB less peak power than ASK, even though their average power requirements are similar. Popular choices for communication systems are PSK, DPSK, and noncoherent FSK.

Quadrature PSK (QPSK) uses quadrature multiplexing principles to offer twice the bit rate capability as that of binary PSK (BPSK) within the same bandwidth. Delay of either the in-phase (I) or quadrature (Q) signal by one bit period results in offset-keyed QPSK (OQPSK). An advantage of OQPSK is that it has a more constant envelope than QPSK after bandpass filtering. Both methods offer bandwidth efficiencies of 2 bps/Hz.

Spread spectrum systems purposely spread modulated spectra, independent of the data modulation. Use of spread spectrum allows sharing a given band for multiple transmissions with only minimum interference between users. This type of multiplexing is called code division multiple access (CDMA). Spread spectrum has advantages in secure communications, selective calling systems, and protection against narrowband jamming and selective fading.

Direct-sequence (DS) spread spectrum uses suppressed-carrier modulation of a data-modulated carrier with a PN code at a rate called the chip rate. Bandwidth spreading is dependent on the ratio of the chip rate to the bit rate. An extra level of synchronization is necessary in reception for correct PN code sequence and timing.

Minimum-shift keying (MSK) is an example of coherent FSK in which the phase in the modulated waveform is continuous and the frequency shift is such that there is a one-half cycle difference in one bit interval. MSK has signaling characteristics similar to those of OQPSK. A difference is that the power spectral density of MSK has a wider main lobe but decreases more rapidly farther away from the carrier frequency.

In M-ary PSK, the modulator generates constant-amplitude signals with phase spacings of $2\pi(M - 1)/M$ to transmit M-ary information. After bandpass filtering, such signals do not necessarily have constant envelope. For a fixed noise power, detection at low error rates requires substantial increases in signal power as M is increased. The 8-ary PSK systems are popular choices, but transmitted power requirements for $M > 8$ become excessive for many applications.

Amplitude-phase keying (APK) systems offer less required power for given M than does M-ary PSK for a given probability of error. The generated signal envelope is not constant, and a linear channel is required. Both APK and M-ary PSK systems offer bandwidth efficiencies of $\log_2 M$ bps/Hz. Both are sensitive to channel distortions, and equalization is required for good transmission characteristics. In one particular type of APK signaling, called quadrature partial response (QPR), partial-response signaling is used in both the I and Q signals with quadrature amplitude modulation (QAM).

In contrast to the foregoing methods, M-ary orthogonal FSK offers savings in transmitted power at the expense of increased bandwidth for a given error rate. M-ary orthogonal FSK can be generalized to include other orthogonal signal sets. The performance of M-ary orthogonal signaling approaches that of the Hartley–Shannon result for channel capacity when $M \to \infty$.

Frequency hopping (FH) spread spectrum systems shift the carrier frequency of a modulated waveform in response to a code known to the receiver. Frequencies selected and the rate at which they are selected (called the hopping rate) can be controlled independently to attain wide bandwidths. FH systems are termed fast-hop or slow-hop, depending on the frequency hopping rate relative to the bit rate.

Vector-space concepts may be used to give geometric interpretations to digital communication problems. The concept of a signal-state constellation is useful in M-ary signaling. Conditional probabilities may be used to form a likelihood ratio test. This test yields a signal processing algorithm that results in a minimum probability of error in detection of signals in noise.

Selected References for Further Reading

1. S. Haykin. *Communication Systems*. Second ed. New York: John Wiley & Sons, 1983.
   Chapter 7 covers digital modulation methods in the context of matched-filter detection and maximum likelihood estimation.
2. R. E. Ziemer and W. H. Tranter. *Principles of Communications*. Second ed. Boston: Houghton Mifflin, 1985.
   Chapters 7 and 8 offer an expanded treatment of many of the topics in this chapter with an emphasis on matched-filter detection and maximum likelihood estimation.
3. K. Feher. *Digital Communications: Microwave Applications*. Englewood Cliffs, N.J.: Prentice Hall, 1981.
   A good discussion of both theory and current practice of M-ary PSK and APK methods are covered in Chapters 3, and 6–8, with an emphasis on terrestrial microwave systems.
4. R. W. Lucky, J. Salz, and E. J. Weldon, Jr. *Principles of Data Communication*. New York: McGraw-Hill, 1968.
   Chapters 7–9 cover the topics of ASK, FSK, and PSK in some detail with an emphasis on telephone channels.
5. J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: John Wiley & Sons, 1965.
   A well-known text that covers optimum receiver principles and geometric interpretation of signals in Chapter 4.
6. J. J. Spilker, Jr. *Digital Communications by Satellite*. Englewood Cliffs, N.J.: Prentice Hall, 1977.
   A fairly extensive discussion of generation, detection, and effects of distortion in BPSK and QPSK systems in Chapters 11–13, with an emphasis on satellite channels.
7. G. R. Cooper and C. D. McGillem. *Modern Communications and Spread Spectrum*. New York: McGraw-Hill, 1986.
   Generation and detection of spread spectrum and applications of spread spectrum in communications are discussed in Chapters 8–13.
8. R. E. Ziemer and R. L. Peterson. *Digital Communications and Spread Spectrum Systems*. New York: Macmillan, 1985.
   Good discussions of digital modulation, signal-space methods, and introduction to DS and FH spread spectrum systems. Later portions of this text are more specialized and advanced.
9. I. Korn. *Digital Communications*. New York: Van Nostrand Reinhold, 1985.
   An in-depth discussion of digital modulation methods.
10. W. C. Lindsey and M. K. Simon. *Telecommunication Systems Engineering*. Englewood Cliffs, N.J.: Prentice Hall, 1973.
    An extensive treatment of phase-coherent systems.

PROBLEMS

10.1.1 A binary code is transmitted using (OOK) amplitude-shift keying at a rate of 3 Mbps. The ratio of carrier power (when keyed on) to thermal noise temperature at the receiver input is 10 fW/K. Assuming equiprobable ones and zeros, calculate the expected minimum probability of error if the code representation used is
a) NRZ(L); b) RZ.

10.1.2 a) Derive an expression for the probability of error of ASK if the correlation receiver shown in Fig. 10.2(b) has a reference signal $A \sin(\omega_c t + \theta)$, where θ is a phase error.

b) Using your result for part (a), determine the maximum phase error to stay within the probability of error for that of a noncoherent receiver if $E/(2\eta) = 18$.

10.1.3 A certain proposed binary ASK system uses the following signals for transmission in the presence of additive white gaussian noise:

$$\phi(t) = \begin{Bmatrix} A \sin \omega_c t & \text{for a } 1 \\ \gamma A \sin \omega_c t & \text{for a } 0 \end{Bmatrix} \quad \text{over } (0, T),$$

where $0 \le \gamma < 1$.

a) Derive an expression for the minimum probability of error P_ϵ for this system, assuming that 1's and 0's are equiprobable.
b) Plot P_ϵ versus γ for $E/(2\eta) = 18$. Use a five-cycle logarithmic scale for P_ϵ, and Eq. (I.5) in Appendix I as an approximation for Q.
c) Let $P_{\epsilon 0}$ be the value of P_ϵ for $\gamma = 0$. Determine the largest value of γ for which P_ϵ is no more than 10% greater than $P_{\epsilon 0}$.

10.1.4 A binary NRZ(L) data waveform is transmitted using (OOK) amplitude-shift keying at a rate of 1 Mbps. Assume that 1's and 0's are equiprobable. Determine and sketch the power spectral density about the carrier frequency f_c for each of the following:

a) An alternating series of 1's and 0's;
b) A random series of 1's and 0's (cf. Example 8.13.2).

10.2.1 A given binary FSK system operates in the presence of additive white gaussian noise with (two-sided) power spectral density $\eta/2 = 1$ fW/Hz. Assume equiprobable ones and zeros, and $T_b = 0.8$ msec. The signaling waveforms are

$$\phi_1(t) = A \cos (1.2 \times 10^6 t)$$

$$\phi_2(t) = A \cos (1.0 \times 10^6 t)$$

and $A = 10\ \mu$V across 1 ohm.

a) Find the probability of a bit error, using coherent detection.
b) Find the probability of a bit error, using noncoherent detection.
c) Estimate the bandwidth required.

10.2.2 A binary FSK signal with discontinuous phase is generated by switching between two oscillators. The resulting output can be written as [cf. Eq. (10.9)]:

$$\phi(t) = Ap_1(t) \cos (\omega_1 t + \theta_1) + Ap_2(t) \cos (\omega_2 t + \theta_2),$$

where $p_1(t)$ is a binary switching waveform that is either 0 or 1 over $(0, T)$ and $p_2(t) = 1 - p_1(t)$.

a) Find the power spectrum for a periodic alternating sequence of 1's and 0's.
b) Plot the power spectrum determined in part (a) for $\Delta fT = 1$ and for $\Delta fT = 2$.

10.2.3 Determine $R_\phi(\tau)$ for $\phi(t)$ in Problem 10.2.2, and then find $S_\phi(\omega)$ for equiprobable 1's and 0's. Assume that θ_1, θ_2 are uniformly distributed over $(0, 2\pi)$, that the two oscillators operate independently, and that $\omega_c = (\omega_1 + \omega_2)/2$ is much greater than $\Delta\omega = (\omega_2 - \omega_1)/2$. [*Hint:* Separate out the average value of $p_1(t)$, $p_2(t)$.]

10.2.4 Let $\phi_1(t), \phi_2(t)$ be two binary FSK signals over $(0, T)$ with a carrier frequency ω_c and a frequency difference of $2\Delta\omega$, where $\Delta\omega \ll \omega_c$.

a) Show that $\phi_1(t), \phi_2(t)$ are orthogonal over $(0, T)$ if $2\Delta\omega T = n\pi$.
b) Plot $\log_{10} P_\epsilon$ versus $2\Delta\omega T$ for $E/\eta = 16$ in Eq. (10.15).

10.2.5 FSK transmissions are permitted in some high-frequency amateur radio bands using either $2\Delta f = 170$ Hz ("narrowband") or $2\Delta f = 850$ Hz ("wideband") and with $2\Delta fT = 3.75$, 18.75, respectively. Suppose that the rms (noise-free) signal strength at the receiver input is 8 μV across 50 Ω and that these transmissions are received in the presence of additive white gaussian noise with (two-sided) power spectral density $\eta/2 = 1$ fW/Hz.

a) Compute the minimum probability of error, using coherent demodulation; assume equiprobable 1's and 0's.
b) Repeat part (a), using noncoherent demodulation.
c) Why do you suppose the values of $2\Delta fT$ are chosen to be such fractional values?
d) What are the relative advantages of each choice of $2\Delta f$?

√ **10.2.6** Use a $N = 512$ FFT to compute the power spectrum of the binary FSK waveform $\phi(t) = \sin[2\pi(f_c \pm \Delta f)t]$ over (0, 2). Let $f_c = 64$, $(f_c + \Delta f)$ correspond to 1, $(f_c - \Delta f)$ correspond to 0. For a simulation pattern, subdivide the interval (0, 2) into four equal durations and use the following sequences: 1010, 1100, 0001, 0111. Use equal weighting to average the four power spectra. Plot the averaged power spectrum over $(7N/32, 9N/32)$ for $\Delta f = 0.25, 0.5, 1.0$.

10.3.1 This problem examines the sensitivity of the equiprobable assumption on the net probability of error. Suppose that a coherent BPSK (PRK) system operates with $E/\eta = 10$ and probabilities of 0's and 1's equal to P_0, P_1, respectively.

a) Find the net probability of error, P_ϵ, for $P_1 = 0.4, 0.5, 0.6$, assuming that the detection threshold is set at zero in each case.
b) Find the additional signal power required, in decibels, to maintain P_ϵ for $P_1 = 0.5$ when $P_1 = 0.4, 0.6$.
c) What is P_ϵ for the optimum threshold settings?

10.3.2 A binary code is transmitted using BPSK (PRK) at a rate of 1.5 Mbps in the presence of additive white gaussian noise. The ratio of the carrier power to thermal noise temperature at the receiver input is 1 fW/K. State any assumptions you make in your solution.

a) Calculate the minimum probability of error.
b) Calculate the S/N ratio (in decibels) at the receiver input.
c) Calculate the change in input S/N (in decibels) for a one-order-of-magnitude increase in the probability of error.

10.3.3 In remote sensing experiments using the Nimbus F satellite, signals are sent once every minute from each transmitter for a duration of 1 sec [see, e.g., the reference given for the design problem in Chapter 8]. This 1-sec transmission consists of 320 msec of unmodulated carrier for location and acquisition, followed by 640 msec of PSK-modulated Manchester-encoded data at 100 bps and with $\Delta\theta = \pi/3$. For simplicity here, assume that the data waveform results in a periodic, alternating $+\Delta\theta$, $-\Delta\theta$ phase shift.

a) Find the Fourier series of this PSK waveform in terms of $\Delta\theta$.
b) Does your result in part (a) agree with that of Drill Problem 6.5.1?
c) Does your result in part (a) agree with Fig. 10.13?
d) Plot the power spectral density of this PSK waveform for $\Delta\theta = \pi/3$.

10.3.4 A given BPSK (PRK) system transmits data in the presence of additive white gaussian noise with a probability of error of $P_e = 10^{-5}$. This PSK system is tested for $\Delta\theta < \pi/2$.

a) Plot P_e as a function of $\Delta\theta$ for $0 \leq \Delta\theta < \pi/2$.
b) The satellite system described in Problem 10.3.3 must meet a design goal of $P_e = 10^{-4}$ while allocating a fraction of the transmitted power to the carrier for PLL tracking. What fraction of the transmitted power can be allocated to the carrier under these conditions?

10.3.5 Consider the Costas loop shown in Fig. 10.16. Let the input be $p(t) \cos(\omega_c t + \theta_i)$. Write a signal for each path in the block diagram and show that the Costas loop keeps the VCO centered at the suppressed-carrier frequency, insensitive to $p(t)$, and thus provides an acceptable reference signal for the synchronous detector.

10.3.6 We wish to transmit the bit stream 1011000101110010 using DPSK.

a) Determine the encoded sequence. Assume a 1 as the extra starting (preamble) digit.
b) Use the receiver in Fig. 10.17 to demodulate the DPSK-encoded signal of part (a).
c) Introduce an error in the fourth digit (not including the preamble digit) in the DPSK-encoded bit stream in part (a); then repeat part (b). Explain your result.

10.4.1 Rank the binary digital modulation systems covered, assuming that the primary criterion is a) receiver simplicity, or b) lowest peak power requirement for a given error rate (10^{-4}).

10.4.2 Compare the average power requirements, expressed in decibels, for BPSK (PRK), DPSK, coherent (orthogonal) FSK, noncoherent FSK, coherent ASK (OOK), and noncoherent ASK (OOK) for NRZ signaling at 300 bps, assuming $\eta/2 = 10^{-14}$ W/Hz and $P_e = 10^{-5}$.

10.4.3 A given high-frequency transmitter is peak power limited to 1 kW. Path losses to the receiver are estimated to be 135 dB. The (two-sided) power spectral density of the additive white gaussian noise is $\eta/2 = 10^{-18}$ W/Hz. Assume NRZ data signaling at 1 Mbps, and equiprobable ones and zeros.

a) Determine the minimum probability of error using ASK (OOK), coherent (orthogonal) FSK, and BPSK (PRK).
b) Estimate the input S/N ratio (in decibels) for each choice in part (a), using a first-nulls bandwidth for ASK, BPSK, and a $4/T_b$ approximation for FSK.

10.5.1 A direct-sequence BPSK spread spectrum system with average power P uses NRZ data signaling and operates in the presence of continuous-tone jamming with average power P_J. The jammer operates at the center frequency of the BPSK transmission. Assume that the spreading code rate is four times the data rate, and that the period of the spreading code can be assumed to be infinitely long.

a) Sketch the power spectral densities corresponding to Eq. (10.37).
b) Sketch the power spectral densities corresponding to Eq. (10.38).
c) Assume an ideal IF filter with minimum (Nyquist) bandwidth for the NRZ data signaling. Set up the integration for the signal-to-jammer power ratio. (*Optional:* Solve the integration numerically.)

10.5.2 A direct-sequence BPSK spread-spectrum system with average power P uses NRZ data signaling and operates in the presence of continuous-tone jamming with average power, P_J. Assume that the spreading code rate is 100 times the data rate, and that the period of the spreading code can be assumed to be infinitely long. Also assume an ideal IF filter with a bandwidth equal to the first-nulls bandwidth of the BPSK-modulated waveform.

a) Determine the ratio of the jammer power at the output of the IF filter to that at the receiver input.

b) Show that the jammer power is most effective at the center frequency of the BPSK transmission.

10.5.3 Two direct-sequence BPSK spread spectrum systems operate with NRZ data signaling and PN spreading codes. Block diagrams of the two transmitters are shown in Fig. P–10.5.3(a). System 1 transmits data $d_1(t)$ using a carrier frequency ω_1, an average power P_1, and a spreading waveform $c_1(t)$. System 2 transmits data $d_2(t)$ using a carrier frequency ω_2, an average power P_2, and a spreading waveform $c_2(t)$. Assume that $\omega_1 \approx \omega_2$, and that $d_1(t)$, $d_2(t)$, $c_1(t)$, $c_2(t)$ are statistically independent. The data rates are nominally the same for both systems, and the chip rates are nominally the same for both systems. Also assume that the inequality $(1/T_{ch}) \gg (1/T_b)$ holds. A despreading receiver is set to demodulate the signal transmitted by System 1; a simplified block diagram of the receiver is shown in Fig. P–10.5.3(b). Determine the ratio of the desired signal power to the interfering signal power at the output of the IF filter for minimum (Nyquist) IF filter bandwidth. State any assumptions you make in your solution.

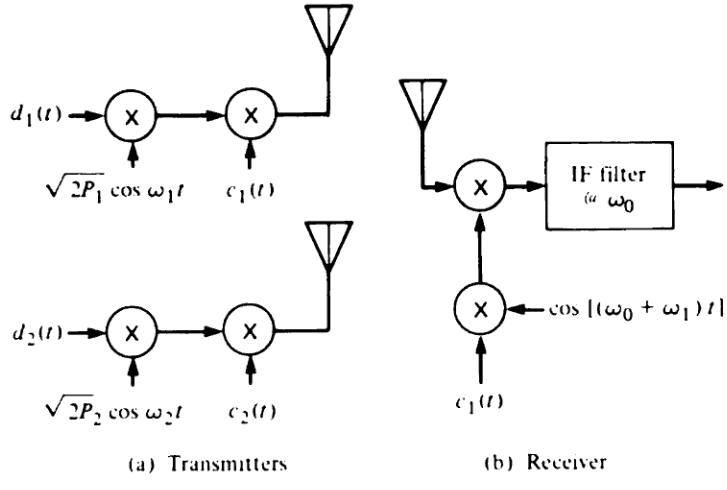


Figure P–10.5.3.

10.6.1 Data streams $d_1(t)$ and $d_2(t)$ are composed of the odd-indexed and even-indexed bits, respectively, of the serial bit stream given in Problem 10.3.6. Using Fig. 10.24(d), sketch the phase of the resulting modulated waveform with respect to carrier (Gray encoded) if the modulation used is

a) QPSK;

b) OQPSK.

10.6.2 Investigate a proposed upgrade in GOES data transmissions to ground stations (see Example 10.3.1) by doubling the data rate and using QPSK. The bandwidth requirements are to remain fixed. Assume that Gray coding would be used. Compute the satellite transmitter power requirement for this proposed change for a) a fixed symbol error rate; b) a fixed BER.

10.6.3 A 10 Mbps QPSK system using NRZ data signaling and Gray coding operates in the presence of additive white gaussian noise. The receiver has a noise bandwidth equal to two-thirds of the first-nulls bandwidth of the modulated signal waveform, and the receiver S/N is 12 dB. Compute the probability of symbol error and the BER of the system.

10.6.4 The block diagram of a fourth-power loop used for QPSK carrier recovery is shown in Fig. P–10.6.4. Analyze the operation of this loop, assuming a noise-free input.

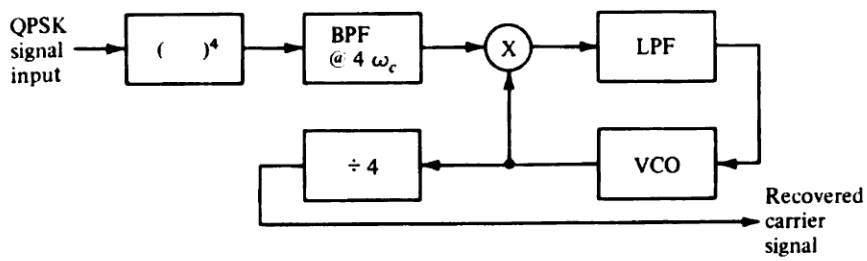


Figure P–10.6.4.

10.6.5 Sketch a complete block diagram of an OQPSK receiver for NRZ binary data, including both carrier recovery and symbol timing recovery subsystems.

10.6.6 Find $v_e(t)$ for the Costas loop in Fig. 10.27, and sketch $v_e(t)$ as a function of the error. Does this sketch have similarities to the error signal in a PLL? Explain.

10.7.1 Repeat Problem 10.6.1 for MSK modulation. Assume that the bit preceding the sequence shown is a zero.

10.7.2 Repeat Problem 10.6.3 for MSK modulation.

10.7.3 a) One measure of equivalent bandwidth is that required for a rectangular power/energy spectral density having the same power/energy as the signal [cf., e.g., Fig. 4.14]. Compare equivalent bandwidths for PSK, QPSK, and MSK on this basis.

b) Another way to measure bandwidth analytically is in terms of second moments [cf. Eq. (3.81)]. Compute the equivalent bandwidth of MSK on this basis, and compare it with the first-nulls bandwidth.

c) Why can't we use the approach in part (b) for PSK and QPSK?

√ **10.7.4** Compute the percentage power containment for QPSK and MSK for bandwidths $W = 1/T_b, \ldots, 8/T_b$ and present your results in a table.

10.7.5 The block diagram of a carrier recovery system for MSK is shown in Fig. P–10.7.5. Assuming a noise-free signal at the input, show that, except for a phase ambiguity of π radians, this system supplies the desired coherent reference signals for detection of MSK.

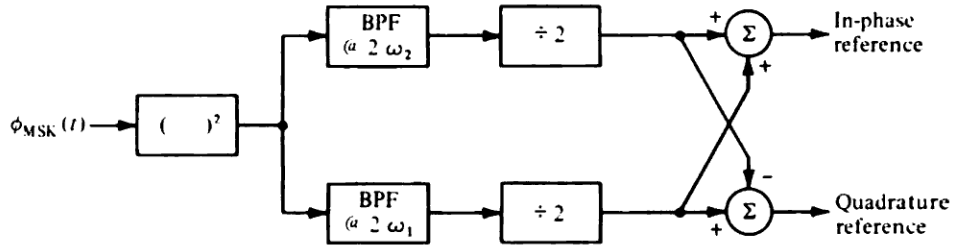


Figure P–10.7.5.

10.7.6 A certain scrambling method in a serial MSK data transmission system is shown in Fig. P–10.7.6. Consider the bit stream 1111000110100101. Take the odd-indexed bits as $d_1(t)$ and the even-indexed bits as $d_2(t)$. The three-stage PN generator shown in Fig. 7.51 is used to generate the scrambling code. Both the data sequences and the code sequence are formatted using NRZ, with zeros represented as -1 and ones as $+1$. The PN generator is triggered from the bit stream so that it shifts to its next state as the waveform for $d_2(t)$ is applied to the input port of the scrambling mixer.

a) Estimate the -3-dB bandwidth of the system. State your assumptions.

b) Sketch an excess-phase trellis diagram for the output, $\phi(t)$, of the system. Assume an initial modulator phase of zero degrees and an initial (preamble) data state of 0.

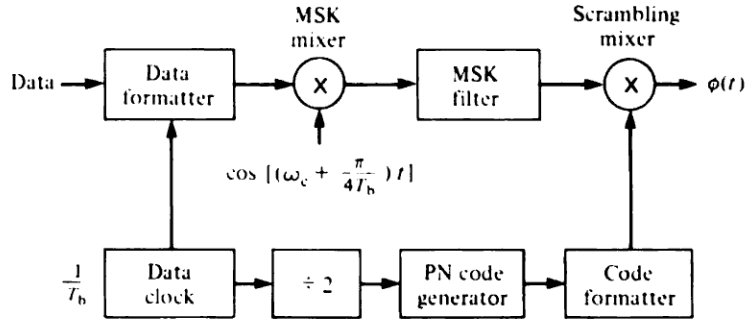


Figure P–10.7.6.

★ **10.8.1** A proposed receiving station for a space experiment requiring a data rate of 0.3 Mbps has a bandwidth of 1 MHz and a receiving antenna gain of 48 dB. The spacecraft antenna gain is 6 dB, and path losses (in decibels) are $\alpha = 100 + 20 \log_{10} d$, where d is in miles. Assuming that the average transmitter power is 10 W and the power spectral density of the noise is $\eta/2 = 2 \times 10^{-21}$ W/Hz, estimate the maximum distance over which communication can be maintained with a BER of 10^{-5} and using

a) M-ary othogonal FSK;

b) QPSK.

★ **10.8.2** Plot the average power required, for fixed noise spectral density and bit transmission rate, versus bandwidth for M-ary FSK, $M > 2$, $P_e = 10^{-4}$ and $P_e = 10^{-5}$. Use a logarithmic scale for bandwidth, linear for power. This graph presents tradeoffs of S/N, bandwidth and probability of error in M-ary FSK systems.

★ **10.8.3** Plot f_b/B versus E_b/η with data taken from Table 10.2 for $P_e = 10^{-4}$ and $P_e = 10^{-5}$; plot in log-log coordinates. Also plot the Hartley–Shannon channel capacity [Eq. (9.9)] bound in these coordinates. Because $N = \eta B$, $S = E_b f_b$, and the limiting value of f_b is C, the limiting case for f_b/B is

$$\frac{C}{B} = \log_2\left(1 + \frac{EC}{\eta B}\right) \quad \text{or} \quad \frac{E}{\eta} = \frac{2^{C/B} - 1}{C/B}.$$

This graph contrasts the performance of M-ary FSK with that of an ideal system.

★ **10.8.4** Using a change in the order of integration and a change of variables, show that Eq. (10.11) follows from Eqs. (10.94) and (10.95) for the case $M = 2$.

10.9.1 A FH spread spectrum system uses binary (NRZ) FSK data modulation at 5 kbps and three frequency hops per bit interval. Frequency hopping is governed by a maximal-length PN sequence generated by a 11-stage feedback shift register. The frequency shift of the FSK modulator, Δf, is set for $\Delta f T_b = 1$. Frequency differences between hops are set at $4\,\Delta f_h$, where $\Delta f_h T_h = 1$. Determine the net bandwidth required by the system. State your assumptions.

10.9.2 A given slow-hop FH/BPSK spread spectrum system uses BPSK data modulation at a rate of 200 kbps and four FH frequencies with frequency spacings of 200 kHz. Assume that the phase of the frequency synthesizer is randomly distributed for each successive hop. The frequency hop rate is 5 kHz.

a) What bandwidth is required by the system?
b) Sketch the power spectral density of the transmitted signal waveform.

√ **10.9.3** Compute and graph the power spectrum of a FH spread spectrum system using binary FSK for the data modulation and four hopping frequencies. Use a 128-point FFT, $T = 1$, hop frequencies given by $f_h T = 26, 30, 34, 38$. Use square-wave data modulation at maximum rate and set $\Delta f T = 1$.

10.10.1 Suppose an upgrade for the satellite data transmission system in Example 10.3.1 is under consideration to handle possible increases in data rates. It is decided that the transmitter power can be increased, but cannot exceed 200 W. Otherwise, no changes other than the modulation are to be made to the system, including bandwidth. Determine if you can meet the given BER within the transmitter power limitation; if not, determine the best BER you can achieve. Assume Gray coding. For each case, list the data rate, transmitter power, and expected BER. Cases to be considered are

a) $M = 4$ PSK;
b) $M = 8$ PSK.

10.10.2 A block diagram of a possible $M = 8$ PSK modulator is shown in Fig. P–10.10.2. The serial-to-parallel converter changes the input bit stream at rate f_b into three polar binary waveforms A, B, C, each at a rate $f_b/3$. The C waveform controls two amplifiers, each with two preset gains (the amplifiers are assumed to be identical). When $C = 1$, the gain of the upper amplifier is large and the gain of the lower amplifier is small; when $C = 0$, the converse is true.

a) Sketch a phasor diagram for this modulator when the amplifier gains are in the ratio 2:1.
b) Determine the amplifier gain settings ratio for 8 equal modulation angles.

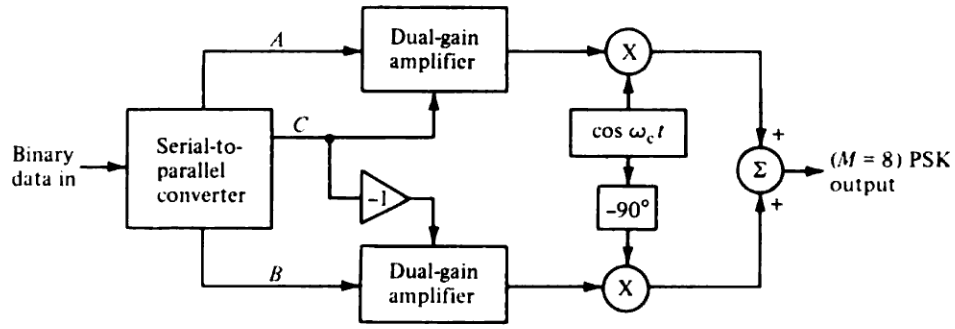


Figure P–10.10.2.

10.10.3 A telephone channel allows signal transmission in the frequency range 600–3000 Hz. Assuming that the carrier frequency is 1800 Hz, show that a) 2400 bps can be sent using four-phase PSK with raised-cosine shaping; and that b) 4800 bps can be sent using eight-phase PSK with 50% sinusoidal roll-off. c) What is the −6-dB bandwidth for each case?

10.10.4 A data transmission system is required to handle a data rate of 64 kbps using M-ary PSK with coherent detection and a BER of 10^{-4}. Assume additive white gaussian noise, full-cosine roll-off ($\alpha = 1$) spectral shaping (for ISI), a −6-dB bandwidth of 16 kHz, and Gray coding.

a) Determine the minimum E_s/η ratio required to meet the system requirements.
b) Repeat part (a) if the data rate requirements can be relaxed somewhat (but you decide and specify where the next "break" in system performance occurs).
c) Convert your answers to S/N requirements for the system.

10.11.1 Explain why there is a significant degradation in performance in an APK system when the modulated signal is passed through a nonlinear (saturation-type) amplifier. Consider both the case of a band-limited APK signal and one that is not band-limited.

10.11.2 Binary data at 9600 bps is transmitted over a conditioned telephone line using a bandwidth of 2400 Hz. Calculate the average S/N necessary to maintain $P_e = 10^{-5}$ in using (a) 16-ary PSK, and (b) 16-ary APK (see Fig. 10.45). Assume that the noise is additive white gaussian noise.

10.11.3 Let the distance from the origin in a signal-state constellation represent the square root of signal energy. Assume that each separation in ϕ_1 and ϕ_2 in the 16-APK signal-state constellations shown in Fig. 10.42 is equal to 2Δ. Determine Δ in terms of the average symbol energy E_s for each case, assuming that all signal states are equiprobable.

10.11.4 The attenuator following QPSK modulator 1 in Fig. 10.44 is set at $K = 1/2$ in that diagram. Sketch the resulting signal-state constellations if the gain of this attenuator were set at

a) $K = 1/4$;
b) $K = 3/4$.

★ **10.13.1** It is given that the probability of error for the binary PSK system in Fig. 10.49 is 10^{-4} with $P_0 = P_1 = 0.5$. Determine the corresponding probability of error geometrically for ASK(OOK) and FSK using the same average transmitted power.

★ **10.13.2** A signal-space constellation for QPSK is shown in Fig. P–10.13.2. Using a signal-space approach, derive the minimum probability of error of this system when operating in the presence of additive white gaussian noise with power spectral density $S_n(\omega) = \eta/2$. Assume that all four signal states are equiprobable; note that the energy has been normalized to give an average energy per symbol of E_s. [*Hint:* First calculate the probability of a correct decision.]

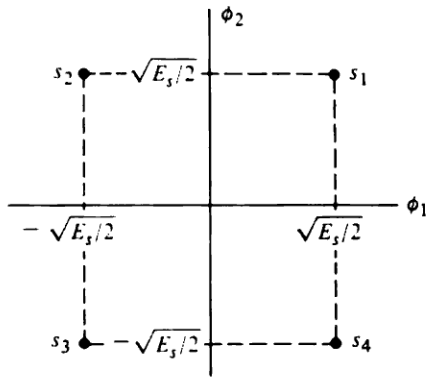


Figure P–10.13.2.

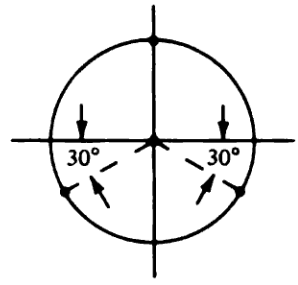
★ **10.13.3** Calculate P_ϵ for the 16-APK signal-state constellation shown in Fig. 10.42(a). Express your answer in terms of E_s, the average energy per symbol. [*Hint:* Classify detection regions into several groups; also see the suggestion in Problem 10.13.2.]

★ **10.13.4** (Union Bound Approximation) In multistate systems, it can become tedious and somewhat difficult to evaluate the probability of error exactly. In such cases, we may be content to find a bound and this problem investigates one, called the union bound approximation. Consider the QPSK case illustrated in Fig. P–10.13.2; suppose we set out to calculate P_ϵ using the following systematic approach. First, remove all signal states from consideration except the pair (s_1, s_2) and find $P_{\epsilon|s_1}$; the area covered is the left half-plane. Next, consider the pair (s_1, s_4) and find $P_{\epsilon|s_1}$; the area covered is the lower half-plane. Third, consider the pair (s_1, s_3) and find $P_{\epsilon|s_1}$; the area covered is a half-plane tilted at $-\pi/4$ to the horizontal. Now we approximate the three quadrants over which P_ϵ is to be computed by taking the union (sum) of the three (area) computations. Of course, some portions of these areas are covered more than once in this procedure and therefore this approximation gives an upper bound. Compute the union bound approximation for the QPSK system in Problem 10.13.2.

★ **10.13.5** (Nearest-Neighbor Approximation) If the probability of error is small, then a further approximation to the method developed in Problem 10.13.4 may be useful. This is explored in this problem.

a) Use the union bound approximation [cf. Problem 10.13.4] to find a bound on P_ϵ for the signal-state constellation shown in Fig. P–10.13.5. Assume all signal states are equiprobable and normalize to an average energy per symbol, designated as E_s.
b) Add a further approximation that the probability of error is governed by only the nearest neighbor (neighbors, if several are equidistant) in the constellation.
c) Plot your results for parts (a) and (b) versus E_s/η and compare.
d) Why does this approximation work fairly well when the probability of error is small?

Figure P–10.13.5.



★ **10.13.6** Repeat Problem 10.13.5(b) for M-ary PSK, $M > 2$.

★ **10.13.7** Repeat Problem 10.13.5(b) for the signal-state constellation shown in Fig. 10.42(a). Express your answer in terms of E_s, the average energy per symbol.

★ **10.14.1** A bipolar signal of amplitude ± 1 V is received in the presence of additive gaussian noise with variance 0.2 V^2. Determine the optimum threshold and the resulting probability of error if one sample of signal-plus-noise is taken. Assume both types of errors are equally costly and:

a) $P_1 = 0.5$;
b) $P_1 = 0.75$;
c) Noise variance of 0.4 V^2, P_1 as in parts (a) and (b).

★ **10.14.2** A single observation $y = s + n$ consists of a signal s which can be either 0 or 1 and additive noise with a pdf given by $p(n) = \exp(-2|n|)$. Determine the optimum detection threshold and corresponding probability of error if a) $P_1 = P_0$; b) $P_1 = 2P_0$.

★ **10.14.3** Compute and plot the probability of error for Problem 10.14.2 as a function of P_1 $(0 < P_1 < 1)$. State any assumptions you make.

★ **10.14.4** In a given experiment, N observations are taken to determine which of the two possible hypotheses H_0, H_1 is correct. In both possibilities the observations are statistically independent zero-mean gaussian random variables. Under H_0 each observable y_i has a variance σ_0^2 and under H_1 each y_i has a variance of σ_1^2. Determine the optimum signal-processing algorithm for detection with a minimum probability of error.

★ **10.14.5** A given signal $s(t)$ is to be detected in the presence of white gaussian noise based on N observations. Using the likelihood ratio test (LRT), show that for a fixed energy constraint the optimum receiver performs a cross-correlation. [*Hint:* Set up a decision based on the possible conditions $y = s + n, y = n$, write the pdf for the noise, and make a substitution of variables.]

★ **10.14.6** Either of two signals $s_1(t)$ or $s_2(t)$ is sent over $(0, T)$ in the presence of additive white gaussian noise with zero mean and variance σ^2. Assume that these two signals are equiprobable, and that they have equal energy over $(0, T)$. Use the likelihood ratio test (LRT) to determine what signal processing should be considered in the receiver.

DESIGN PROBLEMS

Design Problem 10.1 Intelsat Communication Satellite System

The first in a series of Intelsat communication satellites was launched in 1965, followed by Series II, III, IV, IVA, V, and VI designs. Intelsat was founded in August 1964, and is managed and operated by an international consortium, which had 117 members by 1989. The Intelsat system is primarily for long-distance telephone service, with some provisions for television and data services.

Three standard earth station configurations are used within the system: *A* and *B* stations operate in the 6/4-GHz band and *C* stations operate in the 14/11-GHz band. For the 6/4-GHz band, the 500 MHz band 5.925–6.425 GHz is used for the up-link and 3.7–4.2 GHz is used for the down-link. *A* stations have large antenna (26–30 m) and stringent low-noise amplifier requirements; *B* stations have somewhat more relaxed requirements.

The satellite acts as transponder (receiver/transmitter). The 500 MHz band is subdivided into 12 bands of 36 MHz (plus guard band allowance), each handled by a transponder. The method of access is FDMA (frequency-division multiple access), similar to the FDM methods discussed in Ch. 5. FDMA makes available a pool of frequencies and assigns these to users, either on a preassigned or on an assigned-on-demand basis. In the former, a carrier is modulated (using FM) for a group of standard FDM telephone channels (grouped using SSB techniques; groups are 24, 60, 96, . . . , 972 4-kHz channels.) Each assignment has its own bandwidth slot (2.5 MHz for 24 channels, 5 MHz for 60, etc.) and the overall system is referred to as FDM/FM/FDMA. Member countries lease these dedicated channel assignments.

A second type of assignment offered is that of single channel per carrier (SCPC). In this case a 36-MHz transponder is subdivided into 800 channels each with 45 kHz of bandwidth. Each channel can handle 64 kbps and QPSK is used to stay within the allocated bandwidth. In addition, one common signal channel (CSC) operating at 128 kbps (PSK) is time-shared (TDMA) for requests for channel usage and to assign frequencies. Channels are voice-activated, so that each can handle a two-way telephone conversation. This particular subsystem is referred to by the acronym SPADE (*s*ingle-channel-per-carrier *p*ulse-code-modulated multiple-*a*ccess *d*emand-assignment *e*quipment). Channels are made available by request, increasing the efficiency of the satellite transponder in channels where traffic is light and variable.

Your assignment in this design problem is dependent on instructor's requirements, but may include the following.

a) Do additional reading and investigation into the Intelsat series of communication satellites (or another series, as designated by your instructor)—this very brief description has only been to get you started! Draw frequency or time assignment

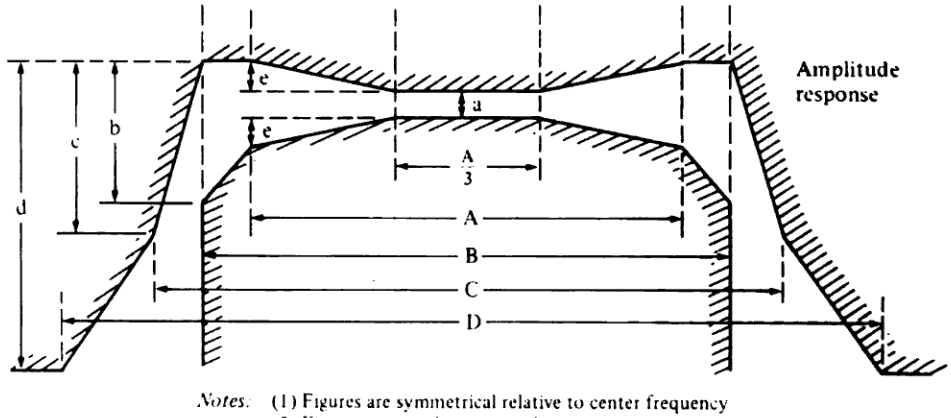
diagrams, classify the methods used, and describe the method(s) of assignment control. Some suggestions for reading are Reference 2 in Chapter 4, Reference 4 in Chapter 7, and Reference 6 in Chapter 10; additional references are cited. Also, see R. L. Freeman, *Telecommunication System Engineering*, Second ed., New York: John Wiley & Sons, 1981.

b) For TV video, two modes of transmission are available: (1) full transponder, and (2) half transponder. A typical spectral mask for a station *B* transmitter amplitude response, adapted from specifications for the Intelsat IV and IVA series of satellites that were deployed in the 1970s (courtesy COMSAT), for this service is given below. (Intelsat has started to deploy the VI series satellites in 1989.)

Mode	Video (MHz)	A (MHz)	B (MHz)	C (MHz)	D (MHz)
Half	17.5	12.6	15.75	18.0	26.5
Full	30.0	24.0	30.0	34.0	50.0

Mode	Video (MHz)	a (dB)	b (dB)	c (dB)	d (dB)	e (dB)
Half	17.5	0.3	2.5	6.5	25	0.1
Full	30.0	0.5	2.5	8.0	25	0.3

Design a (hypothetical) TDM and modulation system to handle as many 64-kbps channels as possible and yet be within the spectral restrictions for each mode for TV video. (You are not permitted to insert a filter between the transmitter and the antenna to accomplish your design objectives.) State any assumptions you make.



Notes. (1) Figures are symmetrical relative to center frequency
(2) Figures are not drawn to scale
(3) Amplitude scale is linear in dB
(4) Frequency scale is linear in MHz

Intelsat Standard B Spectral Mask (courtesy COMSAT).

APPENDIX

A Selected Mathematical Tables

A.1 TRIGONOMETRIC IDENTITIES

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$$
$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$$
$$\sin A \sin B = \tfrac{1}{2}[\cos(A - B) - \cos(A + B)]$$
$$\cos A \cos B = \tfrac{1}{2}[\cos(A + B) + \cos(A - B)]$$
$$\sin A \cos B = \tfrac{1}{2}[\sin(A + B) + \sin(A - B)]$$
$$\sin 2A = 2 \sin A \cos A$$
$$\cos 2A = 2\cos^2 A - 1 = 1 - 2\sin^2 A = \cos^2 A - \sin^2 A$$
$$\sin^2 A = \tfrac{1}{2}(1 - \cos 2A)$$
$$\cos^2 A = \tfrac{1}{2}(1 + \cos 2A)$$
$$\sin A = \frac{1}{2j}(e^{jA} - e^{-jA})$$
$$\cos A = \tfrac{1}{2}(e^{jA} + e^{-jA})$$
$$e^{\pm jA} = \cos A \pm j \sin A$$

A.2 COMPLEX FUNCTION IDENTITIES

$$z = x + jy = \sqrt{x^2 + y^2}\, e^{j \tan^{-1}(y/x)}$$
$$z^* = x - jy = \sqrt{x^2 + y^2}\, e^{-j \tan^{-1}(y/x)}$$
$$|z|^2 = zz^* = x^2 + y^2$$
$$\mathcal{R}e\{z\} = \tfrac{1}{2}[z + z^*]$$
$$\mathcal{I}m\{z\} = \frac{1}{2j}[z - z^*]$$
$$\mathcal{R}e\{z_1 z_2\} = \mathcal{R}e\{z_1\}\mathcal{R}e\{z_2\} - \mathcal{I}m\{z_1\}\mathcal{I}m\{z_2\}$$
$$\mathcal{I}m\{z_1 z_2\} = \mathcal{R}e\{z_1\}\mathcal{I}m\{z_2\} + \mathcal{I}m\{z_1\}\mathcal{R}e\{z_2\}$$

A.3 SERIES

Exponential:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots.$$

Trigonometric:

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \cdots.$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \cdots.$$

Binomial (for $x^2 < 1$):

$$(1 \pm x)^n = 1 \pm nx + \frac{1}{2!}n(n-1)x^2 \pm \frac{1}{3!}n(n-1)(n-2)x^3 + \cdots.$$

$$(1 \pm x)^{-n} = 1 \mp nx + \frac{1}{2!}n(n+1)x^2 \mp \frac{1}{3!}n(n+1)(n+2)x^3 + \cdots.$$

A.4 SUMMATIONS

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

A.5 INDEFINITE INTEGRALS

$$\int \sin ax\,dx = -\frac{1}{a}\cos ax$$

$$\int \cos ax\,dx = \frac{1}{a}\sin ax$$

$$\int \sin^2 ax\,dx = \frac{x}{2} - \frac{\sin 2ax}{4a}$$

$$\int \cos^2 ax\,dx = \frac{x}{2} + \frac{\sin 2ax}{4a}$$

$$\int \sin ax \cos ax\, dx = \frac{1}{2a} \sin^2(ax)$$

$$\int x \sin ax\, dx = \frac{1}{a^2}(\sin ax - ax \cos ax)$$

$$\int x \cos ax\, dx = \frac{1}{a^2}(\cos ax + ax \sin ax)$$

$$\int x^2 \sin ax\, dx = \frac{1}{a^3}(2ax \sin ax + 2 \cos ax - a^2x^2 \cos ax)$$

$$\int x^2 \cos ax\, dx = \frac{1}{a^3}(2ax \cos ax - 2 \sin ax + a^2x^2 \sin ax)$$

$$\int \sin ax \sin bx\, dx = \frac{\sin (a - b)x}{2(a - b)} - \frac{\sin (a + b)x}{2(a + b)} \qquad a^2 \neq b^2$$

$$\int \cos ax \cos bx\, dx = \frac{\sin (a - b)x}{2(a - b)} + \frac{\sin (a + b)x}{2(a + b)} \qquad a^2 \neq b^2$$

$$\int \sin ax \cos bx\, dx = -\frac{\cos (a - b)x}{2(a - b)} - \frac{\cos (a + b)x}{2(a + b)} \qquad a^2 \neq b^2$$

$$\int e^{ax}\, dx = \frac{1}{a} e^{ax}$$

$$\int xe^{ax}\, dx = \frac{1}{a^2} e^{ax}(ax - 1)$$

$$\int x^2e^{ax}\, dx = \frac{1}{a^3} e^{ax}(a^2x^2 - 2ax + 2)$$

$$\int x^ne^{ax}\, dx = \frac{x^n}{a} e^{ax} - \frac{n}{a} \int x^{n-1}e^{ax}\, dx$$

$$\int e^{ax} \sin bx\, dx = \frac{1}{a^2 + b^2} e^{ax}(a \sin bx - b \cos bx)$$

$$\int e^{ax} \cos bx\, dx = \frac{1}{a^2 + b^2} e^{ax}(a \cos bx + b \sin bx)$$

$$\int \left[\frac{\sin ax}{x}\right]^2 dx = a \int \frac{\sin 2ax}{x} dx - \frac{\sin^2 ax}{x}$$

$$\int \frac{dx}{a^2 + b^2x^2} = \frac{1}{ab} \tan^{-1}\left(\frac{bx}{a}\right)$$

$$\int \frac{x^2\, dx}{a^2 + b^2x^2} = \frac{x}{b^2} - \frac{a}{b^3} \tan^{-1}\left(\frac{bx}{a}\right)$$

$$\int \frac{dx}{(a^2 + b^2x^2)^2} = \frac{x}{2a^2(a^2 + b^2x^2)} + \frac{1}{2a^3b}\tan^{-1}\left(\frac{bx}{a}\right)$$

$$\int \frac{x^2\,dx}{(a^2 + b^2x^2)^2} = \frac{-x}{2b^2(a^2 + b^2x^2)} + \frac{1}{2ab^3}\tan^{-1}\left(\frac{bx}{a}\right)$$

$$\int \frac{dx}{(a^2 + b^2x^2)^3} = \frac{x}{4a^2(a^2 + b^2x^2)^2} + \frac{3x}{8a^4(a^2 + b^2x^2)} + \frac{3}{8a^5b}\tan^{-1}\left(\frac{bx}{a}\right)$$

A.6 DEFINITE INTEGRALS

$$\int_0^\infty \frac{\sin ax}{x}\,dx = \begin{cases} \pi/2 & a > 0 \\ 0 & a = 0 \\ -\pi/2 & a < 0 \end{cases}$$

$$\int_0^x \frac{\sin u}{u}\,du \triangleq \text{Si}\,(x) \quad \text{(a tabulated integral as a function of } x\text{)}$$

$$\int_0^\infty \frac{\sin^2 ax}{x^2}\,dx = |a|\pi/2$$

$$\int_0^\infty e^{-ax^2}\,dx = \frac{1}{2}\sqrt{\pi/a}$$

$$\int_0^\infty xe^{-ax^2}\,dx = \frac{1}{2a}$$

$$\int_0^\infty x^2e^{-ax^2}\,dx = \frac{1}{4a}\sqrt{\pi/a}$$

$$\int_0^\infty \frac{dx}{(x^2 + a^2)(x^2 + b^2)} = \frac{\pi}{2ab(a + b)} \qquad a > 0, b > 0$$

$$\int_0^\infty \frac{dx}{ax^4 + b} = \frac{\pi}{2\sqrt{2b}}\left(\frac{b}{a}\right)^{1/4} \qquad ab > 0$$

$$\int_0^\infty \frac{dx}{ax^6 + b} = \frac{\pi}{3b}\left(\frac{b}{a}\right)^{1/6} \qquad ab > 0$$

$$\int_0^\infty \frac{dx}{1 + x^{2n}} = \frac{[\pi/(2n)]}{\sin\,[\pi/(2n)]} \qquad n \text{ an integer, } n > 0$$

Selected References for Further Reading

1. S. M. Selby (ed.). *CRC Standard Mathematical Tables*. Fifteenth ed. Cleveland, Ohio: The Chemical Rubber Co., 1967.
2. H. B. Dwight. *Tables of Integrals and Other Mathematical Data*. Fourth ed. New York: Macmillan, 1961.
3. I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series and Products*. Fourth ed. New York: Academic Press, 1965.

APPENDIX

B Decibels

The decibel is a logarithmic unit of measure used for comparing two power levels. Assigning P_r as the reference power level, the decibel (dB) is defined by the equation

$$(\text{dB}) \triangleq 10 \log_{10} (P/P_r). \tag{B.1}$$

Because the decibel is a logarithmic unit, operations of multiplication and division reduce to addition and subtraction, and powers and roots reduce to multiplication and division. Addition and subtraction, however, require conversion to numeric values. The sign of a logarithm changes when its argument is inverted. Given a power ratio expressed in decibels, the power ratio can be found from the inverse of Eq. (B.1):

$$P/P_r = 10^{(\text{dB})/10}. \tag{B.2}$$

Decibels are also used to indicate absolute power levels by adding a third letter to the notation. If the reference power level P_r is 1 watt, the power P is expressed in "decibels above 1 W" (dBW) by

$$P_{\text{dBW}} = 10 \log_{10} P. \tag{B.3}$$

In a similar manner, if the reference power P_r is 1 milliwatt, the power P is expressed in "decibels above 1 mW" (dBm).

The definition in Eq. (B.1) can be rewritten in terms of voltages V, V_r developed across resistances R, R_r in the following manner:

$$(\text{dB}) = 10 \log_{10} \left[\frac{V^2/R}{V_r^2/R_r}\right] \tag{B.4a}$$

$$= 20 \log_{10} (V/V_r) - 10 \log_{10} (R/R_r). \tag{B.4b}$$

For the case in which $R = R_r$, this becomes

$$(\text{dB}) = 20 \log_{10} (V/V_r). \tag{B.5}$$

If, in addition, the reference voltage level is taken as unity, we have

$$V_{\text{dB}} = 20 \log_{10} V. \tag{B.6}$$

Sometimes Eq. (B.6) is used as a definition of the decibel. This is valid only as long as the proper normalizing factors are considered whenever any comparison is made to power calculations using the definition of Eq. (B.1).

The predominant usage of the decibel in communication systems is for power ratios. Unless otherwise specified, Eqs. (B.1) and (B.2) are assumed to be the defining relations.

DRILL PROBLEM B.1

(a) Convert the following power ratios to dB: 4000, 0.003. (b) Convert the following to numeric values (in terms of power): 29.3 dB, −7 dBW, 27 dBm.

ANSWER: (a) 36.0 dB, −25.2 dB; (b) 851, 0.200 W, 0.501 W.

APPENDIX

C Broadcast Frequency Bands

TABLE C.1 Frequency Designations

ELF	Extremely low frequency	30–300 Hz
VF	Voice frequency	300–3000 Hz
VLF	Very low frequency	3–30 kHz
LF	Low frequency	30–300 kHz
MF	Medium frequency	300–3000 kHz
HF	High frequency	3–30 MHz
VHF	Very high frequency	30–300 MHz
UHF	Ultra high frequency	300–3000 MHz
SHF	Super high frequency	3–30 GHz
EHF	Extremely high frequency	30–300 GHz

	10^6	10^5	10^4	10^3	10^2	10^1	10^0	10^{-1}	10^{-2}	λ, m
ELF	VF	VLF	LF	MF	HF	VHF	UHF	SHF	EHF	
	300 Hz	3 kHz	30 kHz	300 kHz	3 MHz	30 MHz	300 MHz	3 GHz	30 GHz	f

TABLE C.2 Radar Frequency Bands[a]

Band	Frequency	Band	Frequency
HF	3–30 MHz	X	8–12 GHz
VHF	30–300 MHz	K_u	12–18 GHz
UHF	300–3000 MHz	K	18–27 GHz
L	1–2 GHz	K_a	27–40 GHz
S	2–4 GHz	V	40–75 GHz
C	4–8 GHz	W	75–110 GHz
		millimeter	110–300 GHz

[a] E. C. Jordan, (ed), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, Indianapolis, Howard W. Sams & Co., 1985.

AM Band†

Stations are assigned carrier frequencies at 10 kHz intervals from 540 to 1600 kHz. The bandwidth of transmissions is nominally about 10 kHz. Stations in local proximity are usually assigned carrier frequencies which are separated by 30 kHz or more. Interference between transmissions is controlled by a combination of frequency allocation, transmitter power, transmitting antenna pattern, and possible night-time operating restrictions. Required carrier stability is ±20 Hz; licensed power output is for an unmodulated carrier.

FM Band†

Stations are assigned carrier frequencies at 200 kHz intervals from 88 MHz to 108 MHz. Transmissions received beyond the line-of-sight distance are weak and subject to fading. FM receivers tend to suppress weak signals in the presence of strong signals. These effects help to define the station broadcast area. Peak frequency deviation is 75 kHz. Minimum required carrier frequency stability is ±0.002% (that is, ±2 kHz at 100 MHz).

TABLE C.3 Permissible Transmitted Power Levels as Specified by FCC

AM Band:	Broadcast Class	Average Power, kW
	Local	0.1–1.0
	Regional	0.5–5.0
	Clear	0.25–50
FM band: 0.25, 1, 3, 5, 10, 25, 50, 100 kW, depending on class of service (community size) and area coverage for license granted.		
Television-video:	Channels	Effective Radiated Power (average), kW
	VHF (2–6)	100
	VHF (7–13)	316
	UHF	5000
Television-audio: 10% (−10 dB)–20% (−7 dB) of power in video carrier.		

† Ibid.

TABLE C.4 VHF Television Station Allocations[a]

Channel Number	Frequency Span (MHz)	Video Carrier (MHz)[b]
1		Not used
2	54–60	55.25
3	60–66	61.25
4	66–72	67.25
5	76–82	77.25
6	82–88	83.25 (Standard FM band 88–108)
7	174–180	175.25
8	180–186	181.25
9	186–192	187.25
10	192–198	193.25
11	198–204	199.25
12	204–210	205.25
13	210–216	211.25

[a] For more information on television systems, see K. B. Benson (ed.), *Television Engineering Handbook*, New York, McGraw-Hill, 1986, Chapter 13.

[b] Audio carrier is 4.50 MHz higher than the video carrier.

TABLE C.5 UHF Television Station Allocations

Channel Number	Frequency Span (MHz)	Video Carrier (MHz)[a]	Channel Number	Frequency Span (MHz)	Video Carrier (MHz)[a]
14	470–476	471.25	42	638–644	639.25
15	476–482	477.25	43	644–650	645.25
16	482–488	483.25	44	650–656	651.25
17	488–494	489.25	45	656–662	657.25
18	494–500	495.25	46	662–668	663.25
19	500–506	501.25	47	668–674	669.25
20	506–512	507.25	48	674–680	675.25
21	512–518	513.25	49	680–686	681.25
22	518–524	519.25	50	686–692	687.25
23	524–530	525.25	51	692–698	693.25
24	530–536	531.25	52	698–704	699.25
25	536–542	537.25	53	704–710	705.25
26	542–548	543.25	54	710–716	711.25
27	548–554	549.25	55	716–722	717.25
28	554–560	555.25	56	722–728	723.25
29	560–566	561.25	57	728–734	729.25
30	566–572	567.25	58	734–740	735.25
31	572–578	573.25	59	740–746	741.25
32	578–584	579.25	60	746–752	747.25
33	584–590	585.25	61	752–758	753.25
34	590–596	591.25	62	758–764	759.25
35	596–602	597.25	63	764–770	765.25
36	602–608	603.25	64	770–776	771.25
37	608–614	609.25	65	776–782	777.25
38	614–620	615.25	66	782–788	783.25
39	620–626	621.25	67	788–794	789.25
40	626–632	627.25	68	794–800	795.25
41	632–638	633.25	69	800–806	801.25

[a] Audio carrier is 4.50 MHz higher than the video carrier.

TABLE C.6 Cable Television Channel Frequencies

CATV Channel Number	Letter Designator	Band Name	Frequency (MHz)[a]	CATV Channel Number	Band Name	Frequency (MHz)[a]
2–6		VHF-L	(see TV)[b]	37	Hyper	300–306
95	A-5	Mid	90–96	38		306–312
96	A-4		96–102	39		312–318
97	A-3		102–108	40		318–324
98	A-2		108–114	41		324–330
99	A-1		114–120	42		330–336
14	A		120–126	43		336–342
15	B		126–132	44		342–348
16	C		132–138	45		348–354
17	D		138–144	46		354–360
18	E		144–150	47		360–366
19	F		150–156	48		366–372
20	G		156–162	49		372–378
21	H		162–168	50		378–384
22	I		168–174	51		384–390
7–13		VHF-H	(see TV)	52		390–396
23	J	Super	216–222	53		396–402
24	K		222–228	54		402–408
25	L		228–234	55		408–414
26	M		234–240	56		414–420
27	N		240–246	57		420–426
28	O		246–252	58		426–432
29	P		252–258	59		432–438
30	Q		258–264	60		438–444
31	R		264–270	61		444–450
32	S		270–276	62		450–456
33	T		276–282	63		456–462
34	U		282–288	64		462–468
35	V		288–294	65	Ultra	468–474
36	W		294–300	⋮		⋮
				94		642–648

[a] Video carrier frequency is 1.25 MHz higher than the lower edge of band assignment; audio carrier frequency is 4.50 MHz higher than the video carrier frequency for each channel.

[b] Assignments for Channels 5 and 6 can be shifted 2 MHz upward, thus allowing an undesignated allocation for 72–78 MHz (Channel 1).

TABLE C.7 Amateur Bands[a]

Band Designation	Frequency Allocation (MHz)
160 meter	1.800–2.000
80 meter	3.500–4.000
40 meter	7.000–7.300
20 meter	14.000–14.350
15 meter	21.000–21.450
10 meter	28.000–29.700
6 meter	50.0 –54.0
2 meter	144–148
	220–225
	420–450
	1,215–1,300
	2,300–2,450
	3,300–3,500
	5,650–5,925
	10,000–10,500
	24,000–24,500
	48,000–50,000
	71,000–76,000
	165,000–170,000
	240,000–250,000
	above 300,000

Notes
1. Maximum authorized power input to the final transmitter stage is 1 kW.
2. Maximum amplitude modulation allowed is 100%. Except for brief tests or adjustments, the transmitter may not emit a carrier wave on frequencies below 51 MHz unless modulated for the purpose of communication.
3. Single-sideband: lower sideband is customarily used on 80 m and 40 m, upper sideband on the higher frequencies.

[a] For more information, see *FCC Rules and Regulations* Vol. VI, Part 97, or *The Radio Amateur's Handbook*, American Radio Relay League, Newington, Conn., 06111.

TABLE C.8 Citizens Radio Service[a]

These frequency allocations are for fixed and mobile stations intended for short-distance personal or business radiocommunications, radio signaling, and control of remote objects or devices by radio.

Classifications

Class station	Service	Band (MHz)	Maximum Transmitter Input Power (W)	Maximum Average Output Power (W)	Frequency Tolerance (%)
A	General	460–470	60	48	0.0005
B	General	460–470	5	4	0.5
C	Remote control	26.96–27.23	5	4	0.005
	Hobby control	72–76	5	4	0.005
D	Radio telephone	26.96–27.41	5	4	0.005

[a] For more information, see *FCC Rules and Regulations*, Vol. VI, Part 95.

Frequency Allocations
Class A and Class B

Base Station/ Mobile (MHz)	Mobile Only (MHz)
462.550	467.550
462.575	467.575
462.600	467.600
462.625	467.625
462.650	467.650
462.675	467.675
462.700	467.700
462.725	467.725

Class C (mobile only; may employ only amplitude tone modulation or on-off keying of the unmodulated carrier)

1. For control of remote objects or devices by radio:
   26.995 MHz
   27.045
   27.095
   27.145
   27.195
   27.255
2. For radio remote control of any model used for hobby purposes:
   72.16 MHz
   72.32
   72.96
3. For radio remote control of aircraft models only:
   72.08 MHz
   72.24
   72.40
   75.64

Class D
(mobile radio telephone use only)

Channel	Frequency (MHz)	Channel	Frequency (MHz)	Channel	Frequency (MHz)
1	26.965	15	27.135	28	27.285
2	26.975	16	27.155	29	27.295
3	26.985	17	27.165	30	27.305
4	27.005	18	27.175	31	27.315
5	27.015	19	27.185	32	27.325
6	27.025	20	27.205	33	27.335
7	27.035	21	27.215	34	27.345
8	27.055	22	27.225	35	27.355
9[a]	27.065	24	27.235	36	27.365
10	27.075	25	27.245	37	27.375
11	27.085	23	27.255	38	27.385
12	27.105	26	27.265	39	27.395
13	27.115	27	27.275	40	27.405
14	27.125				

[a] Channel 9 is reserved for emergency communications: (a) involving the immediate safety of life of individuals or the immediate protection of property; (b) necessary to render assistance to a motorist.

TABLE C.9 Civil Defense[a]
Disaster Communication Service

Channel	Assigned Frequency (MHz)	Bandwidth (kHz)	Service
1	1.7505	1	radiotelegraph
2	1.7515	1	radiotelegraph
3	1.7525	1	radiotelegraph
4	1.7535	1	radiotelegraph
5	1.7545	1	radiotelegraph
6	1.7555	1	radiotelegraph
7	1.7565	1	radiotelegraph
8	1.7575	1	radiotelegraph
9	1.7615	7	radiotelephone
10	1.7685	7	radiotelephone
11	1.7755	7	radiotelephone
12	1.7825	7	radiotelephone
13	1.7895	7	radiotelephone
14	1.7965	7	radiotelephone

Channel 9 is reserved for "Scene of Disaster."

[a] For more information, see *FCC Rules and Regulations*, Vol. VI, Part 99.

APPENDIX

D Commercial Television Transmissions

Commercial television systems exhibit several aspects of signal design choices and compromises in communication systems. The emphasis here is on these signal design choices and not on the circuits. Television standards vary from country to country and the systems discussed here, unless otherwise stated, are those presently used in the United States.

D.1 BLACK-AND-WHITE TELEVISION

In order to transmit two-dimensional (in this case, picture) information via a one-dimensional coordinate system (in this case, time), it is necessary to employ some type of scanning technique. In television, each picture is divided into 525 lines and is scanned line by line. The time for one complete scan of a picture has been chosen to be $\frac{1}{30}$ sec so that power line interference appears stationary and hence is much less objectionable in the picture display. Each complete picture is called a *frame*.

Even though the human eye interprets picture sequences shown at a rate of about 15 frames per sec or greater as continuous motion, there is some discernible flicker present until frame rates on the order of about 40 per sec are used.[†] To eliminate discernible flicker in the television picture, alternate lines are sent at a rate of 60 per sec. This is called *interlaced scanning* and is illustrated in Fig. D.1. The scan starts at the upper left and proceeds from left to right, retracing quickly after each scan. (The slower vertical scan rate produces a slight tilt to the lines.) This is repeated to the bottom of the display, where only a half-line is scanned. A total of 262.5 lines have now been scanned and this is called a *field*. The time to scan one field is $\frac{1}{60}$ sec. The required horizontal scan frequency is then $(262.5)(60) = 15{,}750$ Hz.

At the end of the field, the cathode ray tube beam retraces rapidly upward (while still zigzagging horizontally) until it reaches the top center of the display screen. The beam is biased off (blanked) during both horizontal and vertical retrace times so this is not visible in the display. Next the beam repeats the horizontal scan at the 15.75-kHz rate with a slow downward motion at the vertical scan rate of 60 Hz. It follows the dashed-line path shown in Fig. D.1 until it reaches the lower right corner of the screen and the vertical retrace is again initiated. The two fields now total 525 lines.

[†] Silent movie film is generally run at 16 frames per sec; sound movie film is run at 24 frames per sec. Theater projectors use a shutter that opens and closes at double the frame rate (i.e., at 48 per sec) to eliminate discernible flicker.

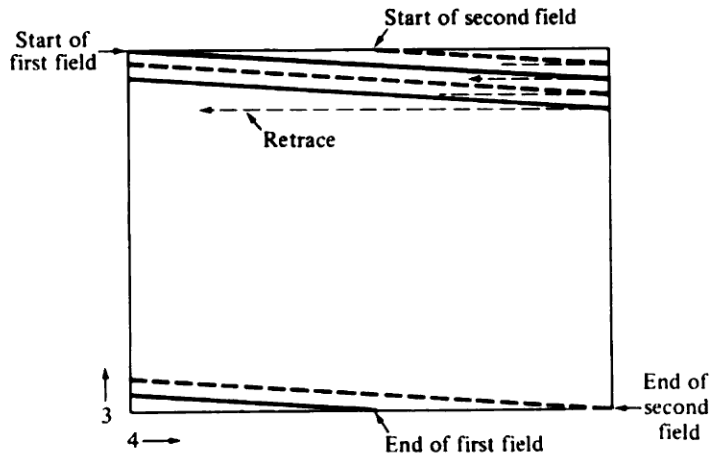


Figure D.1 Interlaced scanning.

interlaced to form the complete picture (frame). In summary, the interlaced scanning method provides a 60-Hz field rate to eliminate flicker while maintaining a 30-Hz frame rate.

The line pattern used for each frame is called a *raster*. (The raster is all that is seen when no picture information is being received.) Although the size of the raster depends on the particular receiver (usually specified as a diagonal measurement), the relative dimensions of width to height—known as the aspect ratio—are standardized at 4:3.

Using the 4:3 aspect ratio, we can consider the complete picture as an array of 700 × 525 dots of varying intensity. Thus there are a maximum of (700)(525)(30) = 11,025,000 picture elements to be sent each second. (Actually, this is a pessimistic choice because about 21 lines are lost during the relatively long vertical retrace time.) Using the criterion that the system rise time must equal one picture element, a system bandwidth requirement ranging from *RC*-type response:

$$B \approx 0.35/t_r = (0.350)(11{,}025{,}000) = 3.86 \text{ MHz}$$

to a much sharper response:

$$B \approx 1/(2t_r) = (0.500)(11{,}025{,}000) = 5.51 \text{ MHz}$$

gives us some idea of the type of bandwidth required. In practice, a bandwidth of 4 MHz is assumed adequate.

Video information for television is transmitted using large-carrier amplitude modulation. Use of double-sideband would then require 8 MHz per channel for the video. However, channel allocations of 6-MHz width for experimental stations as early as 1936 have tended to restrict allowable bandwidths to within these limits. To circumvent these restrictions, a form of vestigial sideband transmission is used for the video information. A simplified spectral diagram of a television transmission is shown in Fig. D.2.

The lower sideband of the video transmission is attenuated below 0.75 MHz and is completely band-limited to 1.25 MHz below the video carrier. (Complete elimination

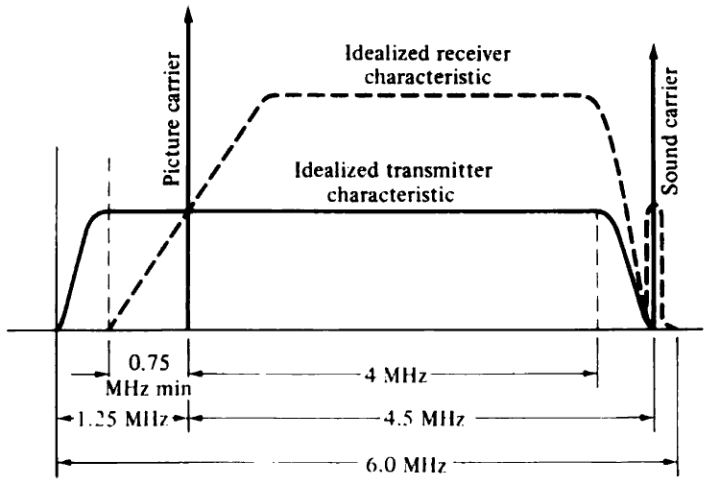


Figure D.2 Simplified spectrum of a television channel.

of one sideband at the transmitter would require much more elaborate and expensive filtering.) The receiver bandpass characteristic completes the vestigial sideband characteristic, as shown in Fig. D.2. The audio information is transmitted using FM, with a peak frequency deviation of 25 kHz, and is centered 4.5 MHz above the video carrier.

The amplitude modulation levels used for the video are shown in Fig. D.3. In the United States, a negative modulation standard is used—i.e., less amplitude corresponds to a brighter scene, while more amplitude corresponds to a darker scene. Because most pictures contain more white than black levels, somewhat greater power efficiency is possible with negative modulation than with positive modulation. Reference black is defined by 70% modulation and the minimum (white) modulation level is 12.5%, as shown in Fig. D.3.

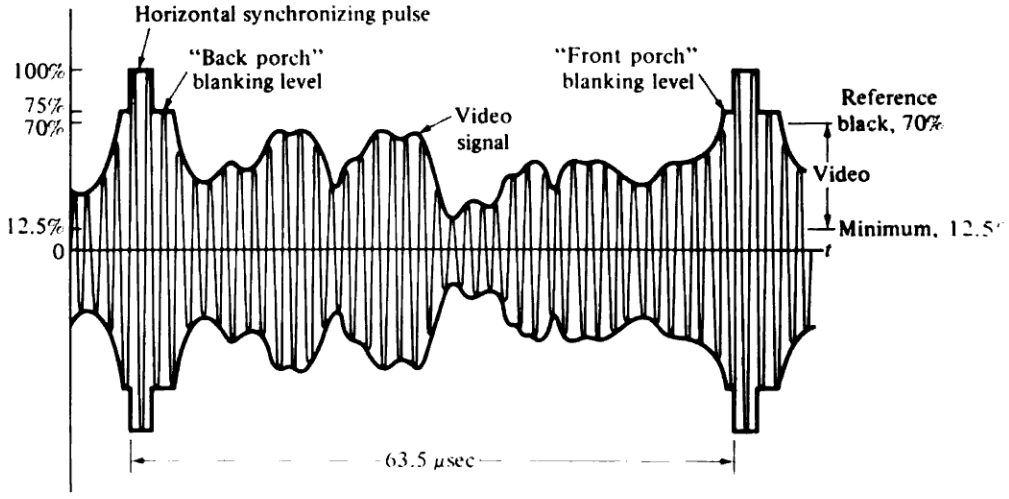


Figure D.3 Video-modulated waveform.

Proper synchronization is necessary to reconstruct a stable picture. Synchronizing pulses are sent at the 100% modulation level to be easily identifiable and are sent at the beginning of each line scan. A blanking level sent at 75% modulation makes sure the CRT beam is cut off so that transients and the horizontal and vertical retraces are not visible in the display.

Vertical synchronization and the allowance for the one-half line scan is effected by generating a pulse train at the end of each field. This pulse train consists of a series of equalizing pulses and vertical synchronizing pulses generated at twice the horizontal scan rate (that is, at 31.5 kHz) as shown in Fig. D.4. The horizontal sweep oscillator is triggered by the leading edge of each pulse at the 31.5-kHz rate. (The fact that it can only sweep half the picture width during this time need not bother us because the CRT is blanked during the entire synchronization interval.) The equalizing pulses are kept narrow so that they will not trigger the vertical sweep oscillator. The vertical synchronization is composed of six consecutive wide pulses. The notches between these pulses are needed to maintain horizontal oscillator synchronization. These wide pulses are applied to an integrator to form a signal having sufficiently large amplitude to separate it from any other signals and to trigger the vertical sweep oscillator. More equalizing pulses and a few pulses at the 15.75-kHz horizontal scan rate follow. By the time the first video trace appears, all transient effects from synchronization have disappeared in the receiver. Note that the 31.5-kHz rate is necessary to cause the half-line horizontal trace variation for the interlaced scanning.

Additional information can be transmitted by the commercial television signal in several ways without interfering with the normal picture and sound. These include use of the horizontal blanking interval, the vertical blanking interval, the audio channel (by using time or frequency multiplexing), and the video channel (by using time or frequency multiplexing). The first and last options are used in color broadcasting,

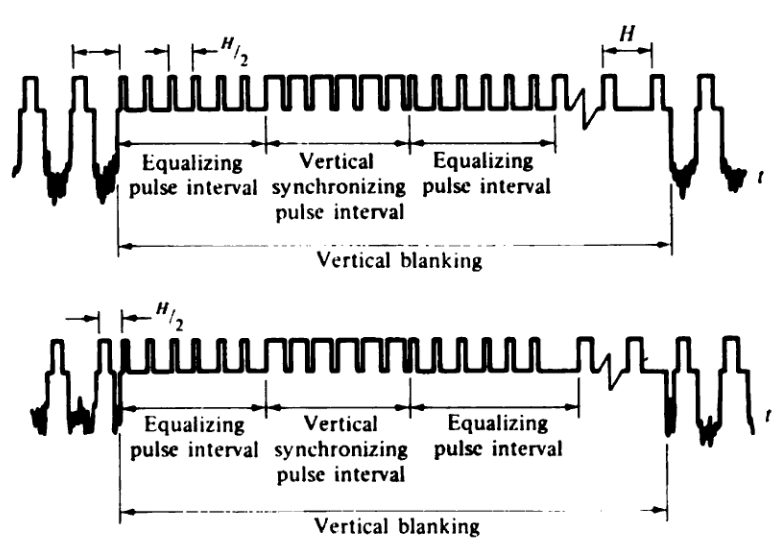


Figure D.4 Vertical synchronization for each field.

as described in a later section. Multiplexing additional information in the audio channel is used by some subscription services. We devote some attention here to the recent interests in use of the vertical blanking interval.

As mentioned previously, the time taken for the vertical retrace interval is equivalent to 21 horizontal lines. The first nine lines are occupied by six equalizing pulses, followed by the vertical synchronizing pulse interval and another set of six equalizing pulses, as shown in Fig. D.4. This leaves lines numbered 10–21 available for other uses. (It is possible to see these lines by adjusting the vertical hold control of the television receiver until the bar separation between pictures rolls into view.)

Some of the lines in the vertical blanking interval have been assigned already. Lines 17, 18 carry vertical-interval test signal (VITS) transmissions. These video and chroma test signals are used to check the quality of network and station transmissions. Line 19 carries the vertical-interval reference (VIR) signals consisting of black, luminance, and chroma references. Some television receivers are equipped with automatic color tint and level circuitry using the VIR signals as a reference.

Line 21 is used by some stations for picture captioning to aid those viewers with hearing impairments. All of line 21 in the first field and one-half of line 21 in the second field is available for this purpose. The NRZ data signal that is transmitted uses standard seven-bit-plus-parity ASCII coding. The decoder extracts the caption data from the vertical blanking interval and displays it near the bottom of the television screen.

Receiver designs can be expected to change in the near future as a result of advances in digital hardware. One system already proposed for production, called improved definition television (IDTV) uses digital memory to double the number of horizontal lines displayed, resulting in 525 lines per field. The new scanning lines to insert between existing lines are derived from combinations of the line above, the line below, and the corresponding line(s) from the preceding field.

D.2 THEORY OF COLOR

Before we proceed directly to color television, it will be advantageous to briefly discuss some of the properties of color and color mixing. It is found that three primary colors are needed to make a color reproduction that is acceptable to human vision. By suitable mixing of these primary colors it is possible to approximate all colors found in nature. The three primary colors chosen are red, green, and blue. Other colors are synthesized by appropriate mixing of these primary colors. If we mix red and green light, for example, we obtain yellow.[†] In a similar manner, red and blue produce a pink-violet color known as magenta while green and blue produce a particular blue known as cyan. If we mix all three (in equal proportions), the result is known as white.

[†] All additions here are assumed to be noncoherent; i.e., addition on an energy basis.

In television, it is customary to describe a picture in terms of its *luminance* (brightness) and *chrominance* (color). The chrominance information can in turn be described in terms of hue and saturation. To demonstrate the use of these terms, let us place the three primary colors at the corners of an equilateral triangle, as shown in Fig. D.5. From the comments on color mixing, it follows that yellow, magenta, and cyan will be located along the sides of this color (chrominance) triangle and white will be at the center.

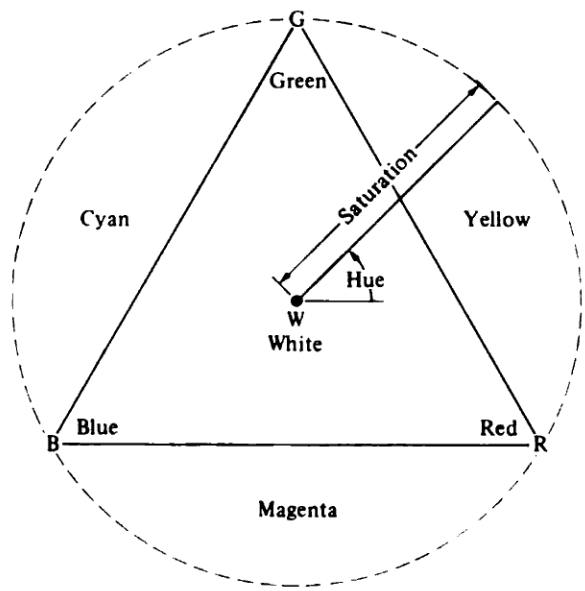


Figure D.5 Color triangle showing hue and saturation.

Now suppose we describe this color information in terms of circular coordinates. This can be done by constructing a circle enclosing the color triangle and defining the colors by a radial magnitude and an angle. The *hue* (color) of the light then varies as the angle and the proportion of pure (i.e., saturated) color to white, called the *saturation,* varies as the radial distance from the center. Note that the saturation of a given color depends on its dilution with white light. Near the center a color is said to be highly desaturated (commonly called pale or pastel) while farther out it is said to be more saturated. However, all points along a line of fixed angle have the same hue.

The luminosity of light is a measure of its incident energy. This may be considered as a third dimension in Fig. D.5, perpendicular to the plane shown. The wavelength response of detectors varies, so it is necessary to also consider the response of the detector. In the case of commercial television, of course, this is the human eye, which is known to be unequally responsive to energy at different wavelengths. A composite of many measurements on the response of the human eye is shown in Fig. D.6. From this we conclude that the response of the human eye (i.e., luminosity as seen by the eye) is greatest at a wavelength of 550 millimicrons (a yellow-green color). The response drops off toward larger and shorter wavelengths, with a more

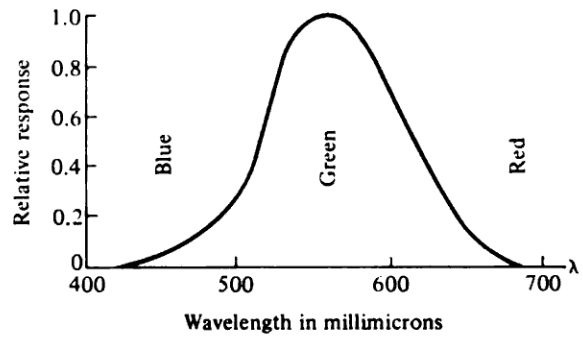


Figure D.6 Response of the standard human eye.†

rapid drop toward blue. In terms of the primary colors, then, the human eye interprets white light not as 33.3% of each but as about 59% green, 30% red, and 11% blue.

D.3 COLOR TELEVISION—THE NTSC SYSTEM

In the late 1940s and early 1950s several different color television systems were being proposed. Most were sequential systems, i.e., the color information was transmitted one color at a time in sequence. For example, one proposed system (CBS) which was approved by the FCC on a trial basis for a few years used a revolving color wheel so that one picture field was scanned in red, the next in green, etc. Such systems, while much simpler than ones in which all the color information is transmitted simultaneously, were found not to be readily compatible with existing black-and-white (B/W) television methods and bandwidth restrictions.

In the early 1950s a committee—the National Television Systems Committee—was formed to try to resolve the problem of compatibility. The resulting NTSC system design was authorized by the FCC in 1953 and is the system presently used in the United States. Instead of transmitting the red, green, and blue signals sequentially, the NTSC system separates the information into luminance and chrominance signals which are transmitted simultaneously using frequency multiplexing. Most of the basic ideas of the NTSC system, with some refinements, have been incorporated into other color television systems used in other countries.

The NTSC system is designed to be compatible with the monochrome (B/W) system based on 525 lines and 60 fields per sec with interlaced scanning. A luminance component Y is formed by the summation of the three primary color components R,

† This is the light-adapted response (photopic vision) corresponding to the brightness of normal daylight conditions. When adapted to darkness (scotopic vision) the peak response of the human eye shifts downward in wavelength by about 50 millimicrons. However, the eye loses some of its perception of color under scotopic conditions.

G, B. The resulting color signals are weighted in the same proportion as the human eye's response to the primary color frequencies so that

$$E_Y = 0.30E_R + 0.59E_G + 0.11E_B. \tag{D.1}$$

All of the information necessary for a monochrome picture is contained in this luminance signal.

In addition to the luminance signal, two color difference signals are transmitted. These color differences are formed by subtracting the luminance. Because the $(G - Y)$ component is smaller than the other two (and therefore subject to larger errors caused by noise), the $(R - Y)$ and $(B - Y)$ components are the two that are sent. These color difference signals are transmitted using quadrature multiplexing (DSB-SC) on a subcarrier placed ω_s rad/sec above the picture carrier. The chrominance (color difference) signal can then be written as

$$(E_R - E_Y) \cos \omega_s t + (E_B - E_Y) \sin \omega_s t. \tag{D.2}$$

The summation of Eqs. (D.1) and (D.2) forms the total video signal, which is then modulated on the picture carrier using vestigial sideband techniques.

At the receiver, the following signals can be detected:†

$$\begin{Bmatrix} E_R - E_Y \\ E_B - E_Y \\ E_Y \end{Bmatrix}. \tag{D.3}$$

Equation (D.1) can be rewritten as

$$(E_G - E_Y) = -\frac{0.30}{0.59}(E_R - E_Y) - \frac{0.11}{0.59}(E_B - E_Y). \tag{D.4}$$

Thus the third color difference signal $(E_G - E_Y)$ can be found by proper weighting and addition (known as "matrixing"). The signals $E_Y, (E_R - E_Y), (E_G - E_Y)$, and $(E_B - E_Y)$ can then be applied to the appropriate control grids of the picture tube.

Now consider the values of the $(R - Y)$, $(B - Y)$, and Y components for six saturated colors. These are listed in Table D.1. The two color differences are transmitted in quadrature at the color subcarrier frequency ω_s and we can use phasor relationships. Writing the color differences in terms of magnitude and phase, we have

$$\sqrt{(R - Y)^2 + (B - Y)^2} \tan^{-1} [(R - Y)/(B - Y)]. \tag{D.5}$$

Note that the magnitude in Eq. (D.5) is a measure of color saturation and the phase is a measure of the hue. The total maximum instantaneous amplitude is proportional to the sum of the luminance component, Y, and the magnitude of Eq. (D.5). This quantity is listed in the last column of Table D.1.

† It is possible to use detection along axes that are rotated by 33° (called I and Q axes) to gain some added bandwidth in the color difference signals. This will be discussed shortly. Modern receivers for home entertainment use generally employ the method outlined here and sacrifice the full color bandwidth (and hence definition) available.

TABLE D.1 Addition of Luminance and Chrominance Components for Six Saturated Colors

Color	R	G	B	Components Y	$(R-Y)$	$(B-Y)$	Magnitude/Phase [using Eq. (D.5)]	Total Maximum Magnitude
Blue	0	0	1	0.11	−0.11	0.89	$0.90\,\underline{/353°}$	1.01
Red	1	0	0	0.30	0.70	−0.30	$0.76\,\underline{/113°}$	1.06
Magenta	1	0	1	0.41	0.59	0.59	$0.83\,\underline{/45°}$	1.24
Green	0	1	0	0.59	−0.59	−0.59	$0.83\,\underline{/225°}$	1.42
Cyan	0	1	1	0.70	−0.70	0.30	$0.76\,\underline{/293°}$	1.46
Yellow	1	1	0	0.89	0.11	−0.89	$0.90\,\underline{/173°}$	1.78

For a monochrome transmission, we have $R = B = G$, and so $Y = 1.0$. The last column in Table D.1 shows, however, that the transmission of a fully saturated color requires larger instantaneous peaks, which may be as high as 78% over that required for monochrome. Because the peak modulation level is already 70% for monochrome, this is an intolerably large peak overload at the transmitter and it is necessary to reduce the amplitude of the color difference signals.

There are two color difference signals so we can introduce two independent scaling constants, a_1 and a_2, to reduce the peak modulation. For monochrome transmissions, $Y = 1.0$ and the maximum modulation index is 70%. A reasonable peak overload due to color differences is taken to be 33% (i.e., modulation index of 93%). Thus we can write the following equation for the maximum peak modulation level:

$$Y + \sqrt{a_1^2(R-Y)^2 + a_2^2(B-Y)^2} = 1.33\,. \tag{D.6}$$

From Table D.1, it can be seen that the two saturated colors resulting in the largest possible overloads are yellow and cyan. Substituting the values for yellow and cyan successively into Eq. (D.6) yields two equations in the two unknown scaling constants a_1, a_2:

$$0.89 + \sqrt{a_1^2(0.11)^2 + a_2^2(0.89)^2} = 1.33\,, \tag{D.7}$$

$$0.70 + \sqrt{a_1^2(0.70)^2 + a_2^2(0.30)^2} = 1.33\,. \tag{D.8}$$

Solving Eqs. (D.7) and (D.8) simultaneously yields $a_1 = 0.877$, $a_2 = 0.493$. The total video signal, including the corrected chrominance signal, which is used to amplitude modulate the transmitter carrier, is then[†]

$$E_T = E_Y + 0.877(E_R - E_Y)\cos\omega_s t + 0.493(E_B - E_Y)\sin\omega_s t\,. \tag{D.9}$$

The corrected chrominance values are listed in Table D.2. These results can also be portrayed by constructing a phasor diagram of the phase relationships at the color subcarrier frequency ω_s, as shown in Fig. D.7.

[†] At the receiver, the $(R - Y)$ component is weighted—relative to the luminance signal—by $1/0.877 = 1.14$, and the $(B - Y)$ component by $1/0.493 = 2.03$.

TABLE D.2 Phase Relationships for Corrected Chrominance Signal

Color	Components $0.877(R - Y)$	$0.493(B - Y)$	Magnitude/Phase [using Eq. (D.5)]
Blue	−0.0965	0.439	$0.450\ \underline{/347°}$
Red	0.614	−0.148	$0.631\ \underline{/103.5°}$
Magenta	0.517	0.291	$0.594\ \underline{/60.7°}$
Green	−0.517	−0.291	$0.594\ \underline{/240.7°}$
Cyan	−0.614	0.148	$0.631\ \underline{/283.5°}$
Yellow	0.0965	−0.439	$0.450\ \underline{/167°}$

A question that arises is how much bandwidth to allow for the chrominance signals. Fortunately the color resolution of the human eye is less than for black-and-white. In fact, studies of human vision have shown that it is not even the same for all color combinations. Finer chrominance detail can be resolved in orange and cyan than in green and magenta. To take advantage of this, the chrominance subcarrier reference can be advanced (at the transmitter) in Fig. D.7 by 33°. The new axes so formed are called the I and Q axes and are shown in Fig. D.7. For the Q signal, a comparatively narrow bandwidth of 0.5 MHz is found to be adequate. DSB is used, resulting in a 1-MHz modulated bandwidth. A much wider bandwidth is needed for

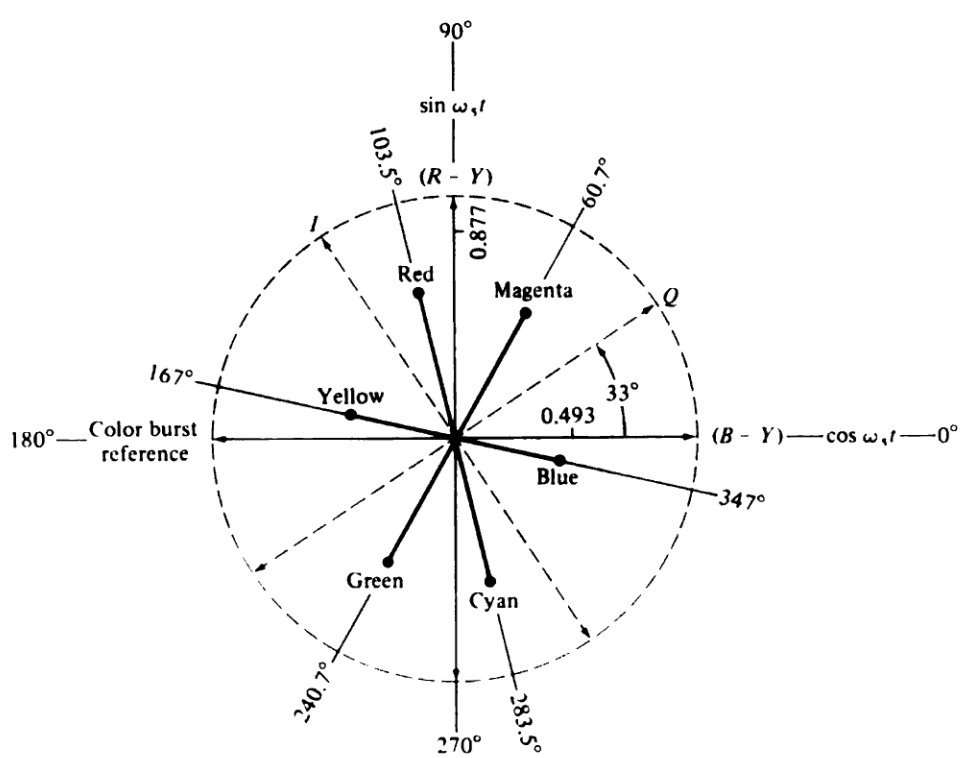


Figure D.7 Phasor diagram of chrominance phase relationships.

the I signal. In order to accommodate the wider bandwidth, vestigial sideband modulation is used to restrict the I signal bandwidth to about 500 kHz above the subcarrier frequency. A simplified spectral diagram of a color television transmission is shown in Fig. D.8.

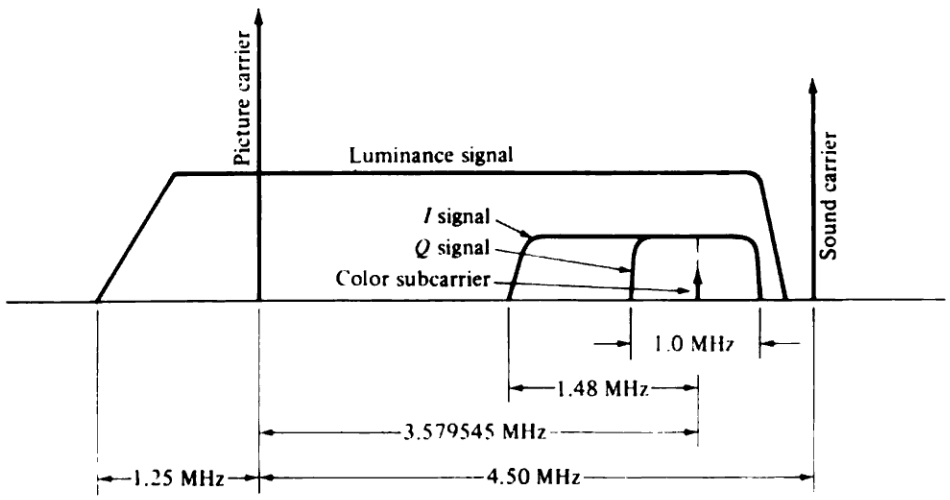


Figure D.8 Color television spectrum showing I and Q spectra.

At the receiver, demodulation along the I and Q axes should be used to take advantage of the wider bandwidth of the I signal. However, this makes receiver design more difficult and generally this method of demodulation is not used. As noted earlier, it is simpler to demodulate along the $(R - Y)$ and $(B - Y)$ axes and then derive the $(G - Y)$ component by matrixing. The two demodulators use the same bandwidth so that there is some loss in color definition which is most noticeable in the orange and cyan hues (see Fig. D.7).

The next question that arises is that of the choice of the color subcarrier frequency. A subcarrier at a frequency that is too high would result in crowding the color sidebands too near the 4-MHz video bandwidth limit from the picture carrier. The possibility of heterodyne products between the audio carrier and the color-signal sidebands is also increased with a high subcarrier frequency. On the other hand, if the subcarrier frequency is too low, interference might result from heterodyne products between the main picture carrier and the color-signal sidebands. (In practice, the picture carrier is on the order of about 10 dB larger than the audio carrier, favoring a choice near the high end of the video bandwidth.) The spectral components in a typical video signal are much stronger toward the lower frequencies; therefore, it is easier to insert the chrominance information toward the high end of the video bandwidth and remain compatible with monochrome receivers. A reasonable choice based on these considerations is in the vicinity of 3.5 MHz. To understand the choice of the specific frequency chosen (3.579545 MHz) requires a more careful consideration first of the spectrum of monochrome television transmissions.

A photograph of a spectrum analyzer display of a commercial television transmission is shown in Fig. D.9(a). An expanded portion of the spectrum near the picture carrier is shown in Fig. D.9(b). Note that as a result of the horizontal sweep sampling, the luminance spectral information tends to cluster about the harmonics of the horizontal sweep frequency. One can observe these very definite sidebands introduced by the sampling process at all multiples of the horizontal sweep frequency.† The reference or zeroeth harmonic is the main picture carrier.

The chrominance information is also sampled at the horizontal sweep rate, but its reference is at the color subcarrier frequency. Therefore, we can choose this subcarrier frequency in such a way that the spectra of the chrominance signals lie exactly between the spectra of the luminance signal. In this way it is possible to transmit both

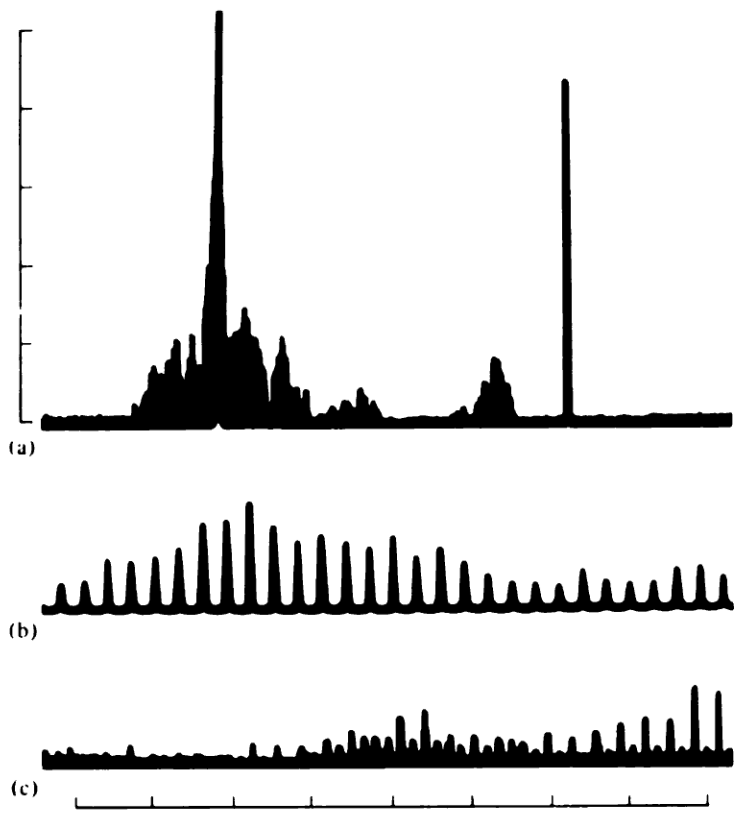


Figure D.9 Spectrum analyzer display of a commercial television transmission: (a) 1 MHz per division; (b) 50 kHz per division near the picture carrier; (c) 50 kHz per division near the color subcarrier (vertical scale is 10 dB per division).

† If the picture were absolutely stationary, the sidebands about each multiple of the 15.75-kHz horizontal sweep frequency should in turn have harmonic lines spaced at 60 Hz (the vertical sweep frequency). In practice, the picture information generally varies too much to resolve these lines.

the luminance and the two chrominance signals within the same bandwidth as is normally used for monochrome television transmissions.

In order to position the chrominance sidebands between the luminance sidebands, it is necessary to choose the color subcarrier frequency at an integer-plus-one-half times the horizontal sweep frequency. In other words, the color subcarrier frequency must be an odd multiple of one-half the horizontal sweep rate. Another consideration is that we would like to have the sound carrier (at 4.5 MHz) positioned between the color sidebands to avoid possible low-frequency heterodyne effects. Taking the 571st harmonic, we get

$$\frac{15{,}750}{2}(571) = 4{,}496{,}625 \text{ Hz},$$

and for $n = 573$,

$$\frac{15{,}750}{2}(573) = 4{,}512{,}375 \text{ Hz}.$$

Thus the 4.50 MHz audio carrier is not centered, as desired, between the color sidebands. We have two possible choices to correct this situation. First, we could shift the audio carrier frequency up slightly to

$$\frac{4{,}496{,}625 + 4{,}512{,}375}{2} = 4{,}504{,}500 \text{ Hz}.$$

However, this 4.5 kHz shift in the audio FM spectrum would not be compatible with existing monochrome receivers.† The second choice is to lower the horizontal sweep frequency slightly. The required horizontal sweep frequency, f_h, must then satisfy the equation

$$\frac{4{,}500{,}000 + f_h/2}{573} = f_h/2, \tag{D.10}$$

from which

$$f_h = 2\frac{4{,}500{,}000}{572} = 15{,}734.266 \text{ Hz}. \tag{D.11}$$

This new horizontal sweep frequency for color transmission is within 0.1% of the monochrome value (15,750.) and is quite satisfactory with existing receiver designs. The corresponding new field frequency is

$$\frac{2}{525}(15{,}734.266) = 59.94 \text{ Hz}. \tag{D.12}$$

Both the 59.94-Hz field rate and the 15,734.266-Hz horizontal sweep rate are sufficiently close to the 60-Hz and 15,750-Hz rates used in monochrome transmissions to

† A peak frequency deviation of 25 kHz is used in television audio transmissions and the shift of 4.5 kHz is not negligible.

permit synchronization in both types of receivers. However, the new rates are no longer in synchronism with the 60-Hz power line frequency. Thus any power line pickup in the video will appear to "roll" slowly through the picture vertically with about a 17-sec (that is, 1/.06) period when the receiver is tuned to a color broadcast. This is judged to be a minor disadvantage for the sake of compatibility and one that can be minimized with good design procedures.

The color subcarrier is chosen to be the 455th harmonic of one-half the horizontal sweep rate, or

$$(455)\left(\frac{15{,}734.266}{2}\right) = 3{,}579{,}545 \text{ Hz}. \tag{D.13}$$

Because quadrature multiplexing is used for the color information, it is essential that the synchronous detector in the receiver be exactly in phase and frequency with the transmitter. This is accomplished by sending eight cycles—known as the "color burst"—of the color subcarrier during the blanking interval following the horizontal sync pulse (known as the "back porch"). This is illustrated in Fig. D.10.

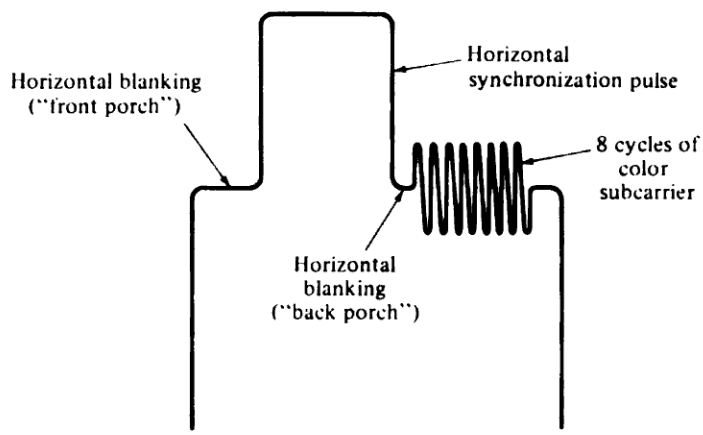


Figure D.10 Color subcarrier burst.

In the color television receiver the eight cycles of the color subcarrier are selected by a delayed gating circuit triggered by the horizontal sync pulse. One method of synchronization is to compare the phase of the color burst with the phase of a local oscillator and correct it accordingly. An alternative approach is to use a very high-Q narrowband filter such that oscillations are started by the color burst and then will continue for the duration of one horizontal sweep. The color burst is not transmitted during the vertical retrace or during any B/W transmission. A "color killer" circuit in the color television receiver is used to detect the absence of the color burst and disables the color amplifiers so that possible color leakage is not displayed in the picture.

The stability of the color subcarrier frequency must be carefully controlled, particularly over the period of one horizontal line. Commercial broadcast station tolerances for the 3.579545-MHz frequency are ±0.0003%, or about ±10 Hz. The

horizontal and vertical sweep frequencies at the transmitter are obtained by appropriate frequency dividers from the color subcarrier frequency.

Up to this point, we have tacitly assumed that the cathode ray tube in the receiver operates as a linear device so that the light intensity output is proportional to the amplitude of the input signal voltage. This is not true in practice and a more realistic relationship is

$$I = kE^{\gamma}, \tag{D.14}$$

where I represents the light intensity output, k is an arbitrary constant, E is the signal voltage, and γ (gamma) is a constant—about 2.2 to 2.5. Thus the light output increases more rapidly than the input signal voltage. If uncorrected, gradations in the highlights in the picture would be expanded while gradations in the shadows would be compressed. To compensate for this, a nonlinear amplifier that has an opposite transfer characteristic dependency is used at the transmitter to effect the needed gamma correction.

Gamma correction is quite straightforward for monochrome transmissions. An amplifier whose input/output transfer characteristic is of the form

$$E_o = k_1 E_i^{1/\gamma} \tag{D.15}$$

is used to amplify the camera signal before modulation at the transmitter. For color transmissions, the three primary color signals are gamma-corrected, using Eq. (D.15), and then coded into the luminance and chrominance signals. This provides for the correct picture dependency in each color. The luminance signal is then given by

$$E_Y = 0.30E_R^{1/\gamma} + 0.59E_G^{1/\gamma} + 0.11E_B^{1/\gamma} \leq (0.30E_R + 0.59E_G + 0.11E_B)^{1/\gamma}. \tag{D.16}$$

The equality in Eq. (D.16) holds when $E_R = E_G = E_B$; i.e., when a monochromatic (B/W) picture is being transmitted. For highly saturated colors, on the other hand, the luminance signal is smaller than it should be and yet the complete color information is present at the receiver. What is happening, of course, is that some of the luminance information for highly saturated colors is actually being sent in the chrominance signals when gamma correction is applied. This results in some lack of definition in scenes with highly saturated colors and some distortion on sudden changes in color as a result of a partial transmission of the luminance information through the narrower-band chrominance system. It is not judged to be a serious drawback in practice, however.

Because there is this cross-coupling effect between the luminance and chrominance channels when gamma correction is applied, it now makes a difference at which reference angle the color subcarrier burst is sent. Of the choices: 0, ±90°, 180°, the last one yields the least luminance cross-coupling and is chosen as the reference angle (see Fig. D.7).

In summary, the NTSC color television system produces color pictures of acceptable definition and hue to the human eye within an overall bandwidth allocation of 6 MHz. It provides good compatibility with monochrome reception and it is reasonably simple to transmit and receive.

D.4 OTHER COLOR TELEVISION SYSTEMS

The NTSC color television system described above was the first system to be adopted for commercial broadcasting on a permanent basis. Other color systems have been devised since the NTSC system was placed in operation. Basically, these systems have many similarities to the NTSC system but they attempt to overcome some of its drawbacks. Two of these systems are described briefly here.

In the PAL (Phase Alternating Line) color television system, the color difference signals $(R - Y)$ and $(B - Y)$ each are transmitted within a bandwidth of 1 MHz using quadrature multiplexing. The $(R - Y)$ channel, however, is reversed in polarity on successive lines at the transmitter. The correct phase is restored by a synchronous polarity reversing switch in the receiver. By this means, effects of differential phase errors occurring in transmission (or in video recording) may be partially corrected by an averaging effect over several lines. (Recall from Fig. D.7 that the definition of the eye along the $(R - Y)$ axis is better than along the $(B - Y)$ axis.) The phase of the color burst alternates between 135° and 225° (that is, 180° ± 45°) to synchronize the polarity reversing switch. The required receiver design for the PAL system is therefore more complicated (and therefore more expensive), with some advantage in better performance. A simplified spectral diagram of a British color television transmission using a 625-line PAL system is shown in Fig. D.11.

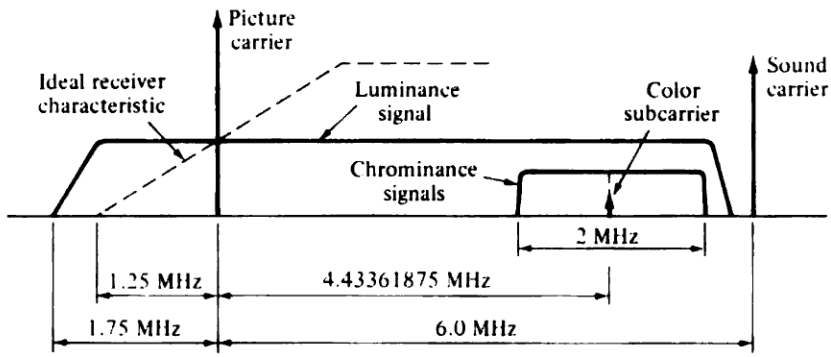


Figure D.11 Simplified spectrum of the British 625-line PAL system.

Another color system is used in France—the Séquential Couleur à Mémoire (SECAM) system. In the SECAM system, the color difference signals $(R - Y)$ and $(B - Y)$ are sent separately on alternate lines, i.e., the $(R - Y)$ signal for one line is transmitted with the line while the $(B - Y)$ signal is discarded, and on the next line the $(B - Y)$ signal is transmitted while the $(R - Y)$ signal is discarded. Frequency modulation of the color subcarrier is used to transmit the color information.

In the SECAM receiver, a delay line and switch arrangement is used to allow a comparison of the previous line of color information with the present one. Thus the color resolution in the vertical dimension is reduced by one-half but the quadrature

multiplexing and its possible cross-coupling between color signals is avoided. The use of FM for the color information makes the SECAM system immune — in contrast to the NTSC and PAL systems — to color distortion arising from amplitude variations. The major drawback of the SECAM system is that the use of FM for the color information tends to cause some video interference pattern effects in monochrome transmissions resulting in a poorer degree of compatibility.

D.5 STEREO TV

Good-quality monaural sound is provided by the FCC television standards, which require an audio channel to be capable of a 50-Hz to 15-kHz frequency response and a 60-dB signal-to-noise ratio. As a result of increasing consumer interest in high-quality stereo audio, the FCC approved a proposed Zenith/dbx system in 1984 for the addition of stereo audio to commercial television transmissions in the United States. Because this system actually provides several additional channels for the transmission of information on the audio carrier of the television signal, it is also known as Multichannel Television Sound (MTS). The MTS system retains compatibility with monaural television receivers.

The principles of MTS are similar in many respects to those used in FM stereo. The sum channel (L + R) occupies the frequency range 50 Hz to 15 kHz to provide compatibility with monaural systems. The difference channel (L − R) is modulated on a 31.468-kHz subcarrier using DSB-SC, and a pilot tone is transmitted at 15.734 kHz (the horizontal scan frequency — abbreviated by H). The pilot tone furnishes a signal that can be detected and doubled in frequency to supply the 2H subcarrier for the L − R channel. It tells the receiver that a stereo signal is present and can be used to control a stereo indicator light. A diagram of the MTS composite audio spectrum is shown in Fig. D.12. This composite audio is applied to the input of the frequency modulator to produce the audio portion of the television broadcast signal.

A portion of the composite audio spectrum has been allocated for transmission of second audio program (SAP) material. The SAP channel is capable of handling an audio frequency range from 50 Hz to 10 kHz and it is often used for second-language programming. It is frequency modulated using a 5H (78.67-kHz) subcarrier and a 10-kHz peak frequency deviation. There is provision in MTS for a fourth channel, called the professional channel. This channel uses a subcarrier of 6.5H (102.27 kHz) and is used primarily by television stations for program coordination purposes.

The L + R channel uses the standard 75-μsec (2.1-kHz) passive preemphasis/deemphasis to maintain compatability with monaural television receivers. To help overcome effects of a noise spectral density that rises with frequency, the L − R channel uses twice the modulation level of the L + R channel. Both the L − R and the SAP channels use 390-μsec (0.41-kHz) passive preemphasis/deemphasis and both are encoded by dbx® noise reduction circuits (cf. Appendix F).

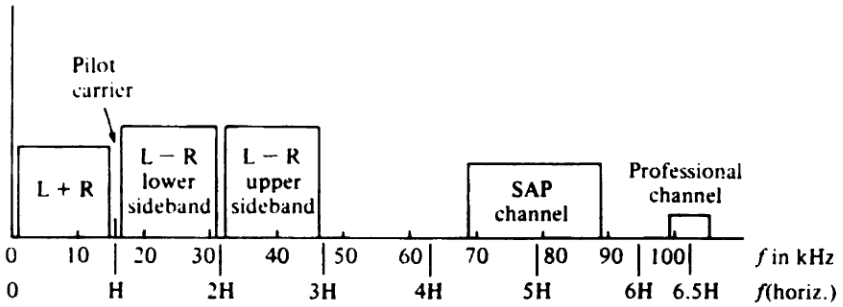


Figure D.12 The MTS composite audio spectrum for stereo television.

The MTS system provides good stereo sound reproduction within the compatability constraints of the FCC television standards. It provides additional audio channels at reduced modulation levels to avoid interference beat-frequency terms with the video signal.

D.6 HDTV†

With development of higher resolution video color displays, there have been recent interests in development of television systems with improved video resolution. Such systems have been termed high-definition television (HDTV). At the time of this writing, 23 proposals for HDTV and other advanced television transmission systems have been submitted to the FCC for evaluation and testing. A FCC decision on broadcast transmission standards for terrestrial systems is expected by 1992.

A goal of HDTV is to offer improved picture definition on the order of 1000 scan lines with 60 fields per sec,‡ or about double that achievable using NTSC standards. Another goal is to offer a "wide-screen" display, one that will require an aspect ratio closer to 5:3 or 16:9 in comparison with the 4:3 ratio in current use. A quick estimate shows that a system that uses twice the number of picture elements (pixels) in each spatial dimension and an aspect ratio of 5:3 instead of 4:3 would require a video bandwidth increase, assuming no other changes, on the order of $(5/4)2^2 = 5$. Because the video already uses most of the 6-MHz channel allocation, this magnitude increase in RF bandwidth is clearly out of bounds. Proposed signal processing methods to try to circumvent these limitations include (1) taking advantage of redundancies in the present system; (2) using time-serial transmission and possible receiver

† References for this section include L. Feldman, "High Definition TV," *Radio Electronics* (February 1989): 35; "High-Definition Television Update," *IEEE Spectrum*, vol. 25 (April 1988): 56–62; Special Issue on Advanced Television Systems, *IEEE Transactions on Consumer Electronics*, CE-34 (February 1988); Special Issue on High Definition Television, *IEEE Transactions on Broadcasting*, BC-33 (December 1987).

‡ For compatibility with NTSC, most of the systems proposed use 59.94 fields per sec.

data storage; and (3) using quadrature multiplexing to a greater extent. Some proposed systems use a combination of these methods; another way to group the proposed systems is whether they are (1) receiver compatible; (2) partially compatible; or (3) not compatible with existing NTSC standards.

Systems that are compatible should display a conventional picture when received by a standard (NTSC) receiver, and require no additional bandwidth beyond the 6 MHz presently assigned to TV broadcast stations. They should offer improved picture definition, improved color rendition and high-quality stereophonic audio when receiving HDTV transmissions. The same requirements hold for partially compatible systems, with the exception that additional bandwidth is needed. Systems that are not compatible would require completely altered transmission and receiver standards. Some examples of proposals in these various categories are described briefly.

An example of a quadrature-multiplexed compatible HDTV system is the Advanced Compatible TV (ACTV-1) system developed by RCA, GE, NBC, and The David Sarnoff Research Center. This system would transmit standard NTSC video, plus the low-frequency components of a time-compressed "side-panel" at each end of the horizontal trace (which may not show on a conventional NTSC receiver). Side-panel high-frequency components are time-expanded and band-limited to about 1 MHz. High-frequency luminance information, 5.0–6.2 MHz, is shifted down to baseband using a 5.0-MHz mixer and then low-pass filtered to 1.2 MHz. These two signals quadrature-modulate a 3.108-MHz subcarrier (395th harmonic of one-half the horizontal sweep rate). A fourth signal carrying vertical-temporal (V-T) luminance information is band-limited to 0.75 MHz and modulated in quadrature with the (main) picture carrier. This V-T signal is needed to smooth the vertical luminance in the reconstructed HDTV picture. The ACTV-1 transmits 1050 interlaced lines per frame with resolution exceeding 400 lines in both spatial dimensions, and is compatible with the existing NTSC standards.

The partially compatible systems use a 6-MHz channel allocation for the standard NTSC transmission, but need additional bandwidth for the HDTV transmission. The Philips High Definition System for North America (HDS-NA) system transmits a NTSC signal in a standard 6-MHz channel and uses a half-channel (3 MHz) to transmit information digitally that is needed for the HDTV video. A possible advantage in using the half-channel is that the augmenting channel could be shared with another HDTV station. The signal power in the augmenting channel can be at a lower level than that needed for the main NTSC signal. Channel allocations are often staggered in urban areas to avoid interference, and a possible use of these adjacent buffer channels would be for these augmenting transmissions.

The Zenith Spectrum-Compatible HDTV (SC-HDTV) proposed system is an example of one that is compatible with current 6-MHz TV channel allocations but does not use the NTSC standards. In this system, video at frequencies above 200 kHz is transmitted in an analog format, while those signals below 200 kHz are sent digitally. Video encoding divides a 30-MHz input luminance signal into four components; two are band-limited to the lowest third of the spectrum (the second uses a differences encoding), one is band-limited to the middle third, and one to the top third of the spec-

trum. The first is sent at full frame rate (59.94) while the others are time-multiplexed at one-fifth frame rate (11.988/sec). Thus, stationary video is refreshed at the 11.988 frame rate but change (motion) is sent at the 59.94 frame rate. This matches the resolving capability of the human eye, which is considerable less for objects in motion than it is for stationary scenes. Two color difference signals ($R - Y, B - Y$) are transmitted at one-third resolution and multiplexed at the one-fifth frame rate. These six video components are time multiplexed into two 3-MHz signals that modulate in-phase and quadrature carriers at the center of the 6-MHz channel allocation, using DSB-SC. A pilot carrier is reinserted to aid carrier regeneration in the receiver.

The digital data in the Zenith proposed system consists of low-frequency video components, synchronization information, high-quality audio, captioning, etc. Error-correcting coding is used, and this digital data is transmitted during the 22–23 lines in the vertical blanking interval at 22.4 Mbps using 16-state QAM on the same suppressed carriers as the analog video. Because this system attempts to optimize use of the spectrum available (no carriers and a relatively flat spectral density) the transmitter power required for a given performance level can be reduced, and thus this system could possibly use channels currently restricted as buffer channels between channels allocated for active use in a given area.

An example of a system that is not compatible with either the existing NTSC standards or the 6-MHz channel allocations is the Multiple Sub-Nyquist Encoding (MUSE) system developed by NHK (the Japan Broadcasting Corporation). This system uses time-serial transmission with a four-field sequence, first transmitting every other pixel of every other line, then the remaining samples of those lines, then every other pixel of alternate lines, etc. It uses two resolutions for the video display, dependent on detected presence of motion. If there is motion, the resolution is that of NTSC systems with 4 MHz of bandwidth. For no motion, resolution is that of a 8.1-MHz (baseband) video bandwidth. The MUSE system uses 1125 scanning lines per frame with 2:1 interlaced scanning, 60 frames per sec, a 16:9 aspect ratio, and requires a bandwidth of approximately 10 MHz. NHK is scheduled to begin some HDTV satellite broadcasting in 1990 using the MUSE system.

This brief discussion gives a sampling of several signal processing systems that have been proposed to the FCC. Next will follow the evaluation and testing phase under carefully specified and maintained conditions. Relative costs, both in terms of economics and bandwidth and compatibility will have to be weighed versus performance and benefits provided. Spectrum availability is one of the big issues for terrestrial broadcast systems in the United States. It is possible that eventually several different standards will evolve, depending on use of terrestrial broadcast, cable, or satellite systems.

Selected References for Further Reading

1. M. Mandl. *Principles of Electronic Communications*. Englewood Cliffs, N.J.: Prentice Hall, 1973, Chapters 8, 9.

2. H. Stark, F. B. Tuteur, and J. B. Anderson. *Modern Electrical Communications: Analog, Digital, and Optical Systems*. Second ed. Englewood Cliffs, N.J.: Prentice Hall, 1988, Chapter 6.
3. C. N. Herrick. *Color Television: Theory and Servicing*. Reston, Va.: Reston Publishing Co., 1976.
4. H. V. Sims. *Principles of PAL Colour Television*. London: Iliffe Books Ltd., 1969.
5. G. N. Patchett. *Colour Television*. London: Norman Price (Publishers) Ltd., 1967.
6. K. B. Benson (ed.). *Television Engineering Handbook*. New York: McGraw-Hill, 1986.

APPENDIX

E Telephone Channels

A basic voiceband telephone channel passes frequency components in the range 0.3–3 kHz with tolerable distortion to the human ear. When voiceband channels are used for data transmission, however, some problems may arise as a result of nonuniform magnitude frequency response and a phase delay that is not linear with frequency. To alleviate this, five types of "conditioned" lines are made available in the United States for data services. Specifications for lines maintained by AT & T are given in Table E.1.[†] (Envelope delay distortion will be defined later.)

TABLE E.1 Some Typical Telephone Parameter Limits[a]

	Attenuation Distortion (Frequency Response) Relative to 1 kHz		Envelope Delay Distortion Relative to Midband	
Channel Conditioning	Frequency Range (Hz)	Variation (dB)	Frequency Range (Hz)	Variation (μ sec)
Basic	500–2500	−2 /+8	800–2600	1750
	300–3000	−3 /+12		
C1	1000–2400	−1 /+3	1000–2400	1000
	300–2700	−2 /+6	800–2600	1750
	300–3000	−3 /+12		
C2	500–2800	−1 /+3	1000–2600	500
	300–3000	−2 /+6	600–2600	300
			500–2800	3000
C3	500–2800	−0.5/+1.5	1000–2600	110
	300–3000	−0.8/+3	600–2600	300
			500–2800	650
C4	500–3000	−2 /+3	1000–2600	300
	300–3200	−2 /+6	800–2800	500
			600–3000	1500
			500–3000	3000
C5	500–2800	−0.5/+1.5	1000–2600	100
	300–3000	−1 /+3	600–2600	300
			500–2800	600

[a] From *Data Communications Using Voiceband Private Channels*, Bell System Publication 41004, October 1973, © American Telephone and Telegraph Company 1973. Reprinted with permission.

[†] "AT & T" is an abbreviation used for the American Telephone and Telegraph Company.

Most long-distance telephone systems use multiple SSB-SC channels that are frequency-multiplexed on a high-frequency (microwave) carrier. The subcarriers are not transmitted. The subcarrier which is reinserted for correct demodulation is not frequency or phase locked but is kept within ±5 Hz of the correct subcarrier frequency.

It is impractical to measure the phase-shift characteristic of a SSB channel because of the lack of phase lock. Therefore a narrowband sinusoidal AM waveform is sent through the channel and the difference between the phase shifts on the upper and lower sideband is detected as a function of the carrier frequency. For low modulating frequencies ($83\frac{1}{3}$ Hz is used by AT & T) this phase difference is approximately the derivative of the phase delay characteristic. The phase delay characteristic can be obtained, except for an arbitrary constant, by integration of the envelope delay. *Envelope delay distortion* is defined to be the maximum difference in the envelope delay over a given band of frequencies.

Delay distortion is quite tolerable for voice communications, but it becomes a serious problem for pulse-type transmissions because not all frequency components arrive at the same time. For example, if a 1-kHz square wave were applied to a channel having appreciable delay distortion, the fundamental and the third harmonic (all higher harmonics are severely attenuated) could undergo very different phase shifts. A typical resulting waveform is shown in Fig. E.1.

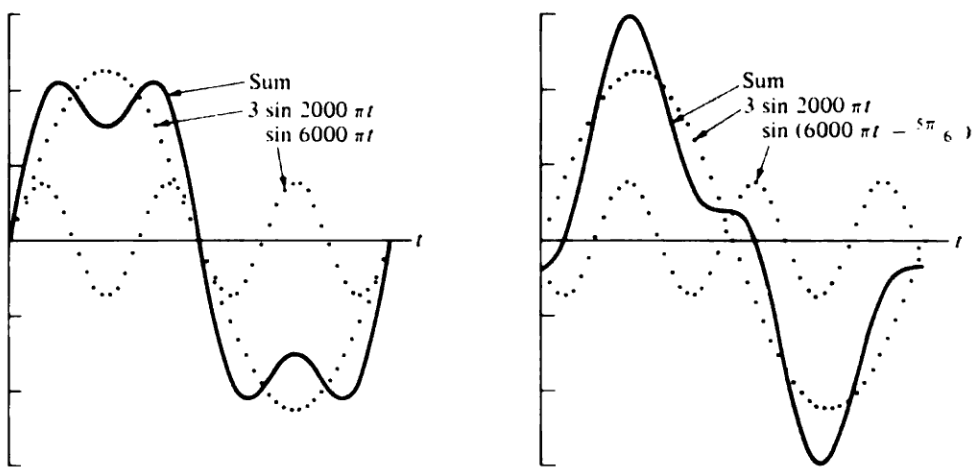


Figure E.1 Input and output waveforms for a channel having delay distortion.

These effects become important in facsimile transmission, the transmission of two-dimensional imagery over commercial telephone systems. Examples include transmission of notices and invoices, weather forecast charts, satellite weather pictures, etc.

Facsimile (fax) transmission uses some of the same basic scanning techniques that are used in commercial television. Unlike television, it requires no vertical retrace, and relative allowances for horizontal retrace are negligible. As a result of some of the characteristics of telephone systems, however, additional problems may arise.

In order to make good use of the limited bandwidth available, analog facsimile transmissions are often amplitude-modulated, using a 2400-Hz carrier and a modulating bandwidth of 1600 Hz. Vestigial sideband filtering about the 2400-Hz carrier is then used to restrict the bandwidth to the range in which the delay distortion is tolerable. Line synchronization is sent by transmitting a low-level carrier (5% modulation) and then modulating the carrier with a step function to increase the modulation index to 100%. As a result of lack of phase coherence, the detected line synchronization can vary appreciably. Note that in commercial television the carrier frequency is high enough that this is not a problem of concern.

As a result of the effects of delay distortion, facsimile systems are increasingly making use of digital encoding methods and frequency-shift (FSK) modulation. Some narrowband transmissions (e.g., EEG waveforms) are sent in analog form using frequency modulation (FM).

APPENDIX F Some Commercial Preemphasis/ Deemphasis Systems

The use of passive high-frequency preemphasis/deemphasis has been practiced in audio recording and commercial broadcasting for years. Recently much more sophisticated preemphasis/deemphasis techniques have been applied successfully to the audio tape-recording field. Some are being proposed for broadcast use (the Dolby-B system is already used by many FM stations).

The three major types of new preemphasis/deemphasis systems in current use are described briefly. All three use a combination of filtering and dynamic-range compression to reduce the effects of the noise. Because the noise effects are most noticeable when the signal level is lowest, all three attempt to raise (boost) the low signal levels.

The Dolby-A system (Dolby Labs., London) divides the audio spectrum into four frequency bands and then applies a 10 to 15 dB boost to individual bands when the audio signal level falls below a preset level. This level is about 45 dB below a standard signal level known as the "Dolby level" and is set with reference to the flux level on the recording tape. The boosted signal is added to the original signal and then recorded. Assuming that the playback and record levels are properly set, the Dolby-A system can increase the net signal-to-noise (S/N) ratio 10 to 15 dB with no perceptible effect on the program material. Block diagrams of the Dolby-A system are shown in Figs. F.1 and F.2. A disadvantage of the Dolby system is that it has a limited dynamic range and cannot handle large orchestral peaks without some amplitude limiting ("clipping"). It is also fairly expensive for widespread use in domestic equipment.

A cheaper version, the Dolby-B system, is designed for domestic tape-recording equipment. To keep the costs down, it replaces splitting the audio spectrum into four bands with a high-pass band above 600 Hz. A graph of the preemphasis characteristic

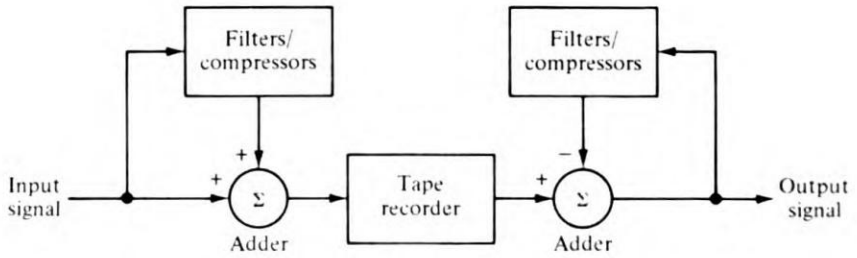


Figure F.1 Block diagram of the Dolby system.

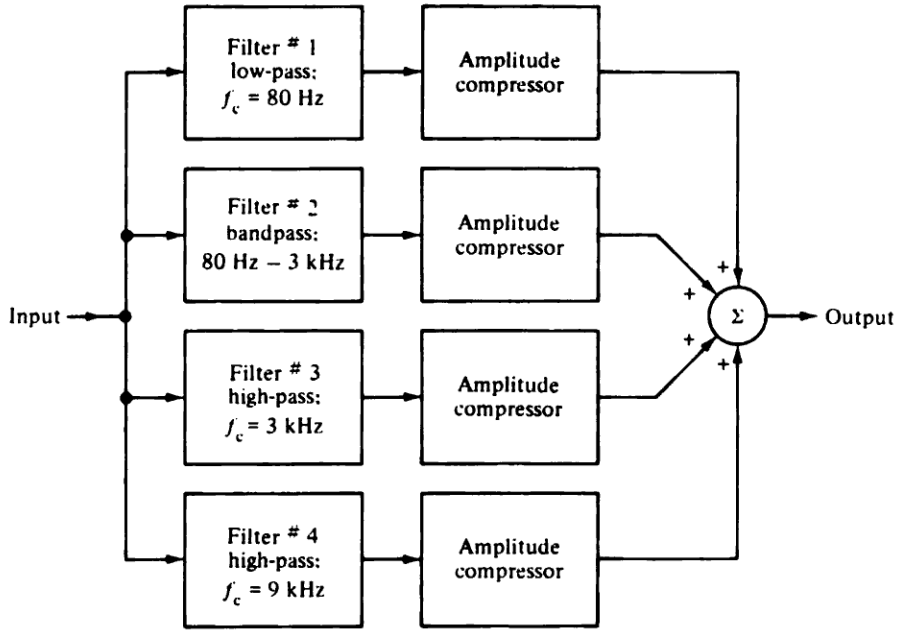


Figure F.2 Filters/compressors for the Dolby-A system.

of a Dolby-B system is shown in Fig. F.3. Its use in home-quality tape recorders is rapidly becoming popular with the availability of IC's. It is also being used for FM transmissions with a 25-μsec time-constant (6.4 kHz), high-pass band. This allows some S/N improvement while retaining compatibility with the 75 μsec fixed preemphasis used in FM transmissions.

The DBX system (DBX, Inc., Waltham, Mass.) treats the audio spectrum as a whole and uses high-frequency preemphasis above 1.6 kHz with a maximum high-frequency boost of 20 dB. A voltage-controlled amplifier with a wide dynamic range

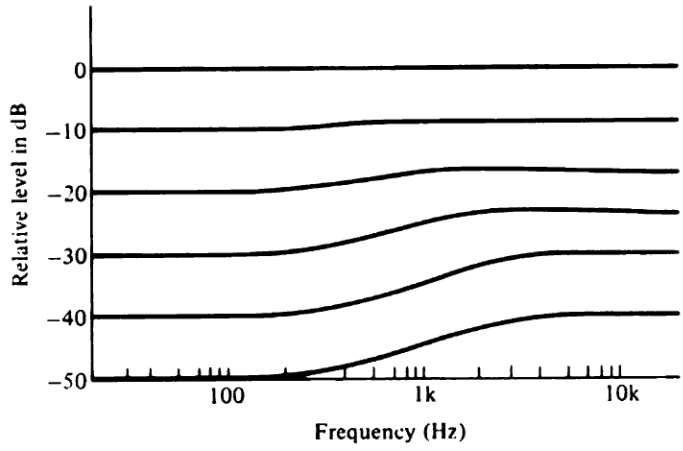


Figure F.3 Preemphasis characteristic of the Dolby-B system.

(130 dB) is used to amplify the signal. The gain of this amplifier is controlled by a rms detector. The overall effect is to detect the level of the signal and compress it by a 2:1 ratio. On playback, the signal is expanded to its original dynamic range and the high-frequency boost is removed. The DBX system offers 20–30 dB S/N improvement and can handle 26-dB signal peaks without distortion. A disadvantage of the DBX system is that the compression is governed by the largest input anywhere in the frequency band.

The Burwen system (Burwen Labs., Inc., Lexington, Mass.) is more elaborate and more expensive than either of the other two systems described above. It is similar in some respects to the DBX system but differs in that it uses precision peak signal detectors to control a wider compressor gain variation and uses low- as well as high-frequency preemphasis. The frequency-weighting preemphasis characteristics are dependent on the recording speed selected. The compression ratio is 3:1 and the net S/N improvement is approximately 50 dB. The Burwen system is expensive and its use is presently limited to high-quality studio use. A block diagram of the Burwen system is shown in Fig. F.4.

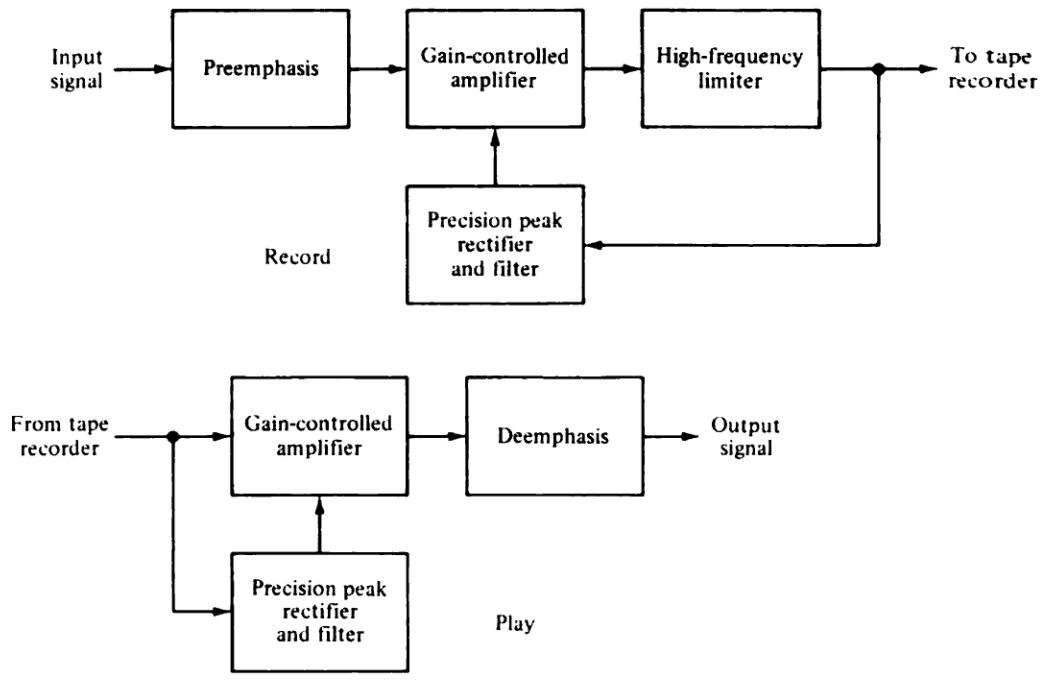


Figure F.4 Block diagram of the Burwen system.

Selected References for Further Reading

K. B. Benson (ed.). *Audio Engineering Handbook*. New York: McGraw-Hill, 1988.

APPENDIX

G A Table of Bessel Functions

Bessel Functions of the First Kind, $J_n(x)$

x	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}
0.0	1.00										
.2	.99	.10									
.4	.96	.20	.02								
.6	.91	.29	.04								
.8	.85	.37	.08	.01							
1.0	.77	.44	.11	.02							
.2	.67	.50	.16	.03	.01⁻						
.4	.57	.54	.21	.05	.01⁻						
.6	.46	.57	.26	.07	.01						
.8	.34	.58	.31	.10	.02						
2.0	.22	.58	.35	.13	.03	.01⁻					
.2	.11	.56	.40	.16	.05	.01					
.4	.00	.52	.43	.20	.06	.02					
.6	−.10	.47	.46	.24	.08	.02	.01⁻				
.8	−.19	.41	.48	.27	.11	.03	.01⁻				
3.0	−.26	.34	.49	.31	.13	.04	.01				
.2	−.32	.26	.48	.34	.16	.06	.02				
.4	−.36	.18	.47	.37	.19	.07	.02	.01⁻			
.6	−.39	.10	.44	.40	.22	.09	.03	.01⁻			
.8	−.40	.01	.41	.42	.25	.11	.04	.01			
4.0	−.40	−.07	.36	.43	.28	.13	.05	.02			
.2	−.38	−.14	.31	.43	.31	.16	.06	.02	.01⁻		
.4	−.34	−.20	.25	.43	.34	.18	.08	.03	.01⁻		
.6	−.30	−.26	.18	.42	.36	.21	.09	.03	.01		
.8	−.24	−.30	.12	.40	.38	.23	.11	.04	.01		
5.0	−.18	−.33	.05	.36	.39	.26	.13	.05	.02	.01⁻	
.2	−.11	−.34	−.02	.33	.40	.29	.15	.07	.02	.01⁻	
.4	−.04	−.35	−.09	.28	.40	.31	.18	.08	.03	.01⁻	
.6	.03	−.33	−.15	.23	.39	.33	.20	.09	.04	.01	
.8	.09	−.31	−.20	.17	.38	.35	.22	.11	.05	.02	.01⁻

Bessel Functions of the First Kind, $J_n(x)$

x	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}
6.0	.15	−.28	−.24	.11	.36	.36	.25	.13	.06	.02	.01⁻
.2	.20	−.23	−.28	.05	.33	.37	.27	.15	.07	.03	.01⁻
.4	.24	−.18	−.30	−.01	.29	.37	.29	.17	.08	.03	.01
.6	.27	−.12	−.31	−.06	.25	.37	.31	.19	.10	.04	.01
.8	.29	−.07	−.31	−.12	.21	.36	.33	.21	.11	.05	.02
7.0	.30	−.00	−.30	−.17	.16	.35	.34	.23	.13	.06	.02
.2	.30	.05	−.28	−.21	.11	.33	.35	.25	.15	.07	.03
.4	.28	.11	−.25	−.24	.05	.30	.35	.27	.16	.08	.04
.6	.25	.16	−.21	−.27	−.00	.27	.35	.29	.18	.10	.04
.8	.22	.20	−.16	−.29	−.06	.23	.35	.31	.20	.11	.05
8.0	.17	.23	−.11	−.29	−.11	.19	.34	.32	.22	.13	.06
.2	.12	.26	−.06	−.29	−.15	.14	.32	.33	.24	.14	.07
.4	.07	.27	−.00	−.27	−.19	.09	.30	.34	.26	.16	.08
.6	.01	.27	.05	−.25	−.22	.04	.27	.34	.28	.18	.10
.8	−.04	.26	.10	−.22	−.25	−.01	.24	.34	.29	.20	.11
9.0	−.09	.25	.14	−.18	−.27	−.06	.20	.33	.31	.21	.12
.2	−.14	.22	.18	−.14	−.27	−.10	.16	.31	.31	.23	.14
.4	−.18	.18	.22	−.09	−.27	−.14	.12	.30	.32	.25	.16
.6	−.21	.14	.24	−.04	−.26	−.18	.08	.27	.32	.27	.17
.8	−.23	.09	.25	.01	−.25	−.21	.03	.25	.32	.28	.19
10.0	−.25	.04	.25	.06	−.22	−.23	−.01	.22	.32	.29	.21

Approximations for small arguments:

$$J_0(x) \approx 1 - \frac{x^2}{4}$$

$$J_1(x) \approx \frac{x}{2}\left(1 - \frac{x^2}{8}\right)$$

$$J_2(x) \approx \frac{x^2}{8}\left(1 - \frac{x^2}{12}\right)$$

APPENDIX

H Stereo AM†

For more than 50 years in the United States interest has been expressed by various individuals and groups in the transmission of stereophonic (stereo) information using amplitude modulation (AM). In the late 1950s and early 1960s several groups (Philco, RCA, Westinghouse, and Kahn Communications) independently petitioned the Federal Communications Commission (FCC) to approve AM stereo broadcasting based on their respective systems. In 1962 the FCC ruled against all of these proposals, deciding instead to encourage the growth and expansion of stereo broadcasting using frequency modulation (FM). Since that time, public demand for stereo broadcasting has increased steadily and—more recently—has expanded into the automobile market.

As a result of the increasing interest in AM stereo, the National AM Stereophonic Radio Committee (NAMSRC) was formed in 1975, and this committee submitted a report on the subject to the FCC in 1977. The FCC conducted evaluations and held hearings on five competing AM stereo systems, with no conclusions on clear overall advantages to selection of any one system. Then a decision was made to let the marketplace decide, allowing AM broadcasters to implement any of the five approved systems. This has slowed development; both manufacturers of receivers and broadcasters have been hesitant to venture into use of any system, because each system requires different transmitter/receiver circuitry. Because no conclusions have been reached on this matter at the time of writing, this appendix summarizes briefly the five competing AM stereo options.

H.1 AM STEREO VERSUS FM STEREO

Before describing the various systems being proposed for AM stereo, we consider briefly what possible advantages AM stereo may have over FM stereo for certain situations. One advantage of wideband FM systems is the noise quieting that can be realized in demodulation, and the problem of an increasing noise power spectral

† Because the material in this appendix depends on a knowledge of both amplitude modulation and angle modulation, we recommend that the reader cover Chapters 5 and 6 before attempting to read this appendix.

Several references to this appendix include the following: D. Mennie, "AM Stereo: Five Competing Options," *IEEE Spectrum* (June 1978): 24–31; and "Whatever Happened to AM Stereo?" *IEEE Spectrum* (March 1988): 17.

density with frequency in FM detection is usually minimized with preemphasis/deemphasis. However, in FM stereo broadcasting, the stereo information is preemphasized prior to the DSB-SC modulation to avoid overmodulation of the transmitter. This results in a S/N degradation of 22 dB (Problem 6.10.5)—in FM stereo compared to FM mono—that partially offsets the FM noise quieting advantage.

Because the modulation in AM transmitters is limited by negative peaks in program material (to avoid overmodulation), many AM stations use audio compression circuits to compress the negative peaks in the program material before modulating the transmitter. In this manner the modulation on positive peaks can be increased (up to 125%), with the result that detected signal output in the receiver is louder in comparison to the noise level. This aids in the perceived signal-to-noise (S/N) ratio. However, this comparative gain is not so definitive because some FM stations also use audio compression to effectively increase loudness within the constraint of a fixed peak frequency deviation. Exceptions to this practice are those stations—many of which are classical music stations—that consider audio compression a type of undesired audio "distortion."

As a result of the higher frequencies used, the performance of FM systems tends to be influenced by multipath propagation. This is particularly noticeable in the rapid fading effects and varying signal intensities in the FM reception in vehicles moving in the vicinity of tall buildings or other structures. This is not as much of a problem in AM reception as long as the receiver is in an area of sufficient net signal strength and has a reasonably good automatic gain control. Also as a result of the frequencies used, FM has, basically, line-of-sight propagation, whereas AM has a greater potential range.

Finally, there is a question regarding adequate frequency response. In AM systems, the FCC requires a check of transmitter response for inputs from 100 to 7500 Hz (even though many AM receivers do not have good responses beyond 5 kHz). Some commercial AM stations have flat responses for inputs up to 10 kHz, and a few have good responses out to 15 kHz. Transmitted bandwidths are governed primarily by transmitter capabilities and possible cochannel interference from other stations in a particular geographical area. With the possibility of AM stereo, there is some concern that many stations will want to increase transmitter bandwidths to handle input frequencies in the 10–12 kHz range. Also, most of the AM stereo options under consideration at this time use some type of nonlinear modulation that creates extra sideband components which may fall outside the bandwidth allocated. One further complication is that 9-kHz spacings between AM station allocations instead of the present 10-kHz spacings are being considered in order to permit more stations in the AM band. These are unresolved problems.

Overall, it appears that AM stereo does hold some promise for improved stereo reception in mobile applications. It is estimated that 80% of the market for AM stereo receivers will be in automobile installations. In this type of application a slightly inferior frequency response and S/N, compared to FM stereo, may be acceptable in exchange for less fading and greater distance. Also, AM stereo receivers should be

simpler, cheaper, and more stable than FM stereo receivers. With an increasing interest in stereo systems in automobiles, perhaps the time for AM stereo has arrived!

The five AM stereo systems that attempt to meet these objectives have been submitted by Kahn Communications and the Hazeltine Corporation, Belar Electronics, Magnavox Company, Harris Corporation, and Motorola. In the following descriptions, each system is referred to by the company name.

H.2 THREE AM STEREO SYSTEMS (KAHN/HAZELTINE, BELAR, MAGNAVOX)

Desirable characteristics for AM stereo systems include compatibility with existing monophonic receivers, good stereo separation, good fidelity, minimal degradation (if any) in fringe-area reception, no increase in required bandwidth, adaptability to existing broadcast transmitter facilities, and low cost and complexity. As will be seen shortly, these characteristics are not as easy to obtain as they might seem at first glance.

H.2.1 Kahn/Hazeltine

In the Kahn/Hazeltine system, the upper and lower sidebands independently carry the right-channel (R) and left-channel (L) information, respectively. A mono receiver center-tuned to the carrier frequency adds the sidebands and detects the envelope, giving compatible (L + R) monophonic reception. Two mono AM receivers will detect stereo—with adequate but not good stereo separation—if the left receiver is tuned slightly below carrier and the right receiver is tuned slightly above carrier. A receiver with a single intermediate frequency (IF) has been developed for this system and gives better stereo reception. It uses envelope detection for the L + R signal and quadrature detection with phase-shifting networks for the L − R signal. These signals are then added and subtracted (this is called matrixing) to yield the L and R signals. During the early 1970s the Kahn system was used on an experimental basis at station WFBR in Baltimore, MD, and station XETRA in Tijuana, Mexico. The Hazeltine Corporation joined with Kahn in 1977 in support of this AM stereo system.

The transmitter in the Kahn/Hazeltine system uses the principles of compatible single-sideband (CSSB) modulation to obtain the independent sidebands. An L + R signal amplitude modulates the transmitter while the L − R signal is shifted by 90° (in practice, L + R is shifted −45° and L − R is shifted 45°) and then used to phase-modulate the transmitter. To give better stereo separation, a second-order correction signal is formed by squaring both the L and R signals and adding their difference to the L − R signal prior to the phase modulation. A 15-Hz tone can be added to the L − R signal, which, upon detection, is used for a stereo indicator light. This system has the potential for good stereo performance under varying propagation conditions (but at the expense of stereo separation in the double-IF version) and requires accurate phase shifting networks in the single-IF version.

H.2.2 Belar

A second AM stereo system is the Belar AM/FM system. The first proposal for this type of system for stereo was made by RCA in 1959 and was tested over station WRCA in New York City. Now known as the Belar system, this system uses narrowband FM to modulate L − R information on the carrier. The peak frequency deviation is 1.25 kHz and preemphasis is used with a 100 μsec time constant. The L + R signal is used to amplitude-modulate the carrier. A negative amplitude-modulation limit of 95% (5% residual carrier) is used to prevent excessive noise problems in the receiver limiter stages. The Belar system contains no provision for a stereo identification signal to operate a stereo indicator light in the receiver. Judging from their experience with FM stereo, however, manufacturers think it likely that potential users will want some type of indicator for stereo broadcasts.

H.2.3 Magnavox

A third AM stereo system is the AM/PM system by Magnavox. In this system, an oscillator at the carrier frequency is frequency-modulated with a 5-Hz stereo identification tone using a peak frequency deviation of 20 Hz (β = 4). The L − R signal is used to phase-modulate this carrier using a peak phase deviation of one radian (β = 1). The L + R signal is used for the amplitude modulation signal to the transmitter. A simplified block diagram of the Magnavox system is shown in Fig. H.1.

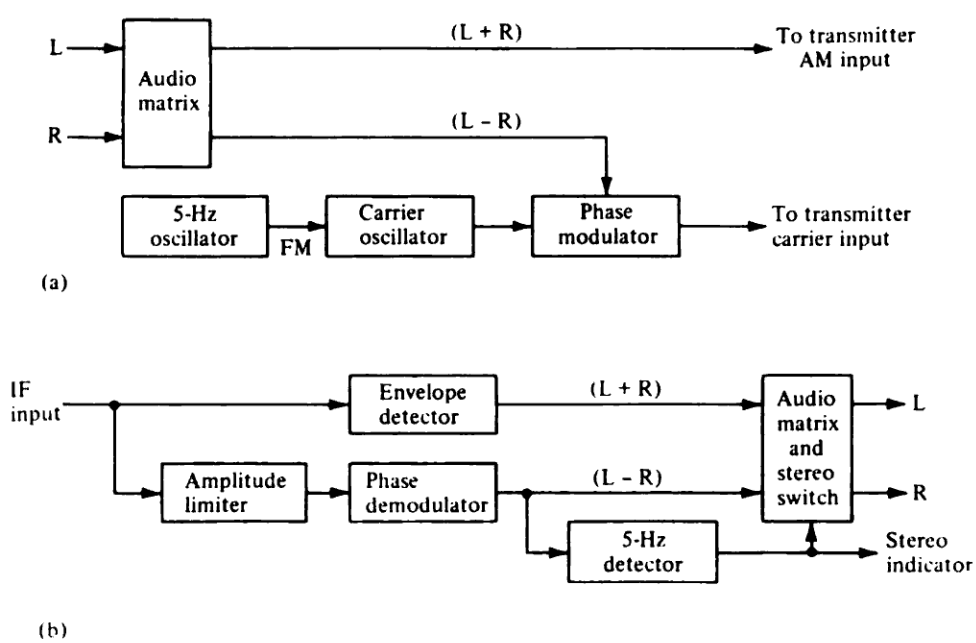


Figure H.1 Simplified diagram of the transmitter and receiver signal processing for the Magnavox AM stereo system.

Detection of the Magnavox AM/PM signal is implemented by using an envelope detector to obtain the L + R signal and using an amplitude limiter and phase detector to obtain the L − R signal. As in the Belar system, a negative modulation limit of 95% (5% residual carrier) is used to prevent excessive noise problems in the receiver limiter stages. The L + R and L − R signals are matrixed (i.e., added and subtracted) to obtain the L and R signals. Detection of the 5-Hz frequency modulation operates a stereo indicator light. Magnavox suggests the possibility of sending digital data (station identification, time, etc.) also using this 5-Hz tone capability.

Early in 1980, the Magnavox system was given preliminary acceptance for AM stereo broadcasting in the United States. Shortly after that decision, however, the FCC was asked to reconsider its preliminary ruling and further tests, hearings and studies were conducted.

Both the Belar and the Magnavox proposals use a combination of AM and angle modulation for stereo transmissions. Both are relatively easy and economical to implement in the receiver as well as the transmitter. Both must maintain at least 5% residual carrier on negative peaks for acceptable angle demodulation. Also, the detected angle modulation could be expected to be noisy if the signal amplitude remains at this low level for any appreciable time interval.

H.3 TWO ADDITIONAL SYSTEMS (HARRIS AND MOTOROLA)

The remaining two systems make use of modifications of quadrature multiplexing. Quadrature amplitude modulation (QAM) offers advantages of good performance in the presence of noise and the transmission and detection of two independent signals within the same bandwidth using a linear combination of the two signals. QAM is used successfully, for example, in the multiplexing of the chroma (color) information in commercial television broadcasting.

While QAM has much to offer for the transmission of stereo signals, a major problem is that of compatibility with envelope detection. Suppose we designate the left-channel signal as $\ell(t)$ and the right-channel signal as $r(t)$. The in-phase component $x(t)$ used to amplitude-modulate the carrier is

$$x(t) = 1 + \ell(t) + r(t) \tag{H.1}$$

and the quadrature signal is

$$y(t) = \ell(t) - r(t). \tag{H.2}$$

A phasor diagram of this result is shown in Fig. H.2. An in-phase detector will detect the AM signal (L + R); and a quadrature detector, the DSB-SC signal (L − R). However, an envelope detector will detect

$$e(t) = \sqrt{[1 + \ell(t) + r(t)]^2 + [\ell(t) - r(t)]^2}. \tag{H.3}$$

The presence of the L − R component in Eq. (H.3) may produce significant distortion in the envelope detector output if the L − R and L + R components are of

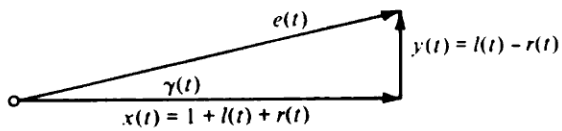


Figure H.2 A phasor diagram of the in-phase and quadrature components in quadrature amplitude modulation (QAM) for AM stereo.

comparable magnitude and if the modulation index is relatively high. Because these are realistic conditions, some modifications are necessary in quadrature multiplexing if compatibility with envelope detection is to be retained.

H.3.1 Harris

In the Harris AM stereo system, the L signal is applied to a balanced mixer whose reference is phase-shifted by −15 degrees with respect to carrier. This is shown in Fig. H.3(a). Similarly, the R signal is applied to a balanced mixer whose reference is phase-shifted by +15 degrees with respect to the carrier. The outputs of the two mixers are added together with the carrier, as indicated by the phasor diagrams in Fig. H.3(b).

The resulting signal envelope from the Harris AM stereo system is

$$e(t) = \sqrt{\{1 + [\ell(t) + r(t)] \cos 15°\}^2 + \{[\ell(t) - r(t)] \sin 15°\}^2}. \qquad \text{(H.4)}$$

The reduction in the L − R component reduces the envelope distortion significantly at the penalty of some S/N loss. The signal power loss in the in-phase component is rather minor:

$$10 \log_{10} (\cos 15°)^2 = -0.30 \text{ dB},$$

but the power loss in the quadrature component is appreciable:

$$10 \log_{10} (\sin 15°)^2 = -11.7 \text{ dB}.$$

On the plus side, the Harris system uses entirely linear modulation methods and, therefore, it is not as prone to the generation of harmonics outside a well-defined bandwidth as are the other four systems. (Note that the second section of the transmitter diagram in Fig. H.3(a) is needed only to match into an existing AM transmitter; the signal leaving the adder in the first section could be broadcast with a linear radio-frequency amplifier with the same results.)

While an envelope detector can do a good job of demodulating the mono signal from the Harris AM stereo system, a stereo receiver would have to have standard quadrature demultiplexing circuitry. A popular way to implement this circuitry is to use a phase-locked loop with both in-phase and quadrature mixers to demodulate the L + R and L − R signals. This is illustrated in Fig. H.3(c). For stereo identification, the Harris system transmits a 20–25 Hz tone at 9% modulation in the L − R channel only. This causes a negligible increase in envelope detector distortion and is relatively easy to separate in the stereo receiver by filtering.

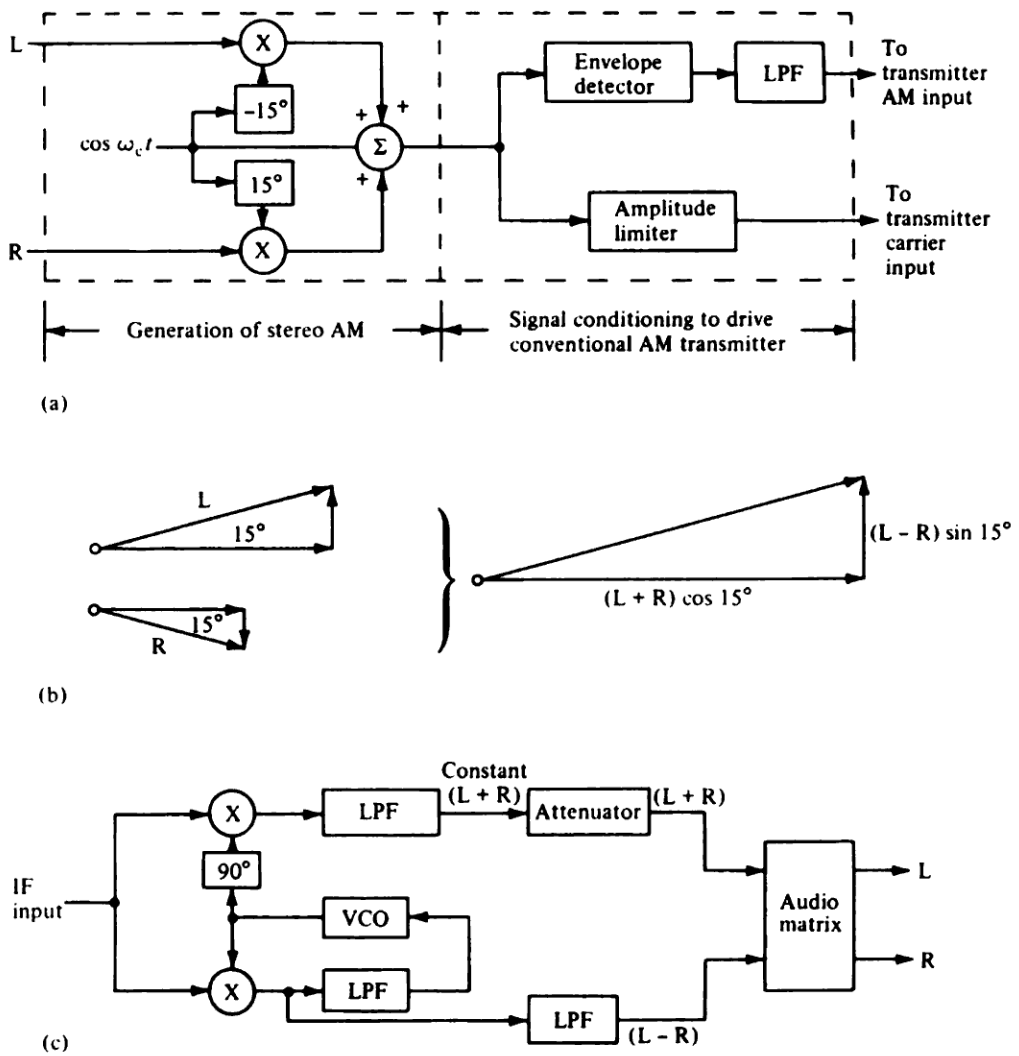


Figure H.3 Simplified diagrams of the transmitter and receiver signal processing for the Harris AM stereo system.

H.3.2 Motorola

The Motorola C-QUAM† system also uses quadrature multiplexing but retains a full 90° between the modulated in-phase and quadrature components. The resulting signal in terms of magnitude and phase is (see Fig. H.2):

$$\phi(t) = \sqrt{[1 + \ell(t) + r(t)]^2 + [\ell(t) - r(t)]^2} \cos [\omega_c t + \gamma(t)] \tag{H.5}$$

where

$$\gamma(t) = \tan^{-1} \left[\frac{\ell(t) - r(t)}{1 + \ell(t) + r(t)} \right]. \tag{H.6}$$

† C-QUAM is the acronym used by Motorola for compatible quadrature amplitude modulation.

From Fig. H.2, note that

$$\cos \gamma(t) = \frac{1 + \ell(t) + r(t)}{\sqrt{[1 + \ell(t) + r(t)]^2 + [\ell(t) - r(t)]^2}}. \tag{H.7}$$

Comparing Eq. (H.7) with Eq. (H.3), we find that the resulting envelope can be forced to be compatible with envelope detection by multiplying by $\cos \gamma(t)$,

$$e(t) \cos \gamma(t) = 1 + \ell(t) + r(t). \tag{H.8}$$

Simplified diagrams of the Motorola AM stereo system are shown in Fig. H.4. After both the in-phase and quadrature components are modulated, they are added and amplitude-limited. This forms a phase-modulated carrier to the transmitter, and the L + R component is applied to the AM input to the transmitter. The resulting transmitted signal is the quadrature amplitude modulation signal described by Eq. (H.8) and is compatible with envelope detection for recovery of the mono signal. For stereo identification, a 25-Hz tone is added to the L − R component prior to modulation.

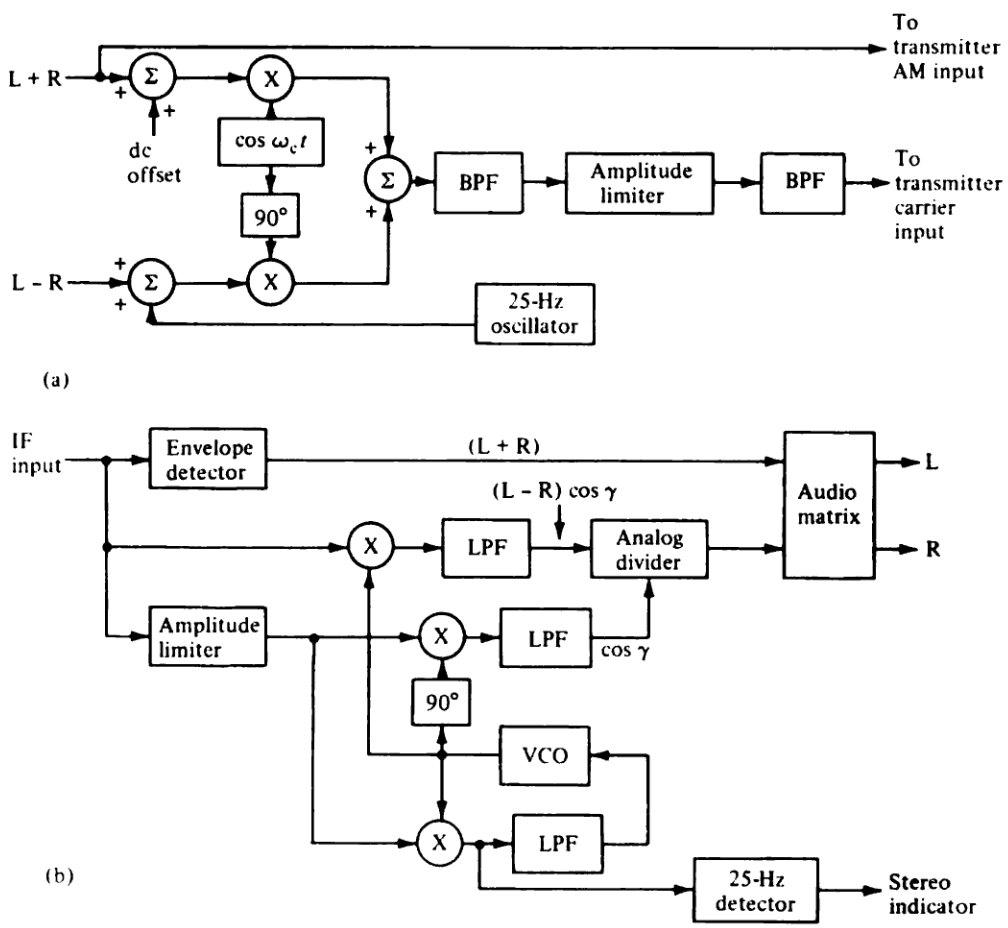


Figure H.4 Simplified diagrams of the transmitter and receiver signal processing for the Motorola AM stereo system.

The burden of compatibility is placed on the stereo receiver in the Motorola system because the term $\cos \gamma(t)$ must be detected and factored out of the L − R component. A simplified receiver diagram is shown in Fig. H.4(b). A phase-locked loop follows the amplitude limiter; thus the in-phase and quadrature components of the phase-modulated carrier are generated. Synchronous detection with the quadrature component gives $(L - R) \cos \gamma$, and synchronous detection with the in-phase component gives $\cos \gamma$. These are applied to an analog divider to yield the L − R signal. The envelope detector output (L + R) and the divider output (L − R) are applied to an audio matrix that provides the desired stereo outputs.

H.4 CONCLUSION

The five AM stereo systems described here have all been field tested. Each system has its own advantages and disadvantages, and much of the weight of a decision concerning the implementation of one of them must be based on the cost efficiency trade-offs of various performance criteria. Among these are the increased bandwidth required, stereo separation, frequency response, and noise under various receiving conditions. Additional factors include not only mono and stereo receiver distortion under skywave-propagation and selective-fading conditions, but also distortion caused by both mistuning and restricted bandwidth conditions. Other considerations that may also be important are the potential reduction of the mono reception area, plans for system development through current and future transmitter and receiver designs, and the use of stereo indication. Trade-offs involving all of these are typical in the design of any new communication system.

APPENDIX

A Table of Gaussian Probabilities

The gaussian probability density function with zero mean and unit variance is

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{I.1}$$

The probability that the gaussian random variable X is greater than a given value x is given by the function $Q(x)$, defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-z^2/2}\, dz\,. \tag{I.2}$$

Note that $Q(0) = \frac{1}{2}$ and $Q(-\infty) = 1$. Some tabulated values of Eq. (I.2) are given in Tables I.1 and I.2. For large values of x, $Q(x)$ can be approximated by

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} \left[1 - \frac{1}{x^2}\right] e^{-x^2/2}. \tag{I.3}$$

The error in this approximation is about -2% for $x = 3$, -1% for $x = 4$, and the approximation becomes increasingly better for larger values of x. For large values of x, a convenient approximation is

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}. \tag{I.4}$$

An approximation with improved accuracy over that of Eq. (I.3) is[†]

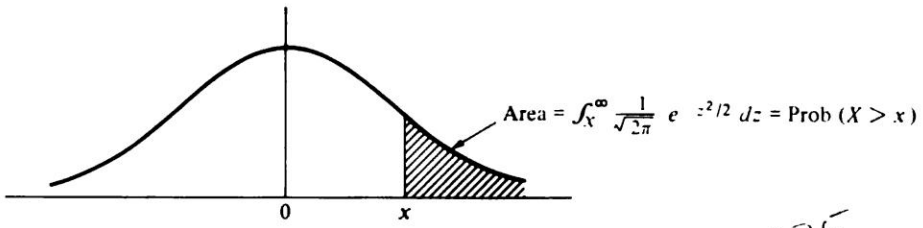
$$Q(x) \approx \left[\frac{1}{(1-a)x + a\sqrt{x^2+b}}\right] \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{I.5}$$

where $a = 1/\pi$ and $b = 2\pi$. This approximation is within -1% for $x > 1$, -0.5% for $x > 2$, and becomes increasingly better for larger values of x.

The error function used in statistics is usually defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt\,, \tag{I.6}$$

[†] P. O. Börjesson and C. W. Sundberg, "Simple Approximations of the Error Function $Q(x)$ for Communications Applications," *IEEE Transactions on Communications*, COM-27 (March 1979): 639–643.



0.025

TABLE I.1 Values of $Q(x)$ versus x.[a]

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139

[a] From J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, New York, Wiley-Interscience, 1971, and D. B. Owen, *Handbook of Statistical Tables*, Reading, Mass., Addison-Wesley Publishing Co., 1962, both by permission; courtesy of the U.S. Energy Research and Development Administration.

TABLE I.2 Values of $Q(x)$ for Large x.

x	$10 \log x$	$Q(x)$	x	$10 \log x$	$Q(x)$	x	$10 \log x$	$Q(x)$
3.00	4.77	1.35E-03	4.00	6.02	3.17E-05	5.00	6.99	2.87E-07
3.05	4.84	1.14E-03	4.05	6.07	2.56E-05	5.05	7.03	2.21E-07
3.10	4.91	9.68E-04	4.10	6.13	2.07E-05	5.10	7.08	1.70E-07
3.15	4.98	8.16E-04	4.15	6.18	1.66E-05	5.15	7.12	1.30E-07
3.20	5.05	6.87E-04	4.20	6.23	1.33E-05	5.20	7.16	9.96E-08
3.25	5.12	5.77E-04	4.25	6.28	1.07E-05	5.25	7.20	7.61E-08
3.30	5.19	4.83E-04	4.30	6.33	8.54E-06	5.30	7.24	5.79E-08
3.35	5.25	4.04E-04	4.35	6.38	6.81E-06	5.35	7.28	4.40E-08
3.40	5.31	3.37E-04	4.40	6.43	5.41E-06	5.40	7.32	3.33E-08
3.45	5.38	2.80E-04	4.45	6.48	4.29E-06	5.45	7.36	2.52E-08
3.50	5.44	2.33E-04	4.50	6.53	3.40E-06	5.50	7.40	1.90E-08
3.55	5.50	1.93E-04	4.55	6.58	2.68E-06	5.55	7.44	1.43E-08
3.60	5.56	1.59E-04	4.60	6.63	2.11E-06	5.60	7.48	1.07E-08
3.65	5.62	1.31E-04	4.65	6.67	1.66E-06	5.65	7.52	8.03E-09
3.70	5.68	1.08E-04	4.70	6.72	1.30E-06	5.70	7.56	6.00E-09
3.75	5.74	8.84E-05	4.75	6.77	1.02E-06	5.75	7.60	4.47E-09
3.80	5.80	7.23E-05	4.80	6.81	7.93E-07	5.80	7.63	3.32E-09
3.85	5.85	5.91E-05	4.85	6.86	6.17E-07	5.85	7.67	2.46E-09
3.90	5.91	4.81E-05	4.90	6.90	4.79E-07	5.90	7.71	1.82E-09
3.95	5.97	3.91E-05	4.95	6.95	3.71E-07	5.95	7.75	1.34E-09

and the corresponding complementary error function is

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x). \tag{I.7}$$

Our particular choice $Q(x)$ is more convenient for work in communication systems.

Using a change of variables, it is easily seen that the relations between Eqs. (I.2), (I.6), (I.7) are

$$Q(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right), \tag{I.8}$$

$$\operatorname{erfc}(x) = 2Q(\sqrt{2}x). \tag{I.9}$$

Index

Acquisition, 630, 657
ACTV-1 (Advanced Compatible TV), 726
A/D (analog-to-digital) conversion, 405, 563
Adaptive equalizer, 580, 664
Algebraic codes, 590
Algorithms, 142, 675
Aliasing, 53, 55, 129, 131, 134
AM (amplitude modulation), 102, 219, 239
  demodulation, 232, 247
  generation, 225, 244
AM broadcasts, 251, 701
AM stereo, 737
Amateur bands, 705
AMI (alternate mark inversion), 410, 569
Amplitude spectrum, 45, 48
Analytic signals, 259
Analog pulse modulation, 399
Analog repeater, 564
Angle modulation, 298
Antenna, 277
  gain and effective area, 204
  noise, 198, 202
Antipodal signaling, 619, 670
Aperiodic signals, 11, 82
APK (amplitude-phase keying), 663, 667
  demodulation, 663
  generation, 663
Armstrong method of FM, 322
ASCII code, 419
ASK (amplitude-shift keying), 412, 608, 626, 670
  demodulation, 610
  generation, 608
Aspect ratio, 709
Asynchronous transmission, 419
Atmospheric noise, 203
AT&T (American Telephone & Telegraph Company), 408, 729
Attenuation losses, 204, 206
Augmenting zeros, 137, 145
Autocorrelation function
  energy signals, 187, 435, 609
  power signals, 180, 505, 509
Available power, 197
Average information, 532
Average value, 177, 471
Averages
  ensemble, 499
  statistical, 471
  time, 177, 500

Balanced modulator, 231
Bandpass noise, 268, 270
Band-limited noise, 189
Bandwidth, 2, 4, 118
  noise-equivalent, 195
  radius-of-gyration, 125
  3-dB, 119
Bandwidth efficiency, 411, 668
  ASK, 610
  FSK, 616
  PSK, 621
  QAM, QPSK, 640
  M-ary FSK, 655
  M-ary PSK, 659
  MSK, 646
Bandwidth vs. S/N trade-offs,
  FM, 347
  ideal demodulator, 540, 687
  M-ary FSK, 653, 677
  PCM, 412, 558
  PPM, PWM, 401
Barker codes, 442, 455
Baseband transmission, 547
Basis functions, 16
Baud, 584, 634
Bayes' theorem, 461
Belar AM stereo system, 740
Bell System M12 multiplexer, 423
BER (bit error rate), 637, 640, 662, 668
Bessel functions, 310, 735
BFO (beat-frequency oscillator), 266
Binary codes, 403
Binary representations, 143, 410
Binary decision, 550, 673
Binary symmetric channel, 536
Binit, 533
Binomial coefficient, 478
Binomial distribution, 478, 481, 566
Binomial series, 694
Bit, 5, 404, 532
Bit error, 660
  vs. symbol error, 587
Bit integrity, 424
Bit-interleaved multiplexer, 421
Bit rate, 584
Bit stuffing, 420, 422
Bivariate gaussian distribution, 497
Block codes, 591
Boltzmann's constant, 190, 204
BPF (ideal bandpass filter), 120
BPSK (binary PSK), 620, 636, 666, 670
BRI (basic-rate interface), 430
BRZ (bipolar return-to-zero), 410, 569
Buffer memory, 422
Burwen compression system, 734
Butterworth filter, 134, 156, 196, 214, 392

Cable TV frequency allocations, 704

Camp-and-wait acquisition strategy, 657
Carrier frequency, 220, 251, 315
Carrier recovery, 623, 639, 685
Carrier signal, 220
Carson's rule, 313
Causal system, 13, 109
CCD (charge-coupled device), 117
CCITT (Consultative Committee on International Telegraphy and Telephony), 419, 430
CDMA (code-division multiple access), 628
Channel, 3, 536
  bank, 409
  capacity, 5, 536
Character, 419
Character-interleaved multiplexer, 426
Chebyshev's inequality, 473
Chip rate, 630, 655
Chopper modulator, 228, 245
Chopper-stabilized amplifier, 233
Chrominance TV signal, 713
Citizens band, 705
Click noise, 353
CNR (carrier-to-noise ratio), 274
Codec, 4
Coding, 3
  algebraic, 590
  convolutional, 593
  correlative, 572
  error-control, 418
Code-rate efficiency, 417, 589
Codeword, 437, 590
Coefficient of determination, 497
Coherent detection, 223, 435, 615, 626
Color television, 714
Comb function, 94
Communication, 1
Communication link, 204
Communication systems, 2, 531
Commutator, 377, 384, 593
Commutator clock rate, 384
Companding, 544
  A-law, 599
  μ-law, 545, 599
Comparator, 397
Comparison of modulation methods
  AM, 281
  binary digital modulation, 626
  CW, 350
  digital modulation, 666
  FM, PPM, 402
Comparison of multiplexing methods,
  asynchronous, synchronous, 419
  FDM, TDM, 384
Completeness, 14, 17, 24, 30
Complex-valued exponential, 23–29, 38
Complex-valued signals, 8, 26, 130
Composite modulation, 326, 383
Concentrator, 420
Conditional probability, 461, 490
Conditioned telephone lines, 729
Consistent estimator, 476, 516
Constraint length [span], 510, 593
Control bits, 422
Convergence, mean-square, 34, 67, 69
Convolution integral, 80, 106, 109, 526
  in frequency, 107
  graphical, 111
  numerical, 114
  periodic [circular], 137
  properties of, 109–114
  with unit step, 110
Convolutional codes, 593
Cooley-Tukey algorithm, 142
Correlation, 495
Correlation coefficient, 494
Correlation function, 179–187
Correlative coding, 572
Correlator, 435, 438, 636
Cosmic noise, 203
Costas loop, 624, 639, 666
Covariance, 493
CPFSK (continuous-phase FSK), 641, 650, 667
Critical angle, 414
Cross-correlation
  energy signals, 187, 435
  power signals, 183, 334, 582
Crosstalk, interchannel, 381, 384
CSSB (compatible single-sideband), 265
Cumulative distribution function, 464, 501
Cumulative power spectrum, 170
CW (continuous-wave) signals, 219, 350, 371

DAC (digital-to-analog converter), 406, 450
DBX compression system, 724, 733
Decibel (dB), 178 *n*, 698
Decision rule, 549, 676
Decision-directed receiver, 583, 639
Decoder, 2, 4, 591, 593
Deemphasis, 355
Delay distortion, 667, 671, 729
Delay-differencing system, 111
Delay modulation, 410, 571
Delta function, 59
Demand multiplexing, 427
Demodulation, 3; *see* Modulation
DE-PSK (differentially-encoded PSK), 624
Despreading, 629
Detection
  coherent [synchronous], 223, 272, 435, 612
  correlation, 182, 438, 441, 619
  envelope, 247, 273, 290, 610, 613
  homodyne, 223
  slope, 329
Detection in noise, 181, 550, 673
Deterministic signal, 12, 457
DFT (discrete Fourier transform), 135, 515
Digitization, 403
Dirac delta function, 59
Direct FM, 323, 327
Dirichlet conditions, 66, 87
Discriminator, 327, 345
Dispersion index, 307
Distortionless transmission, 119
DLL (delay-locked loop), 631
DM (delta modulation), 561
Dolby compression system, 356 *n*, 732
Double-balanced ring modulator, 229
Downconverter, 216
DPCM (differential PCM), 560

DPSK (differential PSK), 625, 627, 666
DQPSK (differential QPSK), 667
DS (direct-sequence) spread spectrum, 629
DSB-LC (double-sideband large-carrier), 239, 350
  demodulation, 247
  generation, 244
DSB-SC (double-sideband suppressed-carrier), 221, 231, 350
  demodulation, 232
  generation, 225
Dual-conversion receiver, 252 *n*, 254
Dual-slope encoder, 405
Duobinary signaling, 572, 665
Duplexer, 565

EIRP (effective isotropic radiated power), 205
Elastic store, 422
Encoder, 2, 4, 593
Energy, 88, 164
Energy signals, 10, 14
Energy spectral density, 164
Ensemble, 499
Entropy, 532, 534, 536
Envelope, 220, 240
  delay distortion, 730
  detection, 240, 247, 264, 273, 610, 613
Equalization, 120, 281, 376, 548, 564, 577
  adaptive, 580, 664
  preset, 580
  zero-forcing, 578
Equally-likely events, 460
Equivalent noise bandwidth, 195
Equivalent noise temperature, 197
Erf (error function), 746
Ergodic processes, 500, 506
Error-control coding, 418
Error probability; *see* Probability of error
Errors, 418
  integral-squared, 19, 669
  mean-squared, 74, 496, 582
Estimation, 476, 516
Euler's identities, 27
Even functions, 32, 97
Even parity, 417
Even symmetry, 32, 108, 139
Expected value, 471
Exponential Fourier series, 25, 57, 82
Exponential pdf, 469
Eye pattern, 559, 585

Facsimile transmissions, 539, 730
Fading, 279
Fast-hop spread spectrum, 657
FCC (Federal Communications Commission), 289 *n*, 315, 356 *n*, 701, 714, 724, 737
FDM (frequency-division multiplexing), 249, 384, 627
FDMA (frequency-division multiple access), 691
FDX (full-duplex) transmission, 4, 617
Feedback encoder, 405
FFT (fast Fourier transform), 141
FH (frequency hopping) spread spectrum, 655
Fiber optic communication system, 413
Filter characteristics, 114
  Butterworth, 134, 156, 196, 214, 392
  deemphasis, preemphasis, 356, 368, 732, 738
  equalizing; *see* Equalization
  ideal, 120
  integrate-and-dump, 436
  matched, 433
  pre-alias, 132
  RC low-pass, 39, 79, 115, 124, 194
  RC high-pass, 213, 328
  sideband, 256, 266
  transversal, 116, 155, 434, 510, 577
Flash [parallel] encoder, 406, 663
FM (frequency modulation), 300, 350, 701, 738
  demodulation, 327
  generation, 321
  multiplexing, 326
  narrowband (NBFM), 301
  noise in, 345
  wideband, 306
FMFB (FM with feedback), 333, 354
FM stereo, 235, 286, 316, 737
Foster-Seeley discriminator, 329
Fourier series, 18
  exponential, 25, 82
  numerical computation of, 50
  coefficients, table of, 35
  trigonometric, 30
Fourier spectrum, 44
Fourier transform, 84
  calculation using impulses, 105
  graphs of transforms, 91
  numerical computation of, 135, 144
  of periodic functions, 92
  properties, 96–109
  of singularity functions, 89
  of sinusoids, 90
Frame, 421, 428, 708
Free-space path loss, 206
Frequency, 9, 23
  carrier, 220, 251, 315
  converter, 227, 322
  designations, 700
  discriminator, 327, 345
  diversity, 279
  domain, 23
  errors, 223, 263
  instantaneous, 299
  mixer, 227
  modulation; *see* FM
  multiplier, 42, 321
  negative, 29, 58
  of occurrence, 457, 485
  resolution, 139
  transfer function, 38, 114, 164, 175
  synthesizer, 237, 286, 655
FSK (frequency-shift keying), 412, 610, 626, 666, 670
  demodulation, 613
  generation, 617
Full-wave rectified waveform, 35, 211, 488

Galactic noise, 203
Gate (rect) function
  aperiodic, 86
  periodic, 35, 46

Gaussian distribution, 482
bivariate, 497
Gaussian probabilities, 483, 746
Gaussian pulse, 11, 95, 126, 148
Generalized functions, 63
Generation; *see* Modulation
Geostationary satellites, 297, 607, 622
Gibbs phenomenon, 68
Graded-index optical fiber, 415 *n*
Gray code, 588, 635, 660
Ground wave, 278
Guard band, 326, 367
Guard time, 391, 395

Half-wave rectified waveform, 35, 489
Hamming codes, 591
Hamming distance, 418, 426, 588
Hamming window, 160
Hann window, 140
Harmonic, 23, 76
distortion, 43
generation, 42, 81
Harris AM stereo system, 742
Hartley-Shannon theorem, 5, 537, 654, 687
HDB (high-density bipolar) codes, 452
HDS-NA (High-Definition System for North America), 726
HDTV (high-definition TV), 725
HDX (half-duplex) transmission, 4, 617
Heterodyne, 227
receiver, 251
Hilbert transform, 260
Histogram, 485, 488
Holding circuit, 66, 447
Homodyne detection, 223
Horizontal TV scan frequency, 720, 724
HPF (ideal high-pass filter), 120
Hue, 713
Huffman coding, 598

I (in-phase) signal, 269, 635, 645, 663, 671, 715, 718
Ideal filter, 120
Ideal demodulator, 540
IDFT (inverse discrete Fourier transform), 136
Idling noise, 562
IDTV (improved-definition TV), 712
IF (intermediate frequency), 251, 331, 657
Image frequency, 252, 254, 292
Impulse,
function, 59, 79
response, 65, 110, 114
sampling, 374
Index of discrimination, 442
Indirect FM, 321
Information, 531
Information transmission, 1, 5, 535
Injected-carrier receiver, 248
Inner product, 16, 669
In-phase component of noise, 269
Instantaneous frequency, 299
Integral-squared error, 19, 669
Integrate-and-dump filter, 436, 629, 639
Integrated power spectrum, 170
Intelsat communication satellite system, 691
Interchannel crosstalk, 381, 384, 386, 448
Interlaced scanning, 708
Intermodulation distortion, 227, 285
Intersymbol interference; *see* ISI
Inverse Fourier transform, 84
Ionosphere, 278
ISDN (integrated services digital network), 429
ISI (intersymbol interference), 386, 555, 559, 577, 584, 667

Jamming, 632, 656
Jitter, 424, 624, 671
Joint probability, 461, 490
Joint moment, 505

Kahn/Hazeltine AM stereo system, 739

Laplacian pdf, 525
Laser diode, 416
LC (large-carrier) modulation, 238, 264, 350
Least-squares filtering, 528
LED (light-emitting diode), 416
Legendre polynomials, 22, 72
Lender precoding method, 574
Likelihood ratio, 675
Limiter, 330, 345
Line spectrum, 45, 77, 82
expressed as impulses, 86
power, 172
Linear
modulation, 300
prediction, 561
regression, 496
systems, 12, 22, 106, 227
time-invariant systems, 13
Link power budget, 416
LO (local oscillator), 252, 332
LPF (ideal low-pass filter), 120
LRT (likelihood ratio test), 675
Luminance TV signal, 713

Magnavox AM stereo system, 740
Magnitude, 298
response, 40, 115
spectrum, 45
Majority-decision rule, 417, 424
Manchester binary representation, 410, 568
Marginal pdf, 491
Mark, 410, 608
Marker word, 425
M-ary modulation, 553, 584, 634
APK, 665, 667, 671
DPSK, 662
FSK, 651, 654
PAM, 404
PCM, 402
PSK, 659, 665, 667, 671
Matched filter, 433, 437
Matched-filter detection, 554, 609, 611, 619, 650
Matrixing, 715, 739
Maximal-length codes, 441
Mean-square
convergence, 34, 68
error, 74, 581
value, 168, 177, 184, 270, 471, 500
Mean time before failure (MTBF), 469
Mean value, 177, 471, 500
Median [statistical], 471
Message switching, 430
MFM (modified FM), 451
Miller code, 410, 571

Mixer, frequency, 227
Mode [statistical], 471
Modem, 4, 617
Modified duobinary, 575, 665
Modulation, 3, 219
  amplitude (AM), 102, 219, 239, 701
  amplitude-phase keying (APK), 663, 667
  amplitude-shift keying (ASK), 412, 608, 626, 670
  angle, 298
  compatible single-sideband (CSSB), 265
  composite, 326, 383
  continuous-phase FSK (CPFSK), 641, 650, 667
  delta (DM), 561
  differential PSK (DPSK), 625, 627, 666
  double-sideband (DSB), 221, 239
  frequency (FM), 300, 350, 701, 738
  frequency-shift keying (FSK), 412, 610, 626, 666, 670
  large-carrier (LC), 239, 264, 350
  linear, 300
  M-ary; *see* *M*-ary modulation
  minimum-shift keying (MSK), 642, 647, 666, 671
  narrowband FM (NBFM), 302, 304
  phase (PM), 299, 319, 350
  phase-shift keying (PSK), 320, 412, 618, 626, 670
  pulse, 371
  pulse-amplitude (PAM), 372, 380, 399
  pulse-code (PCM), 402, 411, 438, 450, 558
  quadrature AM (QAM), 634, 667, 727, 741
  quaternary PSK (QPSK), 634, 638, 647, 667, 671
  single-sideband (SSB), 255, 705, 730
  suppressed-carrier (SC), 220
  vestigial-sideband (VSB), 266, 627, 667, 710
  wideband FM, 306
Modulation constant
  AM, 220 *n*, 241 *n*
  FM, 300, 324, 337
  PM, 299, 320
  PWM, PPM, 392, 395, 401
Modulation index
  AM, 242, 288
  BPSK, 620
  FM, 302, 321
Modulator, 2
  APK, 664
  balanced, 229, 232, 304
  chopper, 228, 245
  FM, PM, 304, 323
  M-ary PSK, 688
  MSK, 648, 649
  power-law, 231, 246
  rectifier, 245
  ring, 229
Modulo-*N*, 136 *n*, 440
Moments [statistical], 472, 504
Motorola AM stereo system, 743
MSK (minimum-shift keying), 642, 666, 671
  demodulation, 685
  generation, 648, 649
MTS (multichannel television sound) system, 724
Multichannel analyzer, 502
Multichannel spectral analyzer, 166
Multipath, 155, 279, 604, 664, 738
Multiple access methods
  code-division (CDMA), 628
  frequency-division (FDMA), 691
  time-division (TDMA), 421
Multiplexing
  bit-interleaved, 421
  character-interleaved, 426
  demand, 427
  FM, 326
  FM stereo, 235, 286, 316, 369, 737
  frequency-division (FDM), 249
  of PCM signals, 419
  quadrature, 224, 634, 645, 714, 741
  time-division (TDM), 377
Multiframe, 422
Multivariate distribution, 497
MUSE (multiple sub-Nyquist encoding) system, 727
Mutually exclusive events, 459, 463
MUX (multiplexer), 409, 419
  Bell System M12 MUX, 423
μ-law companding, 545 *n*, 599

NAMSR (National AM Stereophonic Radio Committee), 737
Natural sampling, 373
NBFM (narrowband FM), 302, 348
Negative frequency, 29, 58, 220, 255
Nearest-neighbor approximation, 689
Nimbus F satellite, 530, 682
Noise, 3, 176
  antenna [sky], 198, 202
  atmospheric, 203
  bandpass, 269
  galactic [cosmic], 203
  partition, 208
  quantization, 541
  shot, 207
  thermal, 190, 206
  white, 188, 507
  see also Signal-to-noise ratios
Noise-equivalent bandwidth, 195
Noise figure, 199
Noise quieting, 347
Noise temperature, 197
Noncoherent FSK, 613, 626
Nonlinear systems, 12, 42, 231, 246
Nonlinear modulation, 310, 392
Nonlinearities, 42, 227, 246, 321, 384, 722
Nonrealizable, 13, 121
Norm, 16, 669
Normalized gaussian pdf, 483, 746
NRZ (nonreturn-to-zero) binary representation, 410, 567, 623
NTSC (National Television Systems Committee), 714
Numerical computation of
  convolution integral, 114, 137
  Fourier series, 50, 57
  Fourier transform, 135, 141
  power spectral density, 171, 515
  probability density function, 485, 502

Nyquist
bandwidth, 622, 665
interval, 129
sampling rate, 129
waveforms, 387, 556

Odd parity, 417
Odd symmetry, 32, 97, 108, 139
OOK (on-off keying), 547, 552, 555, 608, 626, 667, 670, 673
Optical fiber, 413
Optimum systems, 433, 437, 557, 652, 673
OQPSK (offset-keyed QPSK), 638, 644, 647, 666, 671
Orthogonal keying, 611, 642, 651
Orthogonality, 15, 24, 651
Orthonormal signals, 16, 19, 668

Packet switching, 430
PAL television system, 723
Paley-Wiener criterion, 121 *n*
PAM (pulse-amplitude modulation), 372, 381, 391, 399
Parallel encoder, 406
Parallel/serial converter, 405, 636
Parity, 417
Parity-check matrix, 417, 427, 591, 605
Parseval's theorem, 18
exponential Fourier series, 36, 171
Fourier transform, 88, 125, 149, 163, 618
trigonometric Fourier series, 74
Partial-response signaling, 572, 576, 665
Parzen window function, 160
Pascal's triangle, 479
Path losses, 204, 206
PCM (pulse-code modulation), 402, 409, 558
decoder, 407, 438, 450
encoder, 405, 563
telephone system, 407
PDM (pulse-duration modulation), 391 *n*
Peak frequency deviation, 301, 306, 320, 325
Peak phase deviation, 320
PEP (peak envelope power), 288, 292

Period, 11
Periodic gate function, 35, 46, 58
Periodic signal, 11, 33, 82, 91
Periodic steady-state response, 40
Periodogram, 515
Phase
angle, 298, 303
comparator, 334, 344
distortion, 264, 664
errors, 223, 263, 340, 622
modulation, 220, 299, 319
nulls, 335
response, 115
shift, 40, 120, 258
spectrum, 45
trellis diagram, 643, 646
Phase-locked loop; *see* PLL
Phasors, 27, 29, 34, 44, 84, 298, 717
AM, 275, 303
APK, 664
FM, 298, 303, 346, 351
M-ary PSK, 660
bandpass noise, 269
QAM, QPSK, 635, 638, 742
PSK, 620
Photodetector, 416
Pilot carrier, 233, 316
Pixels, 725
Plancherel's theorem, 88 *n*
PLL (phase-locked loop), 236, 333, 354, 424, 442, 612, 621, 658
first-order, 340
second-order, 342
PM (phase modulation), 299
demodulation, 350
generation, 304
PMT (photomultiplier tube), 599, 607
PN (pseudonoise) sequences, 441, 583, 629, 656
Poisson distribution, 481
Polar binary signal, 547, 552, 555
Polar trinary signal, 553
Polarization, 277
Population [statistical], 475
Postdetection bandwidth, 616
Postmodulation filtering, 635
Power, 11, 36, 168
in AM, 242
in FM, 317
Power signal, 11

Power spectral density, 168, 188, 270, 470, 506, 515, 621
FSK, 615
M-ary PSK, 659
MSK, 647
PCM, 567
PSK, 621, 647
QAM, QPSK, 635, 647
Power spectrum, 37, 170
cumulative, 170
numerical, 515
PPM (pulse-position modulation), 391, 396, 524
Pre-alias filter, 132
Precoding, 574, 576
Predetection bandwidth, 616
Preemphasis, deemphasis, 355, 368, 732
Preset equalizer, 580
PRI (primary-rate interface), 430
PRK (phase-reversal keying),
demodulation, 623
generation, 619
Probability, 457, 486
Probability density function (pdf), 467, 486, 501
Probability distributions
binomial, 478
exponential, 469
gaussian, 482
Laplacian, 525
Poisson, 481
Rayleigh, 521
uniform, 477
Probability of error, 549, 552, 670, 674
APK, QPR, 665, 689
ASK, 610
comparisons, 626, 666
DPSK, 625
FSK, 615
M-ary signaling, 586, 652, 660, 665
QAM, QPSK, 637, 665
PSK, 620, 665
Processing gain, 633, 656
Propagation effects, 277
Protocol, 428
Pseudorandom sequences, 486
see also PN sequences
PSK (phase-shift keying), 320, 412, 626, 665, 670
demodulation, 619

generation, 623
Pulse
exponential, 11, 64, 84, 95
gaussian, 11, 64, 95, 126, 148
modulation, 371
rectangular, 11, 35, 64, 87, 91, 95
shaping, 387
$(\sin x)/x$, 35, 47, 64, 91, 95
$(\sin x)^2/x^2$, 64, 91, 95, 567
timing, 390, 395, 400
triangular, 64, 91, 95, 104
PWM (pulse-width modulation), 391, 398

Q (quadrature) signal, 635, 645, 663, 671, 715, 718
$Q(\cdot)$ (gaussian probability function), 484, 551, 746
QAM (quadrature AM), 634, 667, 727, 741
see also QPSK
QPR (quadrature partial-response) signaling, 665
QPSK (quaternary PSK), 635, 647, 660, 664, 667, 670
demodulation, 636
generation, 635
Quadrature, 260
component of noise, 269
multiplexing, 224, 714, 741
phase comparator, 337
Quantization, 402, 522, 542, 553
Quantization noise, 541, 543, 558, 563
Quasi-synchronous, 420
Quaternary baseband transmission, 584

Radar, 155, 254, 308 *n*, 396, 443 *n*, 658, 700
Radio horizon, 279
Radius-of-gyration bandwidth, 125
Raised-cosine, 140, 387
Ramp encoder, 398, 405
Random binary waveform, 508, 512
Random process, 499
Random signals, 12, 457
Random variables, 464
correlation between, 494
transformations of, 487
Raster [television], 709
Ratio detector, 330
Rayleigh criterion, 392 *n*
Rayleigh distribution, 521
Rayleigh's theorem, 88 *n*
RB (return-to-bias) binary representation, 409, 571
RC high-pass filter, 213, 328
RC low-pass filter, 39, 79, 115, 124, 194
Realizable (causal) system, 13, 433
Receiver, 3
dual-conversion, 252 *n*, 254
figure-of-merit, 206
heterodyne, 251
noise-equivalent model, 198
TRF, 251
superheterodyne, 251
Rectangular (rect) function, 86
Rectifier detector, 248
Rectifier modulator, 245
Redundancy, 404, 417
Refractive index, 279, 414
Repeaters, 375, 564
Repetition code, 418
Resistor, 190, 197
Resolution
frequency, 139, 171
time, 392, 395
Response, 8
impulse, 65
magnitude, 40, 115
phase, 115
steady-state, 23, 40
step, 110, 122
RF (radio frequency), 253, 332
Ring modulator, 229
Rise time, 123, 523
RL low-pass filter, 22, 195
Roll-off factor, 388
Root-mean-square (rms), 177, 471, 500
RZ (return-to-zero) binary representation, 403, 409, 568

Sa $(\cdot)$ (sine-over-argument) function, 35 *n*, 47, 64, 91, 95, 98, 567
Sample [statistical], 475
Sample-and-hold circuit, 381, 397
Sample function, 499, 515
Sample mean, 476, 677
Sampling probe, 446
Sampling theorem, 14, 53, 127, 372, 379
SAP (second audio program), 724
Saturation, 713
SAW (surface acoustic wave) filter, 118
SC (suppressed-carrier) modulation, 220
SCA (Subsidiary Communications Authorization), 236, 316
Scanning microwave radiometer, 218
Scanning spectrum analyzer, 49, 157, 238
SC-HDTV (spectrum-compatible HDTV), 726
Schwarz's inequality, 21, 186, 432, 557
Scrambler, 293, 443, 628
SECAM television system, 723
Selective fading, 279, 281, 295
Selectivity, of receiver, 251
Self-clocking, 392, 448
Serial MSK, 648, 686
Serial search, 631, 657
Serial/parallel converter, 634, 663, 688
Sgn $(\cdot)$ (signum function), 90
Shannon's theorem, 5, 536, 651
Shape factor, 122
Shift register, 440, 572, 593
Shot noise, 207
Si $(\cdot)$ (sine integral), 122, 696
Sideband, 221
in AM, 239
in FM, 311
Sideband filter, 256, 266
Sideband power
in AM, 242
in FM, 317
in PSK, 621
Sidelobes, 139
Signal
compression, 544, 733
constellation, 659, 663, 689
space, 16, 668
Signal-to-noise (S/N), 4, 178
AM, 271, 273
budget, 204
comparison of systems, 350, 668
degradation, 276
DM, 563
FM, 345, 353

Signal-to-noise *(cont.)*
FSK, 617, 653
PAM, 399
PCM, 558
PPM, PWM, 401
PSK, 622, 661
vs. bandwidth trade-offs, 4, 402, 541
Signals, 8, 14
analytic, 259
antipodal, 619, 670
carrier, 220
complex-valued, 8, 26, 130
energy, power, 10
orthogonal, 16, 24, 650, 670
periodic, 11, 33, 83
sinusoidal, 9
random, 12, 457
Significant sidebands (in FM), 313
Signum (sgn) function, 90
Simplex (SX) transmission, 3
Sinc (·) function, 47 *n*
Singularity functions, 59
Sky-noise temperature, 202
Sky wave, 278
Slope detector, 329, 332
Slope overload, 562
Slow-hop spread spectrum, 656
Small-signal suppression, 354
Snell's law, 414
Space (in signaling), 410, 608
Space diversity, 279
SPADE system, 691
Spectra, 44
continuous, 83
cumulative, 170
line, 45, 48, 54, 93, 172
power, 37, 170
Spectral density, 85, 718
energy, 164
power, 168
Spectrum analyzer
multichannel, 166
scanning, 49, 157, 238
Spectrum inversion, 264
Speech scrambler, 293
Spike noise (in FM), 352
Split-phase binary representation, 410, 568
Spreading modulation, 629
Square wave, 18, 35, 41, 93
Squaring loop, 623, 685
SSB (single-sideband) modulation, 256, 350, 705, 730
demodulation, 262, 266
generation, 256, 305
Staggered QPSK, 638 *n*
Standard deviation, 472, 501
Standard noise temperature, 198, 200
Start and stop bits, 419, 426
Stationary random process, 500, 504
Statistical averages, 471
Statistical independence, 462, 492, 498, 505, 533
Steady-state response, 23, 40
Step function, 62, 92
Step-index optical fiber, 414
Step response, 110, 123
Stereo
AM, 737
FM, 235, 286, 316, 338, 369, 737
TV, 724
Stuffing
bit, 420, 422
word, 426
Subcarriers, 249, 326, 718, 724
Sunspot activity, 527
Superframe; *see* Multiframe
Superheterodyne receiver, 252
Suppressed-carrier modulation, 220, 281
Sweep generator, 40
Switching systems, 430
SX (simplex) transmission, 3
Symbol
error, 587, 660
rate, 5, 535, 584, 634
timing recovery, 623
Symmetry, use of, 32, 139
Synchronization
matched filter, 438
multiplexer, 421, 425
PAM, 381
PCM, 407
PN sequences, 439, 442
PPM, PWM, 398
PSK, 623, 625, 639
spread spectrum, 630, 657
television, 711, 727
Synchronous detector, 223, 232, 262, 267, 269, 610
Synchronous transmission, 419, 426
Syndrome, 591
Synthesizer, 162, 237, 287, 655
System, 8, 12
causal [realizable], 13, 109, 433
linear time-invariant, 13, 38, 109
nonlinear, 12, 42, 231, 246, 321, 623

Tapped delay line, 116, 434, 561
TASI (time assignment speech interpolation) system, 427 *n*
TASO (Television Allocation Study Organization), 368
T-carrier telephony system, 408, 423
Tchebycheff inequality, 473 *n*
TDM (time-division multiplexing), 377, 384, 408, 628
asynchronous, 419, 426
quasi-synchronous, 420
TDMA (time-division multiple access), 421
Telemetry, 327, 355, 396
Telephone systems, 539, 598
multiplexers, 408, 423
parameter limits, 729
Television; *see* TV
THD (total harmonic distortion), 43
Thermal noise, 190, 197, 558
Threshold effect, 351, 354, 395, 401, 539
Throughput efficiency, 426
Time
averages, 177, 499
delay, 120, 523
domain, 23
duration, 124
resolution, 392, 395
Time-bandwidth product, 124
Time-invariant system, 13, 105
Timing jitter, 624
Tracking mode, 630, 657
Transfer function, 38 *n*; *see also* Frequency transfer function
Transform
Fourier, 84, 135, 141
Hilbert, 260
Transformation of random variables, 487

Transmission
  asynchronous, synchronous, 419, 427
  distortionless, 119
  full-duplex, 4
  half-duplex, 3
  of information, 2, 531
  of noise through systems, 189
  simplex, 3
Transmission efficiency (in AM), 243
Transmission medium; *see* Channel
Transmitter, 3, 705, 710
Transversal equalizer, 577
Transversal filter, 116, 434, 510, 570
TRF (tuned RF) receiver, 251
Triangular function
  aperiodic, 91, 95, 104, 112
  periodic, 35, 181
Trigonometric Fourier series, 30
  sawtooth wave, 33
  square wave, 19, 35
  triangular wave, 33
TV (television), 268, 291, 317, 538, 701, 708, 714

Unbalanced QPSK, 636 *n*
Unbiased estimator, 475, 516
Uncorrelated, 186, 494, 498, 505
Uniform distribution, 477
Union bound approximation, 689
Upper sideband, 221
UHF television frequency allocations, 703

Varactor diode, 322, 324, 370
Variance, 472, 501
Variate, 497; *see also* Random variables
VCO (voltage-controlled oscillator), 40, 236, 324, 336, 613, 617, 631, 658
Vector space, 14, 669
Venn diagram, 459, 463
VHF television frequency allocations, 702
Video (TV), 268, 709
VIR (vertical-interval reference) signal, 712
VITS (vertical-interval test signal), 712
VSB (vestigial sideband) modulation, 266, 350, 627, 667, 710, 731

Walsh functions, 22, 72 *n*
White noise, 188
White spectrum, 89
Wideband FM, 306
Wide-sense stationary random process, 505
Window functions, 139, 516
  Hamming, 160
  Hann, 140
  Parzen, 160
  raised-cosine, 140, 387
Woodward's theorem, 307 *n*
Word stuffing, 426, 452

Zenith stereo TV system, 724
Zero-crossings demodulator, 331
Zero-forcing equalizer, 578

Selected Fourier Transform Pairs

	$f(t)$	$F(\omega) = \mathcal{F}\{f(t)\}$		$f(t)$	$F(\omega) = \mathcal{F}\{f(t)\}$
1.	$e^{-at}u(t)$	$1/(a + j\omega)$	11.	$\cos \omega_0 t$	$\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$
2.	$te^{-at}u(t)$	$1/(a + j\omega)^2$	12.	$\sin \omega_0 t$	$-j\pi[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]$
3.	$e^{-a\|t\|}$	$2a/(a^2 + \omega^2)$	13.	$\mathrm{rect}(t/\tau)$	$\tau\,\mathrm{Sa}(\omega\tau/2)$
4.	$e^{-t^2/(2\sigma^2)}$	$\sigma\sqrt{2\pi}\,e^{-\sigma^2\omega^2/2}$	14.	$\frac{W}{2\pi}\mathrm{Sa}(Wt/2)$	$\mathrm{rect}(\omega/W)$
5.	$\mathrm{sgn}(t)$	$2/(j\omega)$	15.	$\frac{W}{\pi}\mathrm{Sa}(Wt)$	$\mathrm{rect}(\omega/2W)$
6.	$j/(\pi t)$	$\mathrm{sgn}(\omega)$	16.	$\Lambda(t/\tau)$	$\tau[\mathrm{Sa}(\omega\tau/2)]^2$
7.	$u(t)$	$\pi\delta(\omega) + 1/(j\omega)$	17.	$\frac{W}{2\pi}[\mathrm{Sa}(Wt/2)]^2$	$\Lambda(\omega/W)$
8.	$\delta(t)$	1	18.	$\cos(\pi t/\tau)\,\mathrm{rect}(t/\tau)$	$\frac{2\tau}{\pi}\frac{\cos(\omega\tau/2)}{1 - (\omega\tau/\pi)^2}$
9.	1	$2\pi\delta(\omega)$	19.	$\frac{2W}{\pi^2}\frac{\cos(Wt)}{1 - (2Wt/\pi)^2}$	$\cos[\pi\omega/(2W)]\,\mathrm{rect}[\omega/(2W)]$
10.	$e^{\pm j\omega_0 t}$	$2\pi\delta(\omega \mp \omega_0)$	20.	$\delta_T(t)$	$\omega_0\delta_{\omega_0}(\omega)$, where $\omega_0 = 2\pi/T$

Multiples and Submultiples of Units

Prefix	*Symbol*	*Factor*	*Prefix*	*Symbol*	*Factor*
milli	m	10^{-3}	kilo	k	10^3
micro	μ	10^{-6}	mega	M	10^6
nano	n	10^{-9}	giga	G	10^9
pico	p	10^{-12}	tera	T	10^{12}
femto	f	10^{-15}	peta	P	10^{15}

Some Fourier Transforms Corresponding to Given Mathematical Operations

Operation	$f(t)$	$\leftrightarrow$	$F(\omega)$
Linearity (superposition)	$a_1 f_1(t) + a_2 f_2(t)$		$a_1 F_1(\omega) + a_2 F_2(\omega)$
Complex conjugate	$f^*(t)$		$F^*(-\omega)$
Scaling	$f(\alpha t)$		$\frac{1}{\lvert\alpha\rvert} F\left(\frac{\omega}{\alpha}\right)$
Delay	$f(t - t_0)$		$e^{-j\omega t_0} F(\omega)$
Frequency translation	$e^{j\omega_0 t} f(t)$		$F(\omega - \omega_0)$
Amplitude modulation	$f(t) \cos \omega_0 t$		$\frac{1}{2} F(\omega + \omega_0) + \frac{1}{2} F(\omega - \omega_0)$
Time convolution	$\int_{-\infty}^{\infty} f_1(\tau) f_2(t - \tau)\, d\tau$		$F_1(\omega) F_2(\omega)$
Frequency convolution	$f_1(t) f_2(t)$		$\frac{1}{2\pi} \int_{-\infty}^{\infty} F_1(u) F_2(\omega - u)\, du$
Duality: time-frequency	$F(t)$		$2\pi f(-\omega)$
Symmetry: even-odd	$f_e(t)$		$F_e(\omega)$ [real]
	$f_o(t)$		$F_o(\omega)$ [imaginary]
Time differentiation	$\frac{d}{dt} f(t)$		$j\omega F(\omega)$
Time integration	$\int_{-\infty}^{t} f(\tau)\, d\tau$		$\frac{1}{j\omega} F(\omega) + \pi F(0)\, \delta(\omega)$, where $F(0) = \int_{-\infty}^{\infty} f(t)\, dt$

Gaussian Error Probabilities

$Q(x)$	x
10^{-3}	3.09
10^{-4}	3.72
10^{-5}	4.27
10^{-6}	4.75
10^{-7}	5.20
10^{-8}	5.61

90000
9 780201 184983
ISBN 0-201-18498-2